PROCEEDINGS OF SCIENCE

Introduction to Loop Quantum Gravity

Simone Mercuri*
Institute for Gravitation and the Cosmos,
The Pennsylvania State University,
Physics Department, Whitmore Lab, PMB 67, University Park, PA 16802, USA
E-mail: mercuri@gravity.psu.edu

The questions I have been asked during the 5th International School on Field Theory and Gravitation, have induced me to give an account of the premises that I consider important for a beginner's approach to Loop Quantum Gravity. After a description of some general arguments and an introduction to the canonical theory of gravity, I review the background independent approach to quantum gravity, giving only a brief survey of Loop Quantum Gravity.

5th International School on Field Theory and Gravitation
April 20-24, 2009
Cuiabá city, Brazil

*Speaker.

PoS(ISFTG)016

1. Preface

Loop Quantum Gravity (LQG) is a background independent and mathematically rigorous canonical quantization of the gravitational field. The Organizers of the 5th International School on Field Theory and Gravitation have asked me to give an understandable account of the techniques and main results of this theory and I have been pleased to fulfill their requests. Immediately, I realized that the ideas of LQG were capturing the interest of many students, who asked me a lot of clarifications about some mathematical tools and physical aspects of background independent theories. I honestly think to have answered their questions and clarified many details, but, at the same time, I had the impression that many of them were getting confused by the canonical formulation of gravity and by the Dirac quantization procedure.

My impression has been that, generally speaking, the study of canonical quantization of gauge theories and especially gravity presupposes the knowledge of some arguments which often are not treated in details in the basic courses. I had the same experience when I started working on quantum gravity during my PhD. Usually, the approach to General Relativity (GR) proposed in the basic courses is based on the Lagrangian mechanics and the canonical formulation of the theory is usually only marginally described, without deepen into the general pictures of constrained theories. On the one hand, this is understandable from the perspective of academic and practical purposes, but, on the other hand, the Hamiltonian, or better, the canonical formulation of gauge theories and gravity remain obscure arguments among students. This is a common problem, which cannot be neglected when one attempts to describe the problem of quantum gravity and, more specifically, LQG.

Therefore, motivated by the questions asked me during coffee breaks and lunches, I have decided to slightly shift the focus of this proceeding, giving to the readers the opportunity to begin with quantum gravity from what I consider its natural starting point, namely the physical formulation of the problem.

So Section 2 is entirely devoted to describe some simple motivations which induce to formulate a quantum theory of gravity, digressing on the physical implications of such a theory on the existing concepts of space and time. Section 3 contains some preliminary arguments, which I consider as fundamental to understand the following discussion. They are quite simple and well known arguments, nevertheless they have to be necessarily clear before going to face the canonical theory of gravity and the Ashtekar-Barbero formulation of GR. So, I collected here the material I consider important, using a simple language and neglecting many complicated details, in order to give to the readers an easily accessible account of the main definitions and concepts used throughout the paper. In Section 4 I describe the canonical formulation of GR, starting from the 3+1 splitting of space-time. The mathematical procedure that allows to write the Hamiltonian Einstein equations is described in details and the Section concludes with a description of the initial value problem in gravity. Section 5 is dedicated to the connections formulation of canonical GR, better known as Ashtekar–Barbero (AB) formulation of gravity. This argument is particularly interesting in view of quantization, because by using the so-called AB connections, the constraints of GR can be rewritten in a more suitable form for quantization. Interestingly enough, the use of AB connections introduces a quantum ambiguity known as Barbero–Immirzi (BI) parameter, which affects the eigenvalues of geometrical quantum operators. The physical interpretation of the BI

parameter and its correlation with the topological sector of the theory is at present an argument of active discussion. Recently, the idea that this parameter is in fact a field has attracted the interest of researchers (me included!). I think that this could represent an interesting bridge with particle physics and could have consequences in Cosmology that deserve to be studied. For these reasons I invite the readers to refer to the original papers to get more information. Finally, in Section 6 I face the problem of quantization, starting from a brief description of the Dirac procedure and the Wheeler-DeWitt (WDW) equation. The last part of the Section is dedicated to the description of the main ideas of LQG, without entering in the complicated details of the theory. As I have explained before, in this paper I prefer to focus on introductory arguments to LQG, more than on the theory itself, beautifully described in many books and reviews, written by the major experts in this field [1, 2, 3]. For this reason I refer the readers to the standard Literature on LQG, with the hope that this paper could help them to face the argument more confidently.

At the end, I also added two appendices, one on differential forms, which are commonly used in Literature, but sometimes little known among students and one about the topological sector of gauge theories, in order to clarify some concepts which apply in canonical quantum gravity as well.

Throughout the paper, I will use an extremely simple approach, sometimes neglecting some interesting but slightly involved details. This obviously will affect the completeness and the rigor of the discussion, but, I am sure, will be appreciated by the students of the International School of Field Theory and Gravitation, who can find in this paper a usefully simple description of many arguments. My hope is to give them the possibility to get a "first order understanding" of the main concepts of canonical quantum gravity, without being discouraged by the rigorous mathematical formulation of the problem. I strongly suggest interested readers to delve into the "higher order descriptions" of the more complete and rigorous books and reviews cited above.

2. What is Quantum Gravity?

The present knowledge in Physics is a result of the new and revolutionary ideas born in the last century, which later led to the formulation of the two major physical theories describing the four interactions: Quantum Mechanics (QM) and General Relativity (GR). They have, on the one hand, opened the way to a great number of scientific discoveries and technical developments, but, on the other hand, they destroyed the coherence of prerelativistic classical physics [2], because the basic assumptions of each one of the two theories are contradicted by the other. QM is formulated using a Newtonian absolute (fixed, non-dynamical) space-time, on the contrary GR describes the dynamics of space-time itself, which is no more an external set of clocks and rods, but a physical interacting field, namely the gravitational field. The basic physical lesson of GR is contained in the following simple sentence: *Geometry tells matter how to move; matter tells geometry how to curve*, which expresses in a very suggestive way the fact that the theory describes both the dynamics of space-time (or gravitational field) and the motion of the bodies subjected to the gravitational field. But it also contains the seeds of an issue, namely the separation of the physical world in matter and geometry.

This dichotomy in Physics, together with the fact that on one side we have the present description of (what is in general intended as matter) electromagnetic, weak and strong interactions, unified in the language of Quantum Field Theory (QFT) and on the other side gravity (or geometry)

described by the pure classical theory of General Relativity, create a sort of "scientific discomfort" [4]. This is not only a philosophical problem, but assumes the distinguishing features of a real scientific problem as soon as one considers measurements in which both quantum and gravitational effects cannot be neglected. In fact, QM and GR are hugely successful in their own range of applicability, but they seem to be subjected to a sort of "reciprocal exclusion principle". In particular QM describes microscopic phenomena involving fundamental particles, ignoring completely gravity, while GR describes macroscopic systems, whose quantum properties are in general (safely) neglected. Of course there is still no experimental evidence on systems in which neither gravity nor quantum effects can be neglected, but we already know that our current theories are not able to describe such phenomena. This situation is usual in physics and it is in general the prelude to the formalization of a well stated scientific problem [4], in particular the goal of obtaining quantitative predictions about the outcomes of certain measurements on extremely energetic gravitational systems is often referred as the *Quantum Gravity Problem*.

It is clear from what said above that new elements could be necessary in order to make our current theories able to face a certain class of physical phenomena. Then, it is natural to wonder whether these new elements affect low energy processes. In other words, should we expect, even at low energies, small QG corrections to the predictions of our current theories? It is worth stressing that even a little deviation from the predictions of standard physics found in the experimental data of current and future experiments, the presence of which could be traced back to QG effects, would have an enormous impact on the research. We recall as an example the Lamb shift, which motivated and stimulated the studies about QED. From this perspective, it could be important to answer to the following question: *How far we are from an experimental evidence of a QG effect.* We cannot give a completely satisfying answer to it, nevertheless we can use a dimensional argument to state that the new effects should modify the usual predictions with additional terms proportional to the factor $(E/E_{Pl})^n$, where E is the typical energy scale of the experiments, E_{Pl} is the Planck energy ($E_{Pl} \equiv (G_N)^{-1/2} \approx 10^{28} eV$), while n is a positive integer number. At this point one may be surprised by the compelling necessity to quantize gravity felt by physicists, since at the LHC, the most powerful accelerator ever projected, we can reach a very small energy if compared to the Planck scale. Specifically, the ratio between the Planck and the reachable energy is of the order of $E_{LHC}/E_{Pl} \approx 10^{-15}$, that means we are fifteen orders of magnitude below the scale at which we expect to see the quantum effects of the gravitational field. Even though this fact is true, it is absolutely false that this is a good reason for abandoning the program of constructing a consistent theory of QG. The motivations are connected with the fact that there exist in Nature particles of energies much larger than those we can produce with the accelerators, moreover during its evolution our Universe experimented regimes in which the energy available was (most likely) even larger of the Planck scale. Furthermore, even though the factor E/E_{Pl} is extremely small at the present available energies and, consequently, the QG effects cannot be directly experimented, it exists the concrete possibility that some astrophysical phenomena can behave as magnifying glasses being able to make them visible in near future [5, 6]. In other words, the fact that QG effects are expected to be very tiny does not mean they are absolutely untestable, clearly as one should expect the opportunities to make such tests are rather rare. We also emphasize that it could be not necessary to reach the Planck energy to see some QG effects. In this respect, we recall that a class of extremely energetic phenomena called *Gamma Ray Bursts* (GRB) could represent a really

important laboratory to test QG predictions, in fact they seem to be the natural candidates to verify whether the fundamental hypothesis about a discrete structure of space-time will be confirmed by experiments. The peculiar features which make GRB relevant for QG is the extremely wide range of the emitted energies and the cosmological distance of the explosive events.

Concluding, as usual in Physics, from the pure empirical point of view, the new framework possibly introduced by a consistent and complete QG theory could represent a tiny deviation from what we already know, being only a further small modification of those empirical laws which give us a pictorial description of how Nature works. But from the theoretical point of view it could represent not only the completion of that revolution QM and GR introduced in the last century, but also open the way to the discovery of complete new aspects of Nature, exactly what QFT and GR have been doing during the last one hundred years. However, in spite of their great empirical success, QM and GR have left us with a fragmented understanding of the physical world, this requires a new synthesis, which is a major challenge in today's fundamental Physics [2].

In this sense, Quantum Gravity can solve the dichotomy present in our current understanding of physical phenomena and, moreover, it could give us predictions on those regimes in which the quantum and gravity effects merge.

2.1 Why we need a Quantum Theory of Gravity?

Above we introduced the so called QG problem, which gains the status of a true scientific problem as soon as one considers physical systems in which both the gravitational and the quantum mechanical effects play an important role. Moreover, the failure of the existing theories as soon as we push them near their extreme margins of applicability suggests to quantize gravity. We also digressed on the empirical content of such a problem, affirming that obtaining an experimental evidence of a QG effect is extremely complicated, but not impossible; even though we expect a very tiny modification of the existing laws describing the physical systems. Therefore, as often occurs, one may object to the above pleaded motivations, saying that they are mainly suggests by philosophical reasons. Namely, the hope of finding the way to conciliate the basic assumptions of two very different and complementary theories is not really required by a scientific problem, since this attempt has very little to share with Physics, simply because the physical effects are so tiny to be actually undetectable. It could be really so!

For this reason, in this Section three well stated problems, which regards respectively QFT, GR and the merging point of QFT and GR, are discussed in order to emphasize that QG is a true physical problem. It, in fact, can provide information about the behavior of fundamental gravitating quantum system as, for example, a system of gravitating fermions or extremely energetic scattering processes or the Universe itself in its initial expansion.

The questions we present below are generally connected with the fundamental structures of the theories and mainly concern the problem of singularities. In fact, it is worth recalling that the theory describing the gravitational interaction fails in giving a fully satisfactory description of the observed Universe [7]. GR, indeed, leads inevitably to singularities of space-time as a number of theorems mainly due to Hawking and Penrose demonstrate. The singularities occur both at the beginning of the expansion of our Universe and in the collapse of gravitating objects to form Black Holes [8]. Classical GR breaks completely down at these singularities, or rather it results to be an incomplete theory, because it neither gives a description of the singularities themselves, nor

provides the boundary conditions for fields in the singular points. The appearance of singularities in extremal situations reflects both on QFT and on GR itself, generating a more subtle question when we trivially try to merge the two theories as we are going to argue.

This failure of current theories of course represents a good reason to pose a scientific problem and its solution is widely believed to be in the formulation of a consistent and complete theory of QG.

It remains to treat in a more systematic way the problem regarding the empirical content of a QG theory. This argument is postponed at the end of the Section, where we will trace the way to get information about QG from the astrophysical phenomena named Gamma Ray Bursts, motivating the belief that the necessary experimental evidences which could support the QG research are not so far.

2.1.1 Planck scale collisions

The question is:[1] What would happen if we managed to collide an electron-positron pair of energy *per particle* of $10^{28} eV$? We are unable to give an answer to this question, because the energy in the center of mass is greater than the Planck energy. So, what happens in the center of mass of the two particles is completely out of our understanding. But, according to our present physical theories describing the collisions between fundamental particles, there should not be anything peculiar in the setup of such an experiment. However, the same hugely successful theories are not able to provide us with a consistent prediction for the outcome of this experiment. The reason of this failure is strictly bound with the fact that in such a collision we cannot neglect the gravitational properties of the involved particles at the moment of the collision. But, we do not have any scientific information on how taking into account such an effect in the framework of QFT. In other words when the gravitational field is so intense that space-time geometry evolves on a very short time scale, QFT cannot be consistently applied any longer; or, from another perspective, we can say that when the gravitational effects are so strong to produce the emergence of space-time singularities, field theory falls into troubles. Summarizing, we are able to extract numbers (predictions) from QFT when the curved space-time is static (or slowly varying) and non-singular, but we are not able to handle situations in which the gravitational field is so intense to give rise to a fast varying and singular space-time [4].

The incompatibility between QM and GR in treating the proposed scattering problem can be further analyzed. In this respect, we have to remember that GR governs consistently the space-time and particles dynamics, in particular, given the Lagrangian for the matter, once the Einstein equations have been solved, we can predict the trajectories of the particles. But in the framework of QFT particles are asymptotic states of quantum operators and during the collision they do not follow any classical trajectory. The whole dynamics of the collision is contained in the S-matrix, which gives the evolution from the initial ($|in\rangle$) to the final ($|out\rangle$) state. During the interaction, the intermediate state is a pure quantum superposition of all the possible states compatible with the quantum numbers of the initial state; namely, we can associate to particles a semi-classical (fuzzy) trajectory only asymptotically, in other words much before or much later the collision. One could try to apply the formalism of GR to the formally classical trajectories contained in the path-integral,

[1] Some of the ideas presented below are extracted from [4] and [2].

but the problem remains ill-defined and, generally speaking, it will be affected by divergences if the energy of the particle are sufficiently high to generate a significant geometrodynamics. This fact can be pictorially described. The path-integral formulation of QM consists in summing on all the possible trajectories connecting the initial and final quantum state, with a weight proportional to the exponential of the classical action. The major role in the sum is played by those trajectories nearby the classical one, because the contribution of the far away ones is suppressed by the weight factor. If the gravitational field is weak, we can safely approximate the space-time nearby the classical trajectory with the Minkowski flat space-time and assuming that the curvature does not affect the trajectories which enter in the sum. But, if the gravitational field is intense, then also trajectories close enough to the classical one are affected by the curvature of the space-time and this effect should be taken into account in the path-integral sum. The case of fast varying or singular gravitational field is even worst, because in that case the trajectories could fall into a singularity or oscillate very fast, breaking down completely the formalism, by introducing remarkable and uncontrollable effects into the sum.

2.1.2 Singularities

The study of singularities in GR is an absolutely fascinating argument. Here we give an extremely brief account of this huge argument, which represents one of the crucial point suggesting that a quantum theory of gravity is, in fact, necessary. Near singularities classical GR becomes inconsistent and incomplete, as we already stressed before and, differently from Newtonian gravity, they represent an inevitable feature of the theory. Indeed, in Newtonian gravity the $r = 0$ singularity, appearing in the complete collapse (namely, when all the matter reach simultaneously the origin) of a spherical non-rotating shell of dust, can be easily avoided slightly perturbing the spherical symmetry of the collapsing shell, for example giving it a little rotation. On the contrary, in GR the singular behavior of space-time cannot be avoided. All the solutions we have of Einstein equations show a singular behavior. But, since all of them are characterized by symmetries, one may think that, as it happens in Newtonian gravity, the relaxation of symmetries could allow to avoid singular points. But the Hawking, Penrose *et al.* theorems demonstrate that this is not the case. Wald says [9]:

> Although the singularity theorems do not prove that the singularities of classical General Relativity must involve unbounded large curvature, they strong suggest the occurrence in Cosmology and gravitational collapse of conditions in which quantum or other effects which invalidate classical General Relativity will play a dominant role.

So the singularity theorems do not use the natural and, in a certain sense, more physical notion of unbounded density to characterize the singularities of space-time, but the characterization of singularities is based on the notion of incompleteness of geodesics, which, however, contains some unwanted features.

The necessity to derive from such a notion the presence of a singularity of space-time is due to the necessity of a diffeomorphisms invariant criterion. In this respect, it is worth noting that the singular points in $r = 0$ of Schwarzschild or Robertson-Walker metrics, rigorously speaking, are not point of those space-times, otherwise the metrics would not be well defined everywhere

on the manifold M. Moreover, the criterion based on the bad behavior of higher order scalars constructed with the curvature tensor does not work in some cases, so it is clear that a more satisfactory definition has to be introduced.

So far, the criterion of geodesics incompleteness seems to be the most appropriate. Its physical meaning is suggested directly by its definition, indeed a geodesic is said to be incomplete when it is inextendible in at least one direction; namely, it has a finite range for the affine parameter. As a consequence, a particle falling along a time-like or null inextendible geodesic will end its existence within a finite proper time (or it began its existence a finite proper time ago). So, even though a completely satisfying notion of singularity lacks [9], we will call physically singular all those space-time having at least one incomplete geodesic.

At this point, the following famous theorem by Hawking and Penrose can be enunciated:[2]

Theorem 2.1. Singularity theorem of Hawking and Penrose (1970). *Let us suppose that the space-time $(M, g_{\mu\nu})$ satisfies the following four hypothesis:*

1. *$R_{\mu\nu}u^\mu u^\nu \geq 0$ for all time-like or null u^μ;*
2. *it exists at least one point for every time-like or null geodesics at which $R_{\mu\nu}u^\mu u^\nu \neq 0$;*
3. *no closed time-like curve exist;*
4. *at lest one of the following three conditions holds:*
   i) $(M, g_{\mu\nu})$ possesses a compact achronal set without edge, i.e. $(M, g_{\mu\nu})$ is a closed Universe,
   ii) $(M, g_{\mu\nu})$ possesses a trapped surface,
   iii) there exists a point $p \in M$ such that the expansion of the future (or past) directed null geodesics emanating from p becomes negative along each geodesic in this congruence.

Then $(M, g_{\mu\nu})$ must contain at least one incomplete time-like or null geodesic.

The first three conditions of the above theorem are believed to be satisfied in our Universe. The first one in particular can be simply showed to be shared by all those space-times satisfying the Einstein equations and the *strong energy condition*,[3] which seems to be plausible for ordinary matter.

Finally, we can conclude that there are strong evidences on the fact that our Universe is singular; of course we cannot know by which kind of singularity it is characterized, because the above theorem do not give us any insight on this question.[4] Nevertheless, it suggests that an extension toward QG is necessary. It, in fact, demonstrates that the universally accepted theory of gravity cannot definitely give us a complete and consistent description of the evolution of our Universe.

2.1.3 Merging General Relativity and Quantum Field Theory

After having described the problem of singularities in GR, we want to address a number of difficulties coming out when one attempts to quantize gravity using the usual formalism of QFT. The

[2] See § 3.1 for definitions.

[3] Namely for every time-like u^μ we have $T_{\mu\nu}u^\mu u^\nu \geq -T/2$, where $T_{\mu\nu}$ is the energy-momentum tensor of the matter.

[4] The particular kind of singularity we obtain from the Einstein equations depends in general on the particular symmetries of the model.

main source of problems in extending the formalism of QFT to gravity is represented by the double role played by the metric tensor. It, in fact, represents both the dynamical field describing gravity and the tensor describing the causal structure of the background. The nature of these difficulties is not only conceptual, as for example the consideration that a quantum theory of gravity would imply a quantum, namely discrete, structure of space-time itself. But, more practically, they are correlated to the profound difference existing between GR and other classical field theories: while in the latter the background is always considered as an external and fixed structure, the former is the theory describing the dynamics of the background itself. Therefore we cannot assume a given structure of space-time *ab initio*, but we have to invent a formalism which allows to quantize a classical theory in a background independent way. This fact makes GR so peculiar that so far all the attempts to quantize gravity have encountered fundamental difficulties; only in the last years many of these obstacles have been overcome, leading to a consistent QG [1, 2, 3].

As is well known the construction of a QFT on Minkowski space-time of a free or perturbatively interacting field is the only procedure we can control. In particular, given a small number of axioms, the Wightman axioms, we can construct a consistent QFT. Let us begin the description of the issues one would find in applying the usual formalism of QFT to gravity by describing a very simple and well known example, which best illustrates the dichotomy existing in the metric tensor.

One of the Wightman axioms contains the notion of micro-causality. In order to introduce this concept let me consider a scalar field represented by the smeared operator-valued distribution

$$\Phi(f) = \int_{\mathbb{R}^{n+1}} d^{n+1}x \, \Phi(x) f(x) \,, \tag{2.1}$$

where f is a test function of rapid decrease. Suppose now that the supports of the test functions f and f' are space-like separated, then the micro-causality assumption is equivalent to require that:

$$\left[\Phi(f), \Phi\left(f'\right)\right] = 0 \,. \tag{2.2}$$

Physically, the above condition assures that a measurement of the field Φ in the region of space-time contained in the support of the function f cannot be influenced by the measure of the same field in the region contained in the support of the function f'.

Now, the gravitational interaction is described by a self-interacting spin-2 field, it is natural to expect in analogy with the previous example that the following commutation relation holds

$$\left[g(f), g\left(f'\right)\right] = 0 \,, \tag{2.3}$$

where f and f' are two tensorial test functions with compact supports separated by a space-like distance. But, strictly speaking, the above relation makes no sense. The reason being that we cannot give a consistent meaning to the request that the supports of the test functions be space-like separated, unless we know the state of the gravitational field, namely the metric tensor. But, the commutation relation in line (2.3) must hold independently of the state of the gravitational field.

It is clear that the micro-causality condition is the quantum translation of the pure classical concept of causality. Even though it contains information about measurements on quantum field, its basic structure is, however, founded upon the existence of a Minkowski space-time, with its fixed light cone. In a quantum theory of gravity the causality condition and, as a consequence the

micro-causality axiom, is inextricably bound with the quantum dynamics of the gravitational field, so it has to be, at least, modified or completely replaced by a new requirement. In general, the other Wightman axioms are violated in gravity as well. The reason being that the fundamental objects one has to postulate in order to write down the axioms, namely i) a differentiable manifold $\mathbb{R}^{n+1}$, on which is defined a non-dynamical Minkowski metric η, with its fixed future and past causal light cone $J^+ \cup J^-$, ii) a symmetry group, which, in the Minkowski space-time, is the ten parameters Poincaré group $\mathscr{P}$, with its associated infinite dimensional representation acting on the quantum states $U(\mathscr{P})$ and iii) an invariant vacuum state $|\Omega\rangle$, cannot be properly postulated. For a general space-time, in fact, we do not have a symmetry group and a unique invariant vacuum state. Consequently, it neither exist any obvious generalization of the Wightman axioms nor we can rigorously define any Fock-Hilbert space for the quantum states of the theory. Therefore, the whole formalism falls into troubles.

A possible solution to these issues is the splitting of the metric tensor as suggested by a perturbative theory:

$$g_{\mu\nu} = \eta_{\mu\nu} + \gamma_{\mu\nu}\,. \tag{2.4}$$

The basic assumption here is that $\gamma_{\mu\nu}$ represents the dynamical variable describing a self-interacting spin-2 field, while $\eta_{\mu\nu}$ describes the background metric, which, in general, could be any solution of the classical Einstein equations. This method is of course mathematically correct, provides interesting insights on the quantization of a spin-2 field on a fixed (in general curved) space-time and could be useful to describe the interaction between gravitons and matter or the gravitational waves. Nonetheless, it cannot be considered as a good starting point for a complete quantum theory of gravity, because the metric separation in line (2.4) destroys the full general covariance of the classical theory, namely its main constructing principle. More practically, the infinite perturbative series becomes meaningless if the fluctuations become large, in other words, GR is, in general, a non-renormalizable theory. One can hope that the perturbation theory could be actually finite via a possible (magic) cancellation of the divergences, but this hope is unjustified [11]. This means that the resulting theory cannot predict any physical result. We still can advance the hypothesis that the super-symmetric extension of this theory has a chance to be a finite theory, because, as is well known, the super-symmetric extension of a classical field theory features, in general, an higher degree of ultraviolet convergence due to fermionic cancellations [12]. But again the resulting theory is non-renormalizable, even worst neither the eleven dimensional super-gravity theory shows any hoped cancellation property [13]. Then, although we do not have a complete proof of the failure of the QFT formalism in the case of gravity, it is widely believed that the perturbative approach does not provide a completely consistent answer to the problem of QG .

All these issues have been sometimes pushed to the extreme consequences by some authors, who argued that classical GR is correct at the fundamental level. This position is, however, untenable for at least two reasons, the first one is connected with the inevitability of singularities in classical GR; the second one regards the interaction of gravity with quantum matter systems, which is a source of troubles as we are going to show.

The question we want to answer is: what is the curvature of space-time associated to a given quantum state of the matter? Let us suppose that the classical Einstein equation holds at the funda-

mental level, then the most natural candidate to get an answer to this question is:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \left\langle \hat{T}_{\mu\nu} \right\rangle . \tag{2.5}$$

Where in the right hand side we have put the expectation value of the energy-momentum tensor of the matter in a given quantum state. Now suppose that the quantum state of the matter is such that we have probability $1/2$ for the localization of all the matter in a region of space-time denoted as U_1 and the same probability for the localization in another region U_2, disjoint from the region U_1. In other words, we are in the following situation:

$$|\text{matter}\rangle = \frac{1}{\sqrt{2}} |\text{all matter in } U_1\rangle + \frac{e^{i\theta}}{\sqrt{2}} |\text{all matter in } U_2\rangle , \tag{2.6}$$

where U_1 and U_2 are disjoint regions of space-time.

In this physical situation, the gravitational field, according to equation (2.5), would behave like half of the matter were in U_1 and the other half in U_2. Now, if we resolve the quantum state by measuring the position of matter, we will find all the matter either in U_1 or in U_2. Then, the gravitational field should modify in a discontinuous acausal manner, leading to serious difficulties [9]. It is widely accepted the idea that this problem can find its consistent description in the framework of a quantum theory of gravity, which should be able to give an answer to the following question: how does a quantum particle modify space-time? Which is equivalent to the previous question and indirectly contains the subtle problem about the interaction of the quanta of the gravitational field with matter.

So, QG seems to be necessary as soon as we consider the interaction of quantum matter with the gravitational field, but, in trying to quantize gravity as a usual QFT, we have to face a large number of conceptual and technical issues. This is a long standing problem, nevertheless, it does not contain any indication about a fundamental incompatibility between QM and GR principles. In fact, as stressed by Rovelli [2], it is important to distinguish between QM, which is a general mechanical theory and QFT which can be considered as a particular application of the laws of QM to a system with an infinite number of degrees of freedom. As we said above GR is incompatible with the formalism of QFT for the non-existence of a fixed background structure, but this does not means that it is incompatible with QM [14]. So the right question one should pose is the following one: is it possible to construct a quantum theory of a system with an infinite number of degrees of freedom without assuming a fixed background causal structure? The answer is: yes! It is possible as the modern background independent theories of quantum General Relativity demonstrate, but more insights are necessary in order to better understand many issues, in particular those connected with measurements and the consistent introduction of measuring devices in the framework of such theories.

2.2 Space-time, background independence and relationalism in Physics

The question we want to discuss here represents another important "open issue" of the research in QG and regards the fate of classical space-time.

A very interesting feature of GR and QFT is that both are compatible with a classical description of space-time, although they do not share the same construction of such a classical space-time.

According to QFT the strategy to sharply localize a point in space-time requires a limiting procedure on the mass of the devices localizing points; in GR, instead, the localization of a point is a background independent procedure based on the crossing of two geodesics. We will describe in more details these contrasting features in what follows.

The concept of classical space-time is appropriate in Physics as long as the proposed theories allows to localize space-time points sharply. A fundamental requirement for a consistent physical theory is the agreement between the limits on possible measurements as established by the formalism and the in-principle limits imposed by the physical measurements procedure. A possible example which best illustrates this important point can be easily constructed by considering the measurement of the angular momentum vector. In classical physics, the angular momentum of the particle could be sharply measured; in other words, we can, in principle, measure each of the components of the angular momentum vector of a particle simultaneously and with an infinite precision. But, modern physics has radically changed our perspective by imposing some limitations to the measurements allowed by the experimental procedures, namely, we cannot measure all the components of the angular momentum simultaneously. As a consequence, in the description of such a system, the angular momentum has to be described by a non-classical formalism, which incorporates the experimental limitations found.

This aspect deserves to be further discussed and clarified, because it is correlated with the fate of the classical concept of space-time as described below. Fortunately, a good example exists in the history of physics, which involved authoritative physicists as Einstein, Bohr, Landau and Rosenfeld. A lively and fruitful debate, in fact, animated the scientific community during the period immediately after the birth of quantum electrodynamics. The matter of the discussion regarded the measurement of the electromagnetic field in the framework of quantum electrodynamics and led to the formulation of the so-called *Bohr-Rosenfeld criteria* for a consistent theory. Before discussing this matter, we want to focus the attention of the reader on the importance of these kind of arguments for theoretical physics. A simple example can clarify this point: In this respect, we want to stress that it was the study of the synchronization of distant clocks assuming an absolute maximum velocity for signals that led Einstein to special relativity.

2.2.1 Bronstein objection and the fate of space-time in QG

Here, we want to briefly describe the main points of the debate come out in 30s, when a group of really distinguished physicists argued that there was a conceptual disagreement between the uncertainty limits predicted by quantum electrodynamics on a certain class of observations and the mathematical formalism that the same theory adopts for describing these measurement results. Specifically, the key point of the debate was a physical consideration due to Landau and Peierls. They argued that, according to quantum electrodynamics, the electric field in a generic point P can be measured sharply, namely with zero uncertainty, consequently a measurement procedure allowing to measure it sharply must exist. Obviously, if the situation were different then the theory would be inconsistent. Eventually, the conclusion of Landau and Peierls was that quantum electrodynamics must be rejected as physical theory, because such a zero uncertainty measurement procedure is not possible in Nature, so quantum electrodynamics is inconsistent.

The conclusion of Landau and Peierls was motivated by the following consideration: The measurement procedure of the electric field in a generic point can be ideally performed by using an

electrically charged probe, which undergoes the effect of the electric field and accelerates, the outcome consists in the experimental measurement of this acceleration. In principle one can measure the mean value of the electric field in a small region of space, approaching the point asymptotically. Classically, this procedure is perfectly consistent, but, quantum mechanically, it is affected by the Heisenberg principle. In fact, the simultaneous measurements of the localization of the probe, say ΔX and of the variation of its momentum due to the effect of the electric field, say ΔP must to satisfy the following uncertainty relation $\Delta X \Delta P \geq \hbar$. Moreover, the acceleration of the probe introduces another problem, related to the fact that an accelerated charge emits energy. So, the measurement affects itself by modifying the momentum of the probe, namely the outcome of the experiment.

Such an analysis leads to the conclusion that a sharp measurement of the electric field is possible only in a very special situation, namely when the charge and the mass of the probe can be adjusted in such a way that, by a limiting procedure, it is possible to reduce the ratio between the electric charge density and the mass density of the probe to zero.

But, as noted by Landau and Peierls, the above requirement cannot be fulfilled, because, even though it exists in Nature a great variety of particles with different charge/mass ratios, no one can constitute the ideal probe with a charge/mass ratio equal to zero. As a consequence quantum electrodynamics must be rejected and a consistent alternative should be sought.

Bohr and Rosenfeld opposed to this viewpoint, claiming the consistency of quantum electrodynamics. The point is that the generations of particles existing in Nature are not a prediction of the theory, rather they are an outside input; in other words, given the particle content, the theory predicts their mutual interactions. In this sense, the failure pointed out by Landau and Peierls cannot be considered as an inconsistency of the theory, because it can be attributed to an external fact. Therefore, it does not affect the logical structure of the physical theory. If we found particles with a vanishing charge/mass ratio, then we would be able to measure sharply the electric field and quantum electrodynamics would not absolutely be in contrast with such a discovery.

The same argument applied to the gravitational field has a remarkable consequence, related to the fate of the common accepted notion of space-time. In fact, as Bronstein pointed out, the requirement that the charge/mass ratio vanishes for the probe used to measure the field cannot be applied to the gravitational field. The reason being that the gravitational counterpart of the electric charge is the gravitational mass, therefore the equivalence principle obliges the ratio between the "gravitational charge" and the inertial mass of the probe to be equal to one for any particle existing in Nature. Therefore, the equivalence principle seems to put a serious restriction on the possibility to sharply measure the gravitational field.

This fact suggests that an unavoidable fundamental limit on the measurements accuracy of the gravitational field exists, affecting, as a consequence, the fundamental structure of space-time. Furthermore, this could have important implications in the construction of QG, because, as is well known, QM imposes limitations on the simultaneous measurement of conjugate pairs of fields, but, following Bohr and Rosenfeld, no limitation on the accuracy of a measurement of one single field occurs. The Bronstein's argument suggests to consider the possibility that ordinary quantum mechanics could be inadequate to describe the quantum theory of a geometric field, in this sense either a modification of the uncertainty relation comes out naturally from the theory encoding this intrinsic limitation on the measurements of the gravitational field, or, more speculatively, a different quantum mechanics, which reduces to the ordinary one in the appropriate limit, should be sought

to face the problem of QG.

2.2.2 Space-time in classical and quantum mechanics

In accordance to what said above, it is worth stressing once more that QG could radically change the present concepts of space and time. It could be possible, in fact, that new and unpredictable non-trivial outcomes of the theory oblige us to abandon any intuitive representation of space and time. In this respect, it is important to understand how the notion of space and time changes according to the postulates of different physical theories. This is the motivation of the following discussion, where the different notions of space and time are briefly described.

In classical mechanics space-time is an external fixed and flat structure. In particular, space is an Euclidean three-dimensional stage for the physical phenomena, $\mathbb{R}^3$, while time is represented by the oriented one dimensional real axes, $\mathbb{R}$. Space and time can be measured with an infinite precision in classical mechanics. The common idea of space-time acquires a proper operative meaning as soon as we imagine the existence of a dense array of ideal synchronized clocks beating the flow of time, while perfect rods sharply measure the distances among clocks, giving in this way an empiric meaning to the points of space-time.

Quantum mechanics is characterized by a novelty with respect to its classical counterpart; we are referring to the uncertainty principle. This principle establishes the impossibility to simultaneously measure with an infinite precision a pairs of conjugate variables. Spatial coordinates are the conjugate variables to the momenta along the same axis and can be measured with an infinite precision at the price to lose any information on the velocities. Therefore, a subtle question about the evolution of the reference system arises. Indeed, as soon as we have realized that the role of time in quantum mechanics is identical to the one it plays in classical mechanics,[5] we can imagine to construct the same array of dense ideal clocks to give an operative meaning to the space and time of QM. But, if the clocks had a finite mass, one should worry about their evolution, in fact, once we measure their position we lose information on their evolution. This fact suggests the possibility that the space and time of quantum mechanics acquires an operative meaning only in the limit of infinite mass clocks. This does not create any embarrassment, because quantum mechanics, as stressed at the very beginning of this section, completely ignores gravity, then its logical consistency can rely on the idealization of a physical reference frame constituted by infinitely heavy particles.

The study of the space-time of quantum field theory introduces some features which deserve to be deepened. As clarified above the background of quantum mechanics is a classical space-time, in other words we can, with a limiting procedure, measure sharply the position of a particle in a space-time pictorially represented by a dense array of infinite mass clocks. In particular, in order to localize a finite mass particle, we consider an interaction between a probe and the particle. The accuracy of the measure is proportional to the inverse of the energy carried by the probe, namely, if the probe carries an energy, say, $1/\Delta X$ we can localize a particle interacting with the probe with an accuracy of ΔX. The reason being that the probe results to be confined to a region of space-time of size ΔX. This means that the sharp localization implies the injection into the

[5] With the difference that in some extrapolations of the theory an uncertainty relation between time and energy comes out. However, since time is not an observable in quantum mechanics but only an evolution parameter, the time-energy uncertainty relation must be interpreted differently with respect the coordinates-momenta one, which, in stead, is a consequence of the fact that coordinates and momenta are conjugate observables in QM.

physical system of a greater and greater (in principle infinite) energy. Even though this procedure does not create any problem in classical quantum mechanics, as soon as we consider a "relativistic quantum mechanics" a subtle issue crops up. In fact, even though the 4-dimensional picture does not modify the fundamental structure of space-time, still pictorially described by a dense array of extremely massive clocks, it introduces the equivalence between mass and energy, which creates a shortcoming in the above described measuring procedure. Indeed, as soon as the energy carried by the probe becomes higher than the rest mass of the particle being measured, many copies of the original particle are produced as a direct effect of the position measurement (injection of energy into the system).[6]

In order to avoid any misunderstanding, we want to stress that the sharp localization of an infinite mass particle is still possible, because in order to produce copies of the same measured particle an infinite energy is necessary, so, as in ordinary quantum mechanics, no problem exists in the localization of space-time points.

Concluding we can say that even though the space-time background of quantum field theory is the same ideal dense array of clocks that we have already introduced in ordinary quantum mechanics, in quantum field theory the accuracy of a measurement of the position of a particle is however limited, the limit being imposed by the measuring procedure itself. In other words, the coexistence of the Heisenberg uncertainty principle with Special Relativity prevents from obtaining a sharp measurements of the position of a finite mass particle, even though the concept of classical space-time is preserved. However, this appear to be no longer possible when gravity is present. On one side, in fact, quantum mechanics obliges to consider an infinite mass point particle in order to give sense to a sharp localization procedure, but, on the other side, this is incompatible with GR for obvious reasons and the entire sharp localization procedure falls into troubles.

2.2.3 Space-time in General Relativity

General Relativity is the theory describing the dynamics of space-time, in this sense even the shortest discussion implies a deepening into the basic principles and the main ideas which led Einstein to the formulation of his geometrical theory of the gravitational field. We are not referring to the construction of the Einstein dynamical equations, but to the more subtle and complicate "struggle with the meaning of the coordinates."

Generally speaking, in constructing the field equations of GR one has a number of hints as, for example, the fact that the static limit of the field equations must be the Newton law, or the fact that the mass of a particle is the source of the Newtonian gravitational field: Then, the mass being a form of energy, as Einstein himself clarified in special relativity, it is reasonable that the energy-momentum tensor be the source of the relativistic field equations. Finally, the concept that no privileged reference systems exist suggests that the equations have to be covariant under a class of general coordinates transformations. Einstein, some years before the publication of his most famous paper in 1915, learned that the only possible combination of second order partial derivatives of the gravitational field transforming covariantly under general coordinates transformations is the Riemann tensor. It became soon clear to his mind how the equations of motion of GR had to look

[6] From another perspective this is the reason why in place of relativistic quantum mechanics, a quantum field theory is required to describe such phenomena, we need, in fact, a theoretical framework which does not require to fix the number of particles of the system.

like. But the subtlest and most striking aspect of GR, which Einstein dedicated himself to for a long period, regards the philosophical content of the theory: The real novelty introduced by GR is that coordinates have no physical meaning independently from the value of the physical field and the trajectories of the physical particles.

The Equations of GR are, in fact, generally covariant, or, in other words, if $e_\mu{}^a(x)$ is a solution of the field equations, then, given the general coordinates transformation $y = y(x)$, also $e'_\nu{}^a(y)$,

$$e'_\nu{}^a(y(x))\,\frac{\partial y^\nu(x)}{\partial x^\mu} = e_\mu{}^a(x)\,, \tag{2.7}$$

is a solution of the field equations. Physically, this means that the physical laws are the same in all the reference frames, namely in all coordinates systems.

Now, in order to understand the meaning of general covariance, which will be useful for clarifying the structure of space-time of GR, let us consider a region of space-time, say $\mathscr{U}$, containing the event P, and, assigned in $\mathscr{U}$ the system of coordinates $\mathscr{X}$, let us indicate with x_P its coordinates. Let $e_\mu{}^a(x)$ be a solution of the generally covariant field equations and let us assume that

$$R|_P = R(x_P) = 0\,, \tag{2.8}$$

where $R(x)$ is the Ricci scalar. Suppose now that we decide to change our system of coordinates in the region $\mathscr{U}$. Specifically, be $\mathscr{Y}$ the new system of coordinates and $y = F(x)$ the transformation law from one system to the other. Thus $e'_\nu{}^a(y(x))$, obtained from $e_\mu{}^a(x)$ via the relation (2.7), is a solution of the equations of motion too. In other words, $e'_\nu{}^a$ describes the same gravitational field as $e_\mu{}^a$, but in the $\mathscr{Y}$ system of coordinates. Moreover, the Ricci scalar still vanishes around the point P:

$$R'\big|_P = R'(y_P) = R\left(F^{-1}(y_P)\right) = R(x_P) = 0\,. \tag{2.9}$$

Let us now proceed considering the new gravitational field $E_\nu{}^a$ defined as follows

$$E_\nu{}^a(x) = e'_\nu{}^a(x)\,, \tag{2.10}$$

namely as the primed field in the old system of coordinates $\mathscr{X}$. It is worth stressing that the gravitational field described by E is different from the one described by e, in particular, altough the Ricci scalar constructed by e is zero around the point P, namely $R(x_P) = 0$, we cannot draw the same conclusion for the Ricci scalar of the field E. In fact we have:

$$\mathscr{R}|_P = \mathscr{R}(x_P) = R'(x_P) = R\left(F^{-1}(x_P)\right)\,. \tag{2.11}$$

In other words, the Ricci scalar associated to the field E in the space-time point P is given by the Ricci scalar of e calculated in the point $Q = F^{-1}(x_P)$ and for no reason it must be zero.

It results that, if the gravitational field e is a solution of the equations of motion, E is a solution too. The reason is that the field E is described in the system of coordinates $\mathscr{X}$ by the same function describing the gravitational field e in the system of coordinates $\mathscr{Y}$; since the field equations do not change under a coordinates transformation, if e is a solution, so is E.

This is in essence the content of the so called *Einstein's hole argument*. The conclusion of the above described argument is that the generally covariant field equations are not deterministic,

because even though e and E are both solutions of the same field equations, they do not determine the physics at the same space-time point P. For example, while the Ricci scalar associated to the gravitational field e is zero around P, the same scalar calculated using the field E is in general different from zero around the same space-time point. Since we know that classical physics is deterministic we are at a crossroads: either the field equations cannot be generally covariant, or fixing the space-time event P has no physical meaning.

Einstein had the courage to take the right road, i.e. he understood that there is no physical meaning in fixing a particular space-time point on a generally covariant space-time. In this respect, let us consider a solution of the Einstein equations e and two particles moving in this particular gravitational field. The motion of the particles is described by their respective world lines, $x_1(\tau)$ and $x_2(\sigma)$, which are determined by the gravitational field. We suppose, without any loss of generality, that the world lines of the two particles intersect at the space-time event P. Now consider the gravitational field $E = \phi_* e$, where $\phi : M \to M$ is a diffeomorphism; obviously the particles world lines $x_1(\tau)$ and $x_2(\sigma)$ are no longer solutions of the particles equations of motion in the new gravitational field. In fact, the new particles world lines, determined by the gravitational field E, can be easily calculated once the world lines in the gravitational field e have been given, they are:

$$X_1(\tau) = [\phi x_1](\tau) \qquad \text{and} \qquad X_2(\sigma) = [\phi x_2](\sigma) . \tag{2.12}$$

In other words, a diffeomorphism, acting both on the gravitational field and on particles world lines, sends solutions to solutions. Furthermore, as a consequence of the active diffeomorphism, the particles do not intersect anymore in P, but in $Q = \phi(P)$. So, the fixed point P loses its absolute meaning and the right physical entity is the point determined by the intersection between the world lines of the particles. In this sense the theory does not predict the value of the gravitational field around the space-time point P, but around the point determined by the intersection of two world lines. In this sense, the issue contained in the hole argument is solved, the theory is deterministic because it predicts the same value of the gravitational field around the same *physical* (namely determined via a diffeomorphisms invariant construction) space-time point. In other words, the characteristics of the gravitational field e around the intersection of the world lines $x_1(\tau)$ and $x_2(\sigma)$, for example the flatness of space-time around this point, are exactly mimed by the gravitational field E around the intersection of the world lines $X_1(\tau)$ and $X_2(\sigma)$. This means that the theory has a gauge invariance in the sense of Dirac: different solutions, correlated by gauge transformations, represent the same physical situations (see § 3.3). The gauge group is the group of diffeomorphisms, which reflects the fact that the localization of an event is not an absolute procedure, but is related to the particles and fields themselves. We will see in paragraph 2.2.5 that the diffeomorphisms invariance has striking consequences on the theory, which, as a consequence, can be characterized as a partly relational theory.

Therefore the space-time of GR is a classical structure, but it is not absolute as in classical and quantum mechanics. In other words, it is of course possible to localize sharply a point on a generally covariant space-time, but the localization procedure requires the presence of particles and fields. In particular the diffeomorphisms invariant procedure for the localization of an event is based on the possibility to sharply recognize the points of intersection between the world lines of particles.

2.2.4 Space-time in Quantum Gravity

What we are going to say here must be intended by the reader as an attempt to explain why the idea of an absolute limit on the localization of an event has a so long tradition in the quantum gravity research. We cannot of course give a complete and consistent description of space-time in quantum gravity simply because the existing theories give us only a pictorial idea of how quantum space-time could be. So we do not pretend neither to introduce new insights into the quantum gravity problem nor to give a new pictorial description of quantum space-time. Nevertheless, we think it could be useful and interesting to investigate how to give an operative meaning to the localization of an event in a theory that should incorporate both the general covariance and the quantum uncertainty principles.

In this respect, we want to summarize the conclusion obtained before:

1. in order to satisfy the principle of general covariance, the localization of space-time events must be realized via a diffeomorphisms invariant procedure;

2. the quantum uncertainty does not allow to localize sharply a finite mass particle, because this would require the injection into the system of such a quantity of energy that it is impossible to neglect the creation of copies of the analyzed particle.

Let us now construct a physical diffeomorphisms invariant procedure of localization. In order to localize an event we need at least two interacting particles, moving along their world lines, we consider a massive particle and a probe able to interact with the particle. We know from what stated above that the larger is the mass of the particle greater is the accuracy of the localization: the relation between the energy of the probe and the uncertainty in the localization of the particle is $E = (\Delta x)^{-1}$. The physical explanation of this formula is simple: in order to determine the position of an object with a given accuracy Δx we have to use a probe, represented by a massless particle interacting with the system under study, localized in a region at least comparable with the accuracy we require for the experiment. The localization of the probe is proportional to its Compton wavelength, so the accuracy of the experiment is proportional to the inverse of the energy of the probe. It is well known that to reveal smaller and smaller structures in particles physics we have to use larger and larger energy test particles. In this specific context, we have to consider also the presence of the gravitational field, which fixes a limit on the energy of the system. Indeed, it is necessary that the Compton wavelength of the system is larger than its Schwarzschild radius, this means that the energy during the interaction must be smaller than the Planck mass. We consider in this pictorial context the ideal scattering process, which consists in a collision between a massless probe and a particle of mass M, the collision lasts for a time Δt. During this period of time the system is considered frozen. We can compute the typical time of the interaction Δt by a very simple argument: It is the time a signal takes to exchange information between the probe and the particle, so it is at most equal to the distance Δx. The energy contained in the gravitational field during the interaction is $V \approx \frac{\ell_{Pl}^2 EM}{\Delta t}$, where E is the energy of the probe. During the collision the total energy of the system must be smaller than the Planck mass $M_{Pl} = \ell_{Pl}^{-1}$, thus we can write the following relation

$$\frac{\ell_{Pl}^2 EM}{\Delta t} < M_{Pl} = \ell_{Pl}^{-1} . \tag{2.13}$$

Now, remembering that the spatial displacement between the probe and the particle is at most equal to the accuracy of the localization and that the best we can do is to increase the energy of the probe till to the mass of the particle (in order to avoid the creation of copies of the particle under study), by using the relation above we obtain:

$$\frac{\ell_{Pl}^2}{(\Delta x)^3} < M_{Pl}\,. \tag{2.14}$$

From the expression above we deduce that

$$\Delta x > \ell_{Pl}\,, \tag{2.15}$$

namely the best accuracy we can determine the position of a particle is greater than the Planck length.

So space-time in Quantum Gravity is not classical. In other words, taking into account both the general covariance and the quantum uncertainty principle, we cannot sharply localize an event because there is an intrinsic limit given by the Planck length in the accuracy of a localization procedure. As a consequence, all the intuitive concepts about space and time must be abandoned in a Quantum Gravity theory, the resulting space-time structure is genuinely non classical [2, 4]. More radically, we could say that in Quantum Gravity space-time does not exist at all, in its place we have a fuzzy quantum structure, which fits well with the pictorial representation Wheeler gave many years ago: Quantum space-time should appear like a foam. The quantum foam is a state of the gravitational field, which, at that stage, cannot be identified with a measurable structure in a proper sense. In fact, what is generally intended as space-time fits well with its classical actualization. The quantum foam, instead, cannot be considered as a space-time, for example the existence of a minimum length suggests a fundamental discrete (non-continuous) structure. In order to reintroduce the common concept of space-time in Quantum Gravity, namely as a measurement of the time elapsing between two events and the spatial distances separating two disjoint events, we have to give it a different status. We mean, considering it as an "actualization", rather than an "idealization". One way to "actualize" it is through a relational procedure, which is the next argument we want to discuss.

2.2.5 Relational versus Absolute space-time

The discussion about relational or absolute space-time could appear as a pure philosophical one, in stead it regards profoundly Physics and the debate is still open and stimulating [15, 16, 17] (and references therein). This debate is a long standing one and can be traced back to publication of Newton's *Principia Mathematica* in 1687. In his book Newton espoused an *absolute* notion of space-time, according to which the geometry of space-time provided a fixed, eternal and immutable background structure on which particles move. In striking contrast with Descartes, Leibniz, Huygens, who, in stead, espoused the *relational* point of view, according to which space-time has to be intended as a set of "relations" among real objects and events [29] (detailed discussions on this argument can be found in [18, 19, 20]). The Newton's absolute view won against relationalism, supported by the great empirical success Newtonian mechanics had and, for a long period, no doubts were raised on which notion of space-time Physics should be based on. But the general

covariance principle introduced by Einstein and mathematically expressed via the invariance under diffeomorphisms seems to destroy the Newtonian absolute description of space-time. Essentially, in the Einstein theory fixed space-time is replaced by a dynamical structure on which the events are no longer points with assigned coordinates, but interactions between physical particles: in other words relations. So, the relational point of view deserves to be taken in serious consideration as directly suggested by the commonly accepted theory describing space-time and gravitation.

So the question we want try to answer is: to what extent is GR a relational theory? Obviously, to answer to this question we have to capture the differences between Newton and Einstein mechanics, as we are going to do, taking into account the "space-time limitations" we have.

Newton's mechanics is based on the existence and physical definition of inertial reference frames. They play a central role in all the classical theories. In particular, they allow to distinguish between accelerating and uniformly moving point particles. In fact, once fixed an inertial reference system, the distinction between what is accelerating and what is moving uniformly is a property of the geometry of the absolute space-time (background), which is completely independent of the configuration of the matter. In other words, in Newtonian physics there is a clear and absolute distinction between inertial and non-inertial motion. Furthermore, this distinction does not depend on something internal to the physical system, but only on the external geometrical properties of space-time. Physically, we can distinguish between accelerated or uniformly moving particles by looking at the geometry of the reference frame glued to a generic particle: The presence of non-inertial forces allows us to make the distinction. This viewpoint was challenged by Mach, who proposed to eliminate absolute space-time as a cause of distinction between accelerating and non-accelerating motion, replacing it with a dynamically determining procedure. According to Mach, the distinction between accelerating and non-accelerating motion should be determined via the relations between all the structures which compose the whole Universe.

Mach's idea strongly influenced Einstein, who realized that acceleration should be determined with respect to a reference frame dynamically determined by the configuration of the whole system. As a consequence it does not exist any privileged reference frame in the Universe and physical laws must be equivalent in all the frames. Now, since local reference frames are strictly connected with the geometry of space-time, then space-time itself becomes a dynamical field, no more fixed and immutable, but interacting with matter and affected by the matter content of the system. The equivalence of any reference frame, mathematically expressed by the invariance of the theory under general coordinates transformations, suggested the name General Relativity for such a theory of space-time. It is worth noting that, even though Einstein was surely influenced by Mach, we cannot naively conclude that GR is a Machian theory [2], but, on the other hand, we can say that GR is, by construction, a partly relational theory as we are going to motivate.

Generally, a physical theory postulates that Nature is made up of a very large collection of elements. The form of the physical theory is based on the properties acquired by these physical entities. These properties, in an absolute theory, are referred to a fixed structure as, e.g., the Newtonian space-time, where the properties of the particles are described by their coordinates on space-time, evolving with respect to the absolute time t. Put differently, space-time plays the role of background to which the dynamics is referred. The same role is played by a regular lattice, often used in the framework of particles physics, in fact particles or fields are confined on the nodes of the lattice, which does not change in time. The main assumption of a relational theory is that there

does not exist any background; so the question is: to which structure is the dynamics referred? The relational view presumes that the fundamental properties of the elementary entities consist entirely in relations between the elementary entities themselves. Dynamics concerns with the changes of these relations. The pictorial description of a relational theory is a graph [2], where the nodes represents the entities, while their properties are incorporated in the connections between the nodes. The state of the system is just the structure of the links between the nodes and the dynamics is the law which determines the changes of connections (relations) between different nodes. It is important not to confuse the pictorial description of a relational graph with the common image everyone has of a regular lattice. In this respect, it is worth noting that the pictorial description of a relational system is completely abstract, while a regular lattice describes a precise physical situation, i.e. continuous space-time is substituted by a lattice and fields dynamics is referred to this discrete fixed and external structure.

In relational theories also time loses its usual meaning. Evolution is, in general, incorporated in the changing of relations between the physical entities and, since time is the parameter of the evolution itself, it acquires a relational meaning as well. The concept of relational time has a great importance in QG (see [15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30] and references therein).

Having described what we mean when we speak about relationalism, let us now focus the attention on GR. GR is a complicated theory describing all the manifestations of the gravitational interaction, in particular it describes gravitationally dominated subsystems of the Universe as, for example, black holes or gravitationally bound systems. In some models describing subsystems of the Universe, generally we impose conditions on the fields and metric at the boundary, in this case the question whether or not GR is a relational theory is not interesting [29], because imposing the boundary conditions is equivalent to introduce a background. Moreover, the existence of a region of space-time which is external to the system we are modeling would imply that the theory is not a fundamental one. But GR is widely believed to be the theory describing the whole Universe, in other words GR is considered the best candidate for a cosmological theory. This fact is absolutely not in contrast with the fact that the same theory, with suitable boundary or asymptotic conditions, could describe subsystems of the Universe, although one could expect that some modifications appear in the dynamical equations, as for example a non-vanishing cosmological constant. Anyway, this argument is not affected by the choice of a particular dynamics, which regards a completely different problem with respect to the relational content of the theory, mainly related to the fundamental concepts the theory is based on. Therefore, assuming that GR is a cosmological theory, the question whether or not GR is a relational theory acquires a profound meaning.

GR contains a lot of structures, which are fixed *a priori*, they are: dimensions, signature, topology and differential structure. All these structures belong to what is intended as "background," in fact, they can be varied from model to model, but they are fixed and are not subjected to dynamical laws. More precisely, they describe the manifold $\mathscr{M}$; in stead, the metric $g_{\mu\nu}$ and tensor fields $T^{(a)}$ are the dynamical entities of the theory. So, a space-time corresponds to a determination of the manifold, metric and fields, namely $\left(\mathscr{M}, g_{\mu\nu}; T^{(a)}\right)$. But, in order to define a physical space-time, we have to take into account the gauge freedom of the theory, which, as described in paragraph 2.2.3, is encoded in the invariance under the group of diffeomorphisms. Therefore we define a physical space-time as an equivalence class of manifolds, metrics and fields under the action of the

group $Diff(\mathscr{M})$. We denote the equivalence class as $\{\mathscr{M}, g_{\mu\nu}; T^{(a)}\}$. Now, as already mentioned, the points and open sets of the manifold $\mathscr{M}$ are not preserved under the action of the diffeomorphisms group. Diffeomorphisms send points to other points, in this sense the information encoded in the physical space-time is a system of relations between the fields, rather than a collection of the values fields take in the generic points of the manifold. Then, apart from the specification of topology, signature, dimensions and differential structure, GR is a relational physical theory.

It remains to answer a question: which is the physical entity replacing the Newton's absolute space-time? The answer is now simple: it is the gravitational field! In other words, it is the gravitational field that tells objects if they are accelerating or not. This is the profound difference between Newton's and Einstein's mechanics. In Newton mechanics the whole dynamics is referred to an absolute structure external to the system, in Einstein's mechanics the dynamics is referred to the dynamical gravitational field, carrying out the relational idea.

2.3 Possible phenomenological implications of QG

Generally speaking, a quantum theory of gravity should incorporate a sort of *space-time foam* [31, 32]. The physical Nature of this foam depends on the theory describing the quantum effects of gravity, but its phenomenological implications are shared by the different approaches to QG. In particular, a possible candidate for a quantum gravity effect is an energy dependent electromagnetic dispersion relation *in vacuo*. The modified dispersion relation we analyze is of the following form

$$p^2 = E^2\left(1 + f\left(\frac{E}{E_{QG}}\right)\right), \tag{2.16}$$

where E_{QG} is an effective quantum gravity energy, which can be identified with the Planck energy, but in some model, as for example, large extra-dimensions theories, it could be even lesser than the Planck energy. Now, let us suppose for simplicity that the Hamilton equations of motion are approximately valid in the present scenario, then the velocity of the particle follows from equation (2.16) and, for energies much lesser than the Planck scale, it results to be:

$$v = \frac{\partial E}{\partial p} \approx 1 - \xi\frac{E}{E_{QG}}, \tag{2.17}$$

where ξ is a positive or negative factor, which should be fixed by the particular framework.

Before deepening into the phenomenological aspects, let us first briefly review how modified laws for the propagation of particles have emerged independently in different quantum-gravity approaches. It was generally believed that QG effects interest too high energies to be experimented, until some years ago the first suggestions that quantum-gravitational fluctuations might modify particles propagation in an observable way began to appear [33, 34]. As a consequence, different classes of physical phenomena began to be studied. In particular, the effect that a modification in the particles propagation could have on the neutral kaon system [35, 36] was tested in laboratory experiments, fixing a lower limits on parameters analogous to E_{QG} [37]. Other examples of quantum gravitational effects more related to String Theory and LQG can be found in [38] (see also [39] for another string motivated deformation) and [40, 41]. Deformed dispersion relations, consistent with the formula above, arise also in other approaches as the "κ" quantum deformations of the Poincaré symmetry [42, 43] or quantization of point particles on a discrete space-times [44].

Let us now deepen into the problem of finding a physical system which can be used to test such a deformation on the dispersion relation. Surely the deformation on the velocity of the photon in (2.17) is really small for practical purpose, nevertheless it could give a sensible effect if photons of different energies travel for a very long distance before being detected. In particular photons of different energies emitted at the same time acquire a relative time delay over a distance L of the order

$$\Delta t \approx \xi L \frac{\Delta E}{E_{QG}}. \tag{2.18}$$

So, wider is the spectrum of the emitted photons and larger the distance traveled, greater will be the time-delay effect. In this respect, the best candidates to observe such an effect are Gamma Ray Bursts (GRB).

GRBs are, in fact, explosive events at cosmological distances. The typical spectrum of emission is in the range $0.1 - 100$ MeV, but it can extend up to the TeV scale. Moreover, a time structure of the order of the millisecond is typically observed in the light curves. It should result now clear why GRB are good candidates to study the effects of deformation in the dispersion relation. By a simple calculation it results that a GRB with a time structure of the order of the milliseconds, emitting photons of energy of the order of few MeVs and at a distance of $\approx 10^{26}$ m $\approx 10^{10}$ ly, could test the QG structure up to $E_{QG} \approx 10^{19}$ GeV. Sensible sensitivities can be already obtainable from the existing GRB's data and we address the interested readers to [5, 6] and in particular to [45, 46]. It is clear that the same time-delay effects in the arrival time of photons of different energies could be due to a possible tiny photons mass, or to interaction with the interstellar medium, but a simple analysis excludes these possibilities.

We conclude this introductory section stressing the importance that this kind of phenomenological studies could have on the research in QG, as well as a possible bridge with particles physics.

3. Preliminaries

This Section is dedicated to describe some fundamental arguments considered as preliminary, in the sense that they represent the foundations which the forthcoming discussion is based on. They are pretty general and unrelated arguments, very well described in many books, in which many more details can be found. Here I collected the main results and definitions, according to my own experience as student, inspired by my own notes taken during courses and referring to my favorite textbooks. So the description is far from being complete and I exhort the interested readers to refer to the cited Literature for a more complete treatment.

As I said in the Introduction, the focus of this paper is on the canonical formulation of GR. In this respect, it is important to realize that in order to canonically formulate any theory, we have to clarify the causal structure of space-time. This is a trivial task when we treat gauge theories on a flat Minkowski fixed background, but everything becomes more involved when the dynamics of the theory directly concerns the geometrical structure of the space-time itself as in GR. In fact, the identification of the gravitational field with the geometry of space-time implies that some restrictions have to be imposed on its global (causal) structure in order to canonically formalize the theory. In particular, we have to clearly understand under which conditions it is possible to "split"

space-time and describe the dynamics of the gravitational field as the time evolution of a geometrical spatial quantity.[7] It is important to understand that this is a necessary step to consistently define a canonical theory; we recall, in this respect, that, in fact, in classical mechanics, the Hamiltonian can be considered as the momentum conjugate to the time coordinate. This should give an idea of how much complicated the situation is in GR, since the invariance under diffeomorphisms prevents from determining a preferred time coordinate and, in general, it does not exist any global system of coordinates.

For this reason, we consider the study of the causal structure of space-time as the natural starting point for the canonical formulation of gravity. In particular, the scope of the first part of this Section is the introduction of the Geroch theorem, which clarifies under which hypotheses a global time function can be assigned on a generic space-time. In this sense, the Geroch theorem restricts the class of space-times whose dynamics can be canonically formulated, generating a question about the resulting canonical quantum theory: Is canonical quantum gravity applicable to a restricted class of space-times as well, or the classical conditions can be relaxed in the quantum theory? Up to my understanding, this question cannot be rigorously answered until a complete canonical quantum gravity theory has been formulated; nevertheless, since the canonical quantization procedure is, generally speaking, a mathematical tool to face the problem of quantization, once the procedure has been rigorously completed, it is reasonable to expect that the resulting quantum theory will not be affected by purely classical restrictions, which are naturally relaxed by the quantization in a very precise and suggestive sense: In a complete theory of QG there is no room for space-time.

The discussion of the causal structure of space-time and the subsequent canonical formulation of gravity, open the way to another interesting argument, namely the initial value formulation of a theory with gauge symmetries and, in particular, of GR. It should be clear that, in fact, once rewritten GR as describing the "time" evolution of a 3-dimensional geometry, it is expectable that an initial value problem can be consistently formulated by assigning a complete set of initial data on the initial spatial hypersurface. Generally speaking, physical theories are formulated on a fixed Minkowski space-time and the task of the initial value problem is to extract the (unique) evolution of a classical physical system starting from a complete set of initial data, as referred to an external fixed background. In this sense, the case of GR is much more involved. We cannot, in fact, refer the evolution of the initial data to a preferred background, rather the theory describes the evolution of the background itself, once assigned a particular starting configuration for the metric and its time derivatives. In other words, the fact that the gauge symmetries of GR are mathematically described by the group of 4-diffs complicates the formulation of a well posed initial value problem, not only from a mathematical perspective, but also conceptually. Nevertheless, a suitably adapted procedure can be applied to GR in order to assign a well posed initial value problem and extract a unique evolution from the Einstein equations, even though with some limitations.

In order to introduce the reader into this argument, we start by briefly recalling some elements of group theory, useful to introduce gauge theories, which will be used as a useful example in what follows. In particular at to make the general description of the canonical formulation of a gauge

[7] It is worth noting that the word time here does not refer to the quantity measured by using clocks, which, on the one hand, would imply that the space-time metric is completely determined, but, on the other hand, it is the unknown variable in the Einstein equations.

theory more concrete, we will refer to the simple case of the electromagnetic field, which is a gauge theory of the abelian group $U(1)$.

Finally, once the causal structure of space-time has been clarified and the generalities about the electromagnetic theory described, we face the problem of the initial value formulation giving a brief account of the main theorems used to study the problem. Also in facing this argument the electromagnetic theory will be useful, providing a simple, but non-trivial example to show how gauge symmetries enters in the formulation of an initial value problem for a classical system.

3.1 Causal Structure of space-time

The causal structure of flat Minkowski space-time is very simple and intuitive: Once a limit on the propagation of a signal is fixed, we can associate to any event p in space-time a light cone. The future is represented by a half cone, while the past is represented by the other half. The events contained in the future half of the light cone can be reached by a matter particle leaving from p, all these events are generally referred as *chronological future* of p. More generally, all the events lying in the interior of the future light cone together with those on the cone itself represent the *causal future*, physically representing all the events which can be, in principle, influenced by a signal emitted from p.

The causal structure on a generic manifold M is only locally similar to that of flat space, globally, in fact, the situation can be much more complicate. Obviously, in this brief paragraph we cannot give a complete account of the causal structure of a generic space-time. Our aim is to give only a sketch of the problematic coming out studying the causal structure of a general space-time. Non-trivial topologies or space-time singularities in general complicate enormously the situation, so we restrict the discussion by pointing out only those definitions, theorems, and lemmas we consider useful for what will be said below. For a more detailed description of the causal structure and the problem of singularities in GR, we address the reader to the book of Wald [9] from which we extracted the main theorems of this Section and to the book of Hawking and Ellis [10] in which one can find complete demonstrations.

Let me begin giving a simple, but important

Definition 3.1. *The space-time* $(M, g_{\mu\nu})$ *is* time orientable*, if* $\forall p \in M$ *it is possible to make a continuous designation of future and past.*

The simplicity of this definition stands in its intuitiveness, the importance, instead, is connected with the necessity to distinguish a particular class of space-times: in what follows we will always refer to time orientable space-times. It is simple to understand that in general a non-simply connected space-time cannot be time orientable; from the physical viewpoint in a non-time orientable space-time we cannot consistently distinguish between going "forward in time" or "backward in time". Time orientable space-times satisfy the property expressed by the following

Lemma 3.1. *Let* $(M, g_{\mu\nu})$ *a time orientable space-time, then there exists a non-unique smooth non-vanishing time-like vector field* t^μ *on* M*.*

The proof of this Lemma is based on the paracompactness of M and we address the reader to [9] for a complete proof. It is interesting to note that the above Lemma suggests a more useful definition in order to designate time orientable space-times:

Definition 3.2. *The space-time* $(M, g_{\mu\nu})$ *is* time orientable*, if there exists on M a time-like, continuous, non-vanishing vector field.*

For the sake of completeness and selfconsistency of the material presented in this work we recall also the following well known definitions:

Definition 3.3. *A differentiable curve* $\gamma(t)$ *is said to be a* future directed time-like curve*, if at each* $p \in \gamma$ *the tangent vector* t^μ *is time-like and future directed.*

It is simple to generalize this definition to *future directed causal curve*, it is sufficient to replace the adjective "time-like" with "causal". The next definition automatically follows

Definition 3.4. *The set of events that can be reached by a future directed time-like curve starting from p represents the* chronological future *of p, namely*

$$I^+(p) = \{q \in M : \exists \lambda(t) \text{ (future directed time-like curve) with } \lambda(0) = p\,;\, \lambda(1) = q\}\,.$$

Again the definition of *causal future* is the same of that for chronological future except substituting the words "future directed time-like curve" with "future directed causal curve". Finally we remark that

Definition 3.5. *For any subset* $S \subset M$

$$I^+(S) = \bigcup_{p \in S} I^+(p).$$

It is worth noting that, even though in Minkowski space-time $I^+(p)$ consists of the interior of the future light cone, in an arbitrary space-time the situation could be more complicated and the usual properties of flat spaces are in general not applicable to arbitrary space-times (it is simple to construct examples of pathological arbitrary space-times removing points from flat Minkowski space). However, at least locally, the same properties remain valid as stated by the following

Theorem 3.1. *Let* $(M, g_{\mu\nu})$ *be an arbitrary space-time and let* $p \in M$*. Then there exists a convex normal neighborhood of p, i.e., an open set U with* $p \in U$ *such that* $\forall\, q, r \in U$ *there exists a unique geodesic* γ *connecting q and r and staying entirely within U. Furthermore for any such U,* $I^+(p)|_U$ *consists of all points reached by future directed time-like geodesics starting from p and contained within U, where* $I^+(p)|_U$ *denotes the chronological future of p in the space-time* $(U, g_{\mu\nu})$*. In addition* $\dot{I}^+(p)|_U$ *is generated by the future directed null geodesics in U emanating from p.*

The fact that all general relativistic space-times have locally the same qualitative causal structure as Minkowski space-time, means that globally remarkable differences could appear. As a consequence space-time could be not *causally well behaving*. In order to clarify this point consider, for instance, a space-time with topology $S^1 \times \mathbb{R}^3$, constructed identifying the $t = 0$ and $t = 1$ hyperplanes of Minkowski flat space. It is easy to realize that in such a space-time an observer should have no difficulty in altering past events; in fact, the integral curves with tangent vector $t^\mu = (\partial/\partial t)^\mu$ will be closed and time-like. As a consequence, we have $\forall\, p \in M$,

$I^+(p) = I^+(p) = M$ [9]. Although the previous example could seem rather artificial, we stress that many general space-times with the property of allowing closed time-like curves exist and they occur in much less artificial examples than that described above.

From a physical perspective space-times with non-trivial closed causal curves cannot be considered physically realistic, because an observer could alter past events. From a mathematical point of view we have to assure that time-like geodesics do not intersect themselves. But the problem is slightly more complicated than it could appear at a first glance, the reason being that we have also to consider physically unreasonable those space-times in which time-like geodesics come "arbitrarily close" to intersect themselves (without doing it). They could, in fact, violate the causality condition if a small perturbation of the metric occur. Then we characterize physical space-times as

Definition 3.6. *A space-time $(M, g_{\mu\nu})$ is said to be* strongly causal *if $\forall\, p \in M$ and every neighborhood U of p, there exists a neighborhood V of p contained in U, i.e., $V \subset U$ such that no causal curve intersects V more than once.*

So, strongly causal space-time are characterized by the fact that causal curves cannot come arbitrarily close to themselves, but this is not sufficient to assure that one is not on the verge to violate physical causality. For this reason it is in general reasonable to give a stronger notion of causality as follows:

Definition 3.7. *A space-time $(M, g_{\mu\nu})$ is said to be* stably causal *if there exists a continuous non-vanishing time-like vector field t^μ such that the space time $(M, \widetilde{g}_{\mu\nu})$, where*

$$\widetilde{g}_{\mu\nu} = g_{\mu\nu} - t_\mu t_\nu \tag{3.1}$$

possesses no closed time-like curve.

The definition of stable causality avoids that a strong causal space-time could violate causality by perturbing the metric. A perturbation of the metric could, in fact, "open out" the light cone so much that a causal curve can come arbitrarily close to itself. The light cones of space-time $(M, \widetilde{g}_{\mu\nu})$ is strictly larger than that of $(M, g_{\mu\nu})$, consequently if closed time-like curves do not exist for $(M, \widetilde{g}_{\mu\nu})$, surely they will not exist for $(M, g_{\mu\nu})$ too.

For our purposes, the content of the next theorem is particularly important.

Theorem 3.2. *A space-time $(M, g_{\mu\nu})$ is stably causal if and only if there exists a differentiable function f on M such that $\nabla^\mu f$ is a past directed time-like vector field.*

Or, in other words, a stably causal space-time is equivalent to the existence of a global time function. For brevity we do not give the proof of this theorem here, but we address the reader to [9, 10] for the details of the proof. Furthermore it is important to quote the result contained in the following

Corollary 1. *Stable causality implies strong causality.*

This result is, in a certain sense, expectable and allows us to conclude that stable causality is the appropriate notion to be sure that a space-time is not going to violate causality, which is a crucial request for all physical reasonable space-times.

Above we give the definition of causality, namely we studied the collection of events $I^+(S)$ which can be influenced by a set of events S, now we want to study the collection of events "completely determined" by a set of events S. Above all, we have to give the following two definitions:

Definition 3.8. *A set S is said* achronal *if* $I^+(S) \cap S = 0$.

Definition 3.9. *Let S be an achronal set, the* future domain of dependence *of S,* $D^+(S)$*, is the collection of events p such that every past inextendible causal curve passing through p intersect S.*

It is worth noting that the following relations hold $S \subset D^+(S)$ and, S being achronal, $D^+(S) \cap I^-(S) = 0$. The physical importance of the future domain of dependence relies on the fact that, since no signal can travel faster than light, then any signal received in $p \in D^+(S)$ must have been registered on S, therefore giving suitable initial conditions on S, we should be able to predict what will happen in P. On the other hand if $p \in I^+(S)$, but $p \notin D^+(S)$ it is possible to reach p with a signal not passing through S. In general the full domain of dependence of an achronal set S is defined as $D(S) = D^+(S) \cup D^-(S)$ and physically represents the complete set of events which can be completely determined in future and past by fixing initial conditions on S.

Definition 3.10. *A closed achronal set* Σ *of M such that* $D(\Sigma) = M$ *is said a* Cauchy surface.

Now since the edge of an achronal set S is the set of point $p \in S$ such that every open neighborhood U of p contains two point $q \in I^+(S)$ and $r \in I^-(S)$ and a time-like curve $\lambda(t)$ from r to q which does not intersect S; then it follows that the edge of a Cauchy surface is empty. Therefore, by the following

Theorem 3.3. *Let S be a closed achronal set with edge* $(S) = 0$*, then S is a three-dimensional, embedded,* C^0 *submanifold of M,*

we can conclude that the Cauchy surface is a 3-dimensional embedded C^0 submanifold of M. Moreover since Σ is achronal, it represents an "instant of time" of the Universe [9]. As a consequence, we give the following

Definition 3.11. (Wald [9]) *A space-time* $(M, g_{\mu\nu})$ *which possessing a Cauchy surface is* globally hyperbolic.

So, in a global hyperbolic space-time, we can predict or retrodict the entire evolution of the Universe by assigning suitable initial condition on the Cauchy surface Σ. Established the importance of globally hyperbolic space-time, we want to give a criterion for recognize globally hyperbolic space-times. In this respect, the first step is to introduce the so called Cauchy horizons:

Definition 3.12. *Let S be an achronal set, the future Cauchy horizon of S, denoted by* $H^+(S)$ *is*

$$H^+(S) = \overline{D^+(S)} - I^-\left[D^+(S)\right]$$

and let us immediately quote the following

Proposition 3.1. *Being* $H(S) = H^+(S) \cup H^-(S)$ *and being* $\dot{D}(S)$ *the boundary of the future domain of dependence, we have*

$$H(S) = \dot{D}(S)\,.$$

From the above proposition follows the following

Corollary 2. *If M is connected then a non-empty closed achronal set Σ is a Cauchy surface if and only if $H(S) = 0$.*

Proof. *If $H(S) = 0$ then by the proposition follows $\dot{D}(S) = 0$, thus $\overline{D(S)} = int[D(S)] = D(S)$, so $D(S)$ is simultaneously closed and open, but the only sets both open and closed are the empty set and the entire set, so we conclude $D(S) = M$.*

This corollary allows us to enunciate the following theorem, which represent a useful criterion to establish if a surface embedded in a manifold is a Cauchy surface:

Theorem 3.4. *If Σ is a closed achronal edgeless set, then Σ is a Cauchy surface if and only if every inextendible null geodesic intersects Σ and enters $I^+(S)$ and $I^-(S)$.*

Now, it is easy to understand that, if a space-time is globally hyperbolic, then no closed time-like curves can exist in M, because either a closed time-like (or causal) curve never intersects the Cauchy surface Σ violating global hyperbolicity or it intersects Σ twice violating achronality. This fact suggests that global hyperbolic space-times have a "well causal behaviour", as stated by the following

Theorem 3.5. (Geroch 1970 [47]). *Let $(M, g_{\mu\nu})$ be a globally hyperbolic space-time. Then $(M, g_{\mu\nu})$ is stably causal. Furthermore, a global time function, t, can be chosen such that each surface of constant t is a Cauchy surface. Thus M can be foliated by Cauchy surfaces and the topology of M is $\mathbb{R} \times \Sigma$, where Σ denotes any Cauchy surface.*

We refrain from giving the proof of this famous theorem addressing the interested reader to the original reference, nevertheless we want to stress its importance in view of the canonical approach and quantization of gravity. The Geroch's theorem, in fact, allows to recast the gravitational action in the canonical form operating a $3+1$ foliation of space-time, extracting a continuous function, which will play the role of evolution parameter. It will be clear that the dynamical degrees of freedom of the gravitational field are entirely contained in the geometry (metric modulo diffeomorphisms) of the 3-dimensional Cauchy surface. Pictorially, the evolution of the system can be described registering the changes of the 3-geometry going from one Cauchy surface to the next one, following the integral curve of the past directed time-like vector field $\nabla^\mu t$.

3.2 Elements of group theory

Let us now completely change argument, introducing some elements of group theory. Group theory represents one of the main tool to construct gauge theories of physical interactions. Here, we describe some of the elements necessary to this scope, refraining from deepen into the interesting geometrical aspects of gauge theories, related, e.g., to the theory of fiber bundles. This, even though important to appreciate the general mathematical structure of this framework, in fact, would lead us away from the scope of this paper.

So let us start from the following

Definition 3.13. *Let $\mathcal{G}$ and $\mathcal{G}'$ be groups, thus a map $f : \mathcal{G} \to \mathcal{G}'$ is an homomorphism if $\forall\, g_1, g_2 \in \mathcal{G}$ we have $f(g_1 g_2) = f(g_1) f(g_2)$. A homomorphism $h : \mathcal{G} \to GL(V)$ is called a* representation *of the group $\mathcal{G}$ and V is called the* representation space*. The representation is said to be of dimension N if the representation space is N-dimensional.*

A gauge theory can be constructed by using any compact group G, we recall that a group is said *compact* if the parameters which describe the group take values in a compact set. In what follows, we will focus our attention on the compact group $SU(N)$, which is represented by the special group of the unitary $N \times N$ matrices. Its fundamental representation is N dimensional, indeed, if $g \in SU(N)$ is an element of the group, we have that its representation $U = h(g)$ is a unitary $N \times N$ matrix, which acts on a N-dimensional complex space. Using the direct product, we can construct many other representations, which are in general reducible, usually called *tensorial representation*. For instance, let us consider the tensorial representation space $V \otimes V$, on which the elements of the group act as follows: $U_{ik}U_{jl}v_{kl} = v'_{ij}$. In general this representation is reducible, as one can verify constructing group invariants. Another possible representation is the conjugate one, which acts as well as the fundamental representation on the space V, consequently they have the same dimension. Usually, the N-dimensional conjugate representation is denoted by the symbol $\overline{N}$. Remarkably, the conjugate representation of the group $SU(2)$ is equivalent to the fundamental one.[8]

Before going further introducing the so called adjoint representation, we briefly describe the parametrization of the group $SU(N)$. In order to understand which is the most convenient parametrization for this group, we observe that the generic unitary matrix U can be rewritten as

$$U = e^{i\lambda}, \qquad \text{with} \qquad \lambda = \lambda^{\dagger}, \tag{3.2}$$

furthermore if $U \in SU(N) \Rightarrow \mathrm{Tr}\,\lambda = 0$, so the group $SU(N)$ can be parametrized by the hermitian matrices with null trace and, as a consequence, is determined by $N^2 - 1$ parameters.

Let λ^a with $a = 1, 2..., N^2 - 1$ be a basis on the space of the $N \times N$ hermitian matrices, H, with the orthogonality condition

$$\mathrm{Tr}\left(\lambda^a \lambda^b\right) = \frac{1}{2}\delta^{ab} \tag{3.3}$$

and commutation relation

$$\left[\lambda^a, \lambda^b\right] = i f^{abc} \lambda^c, \tag{3.4}$$

easily justifiable by considering that the commutator of two hermitian matrices is an anti-hermitian matrix. The structure constants of the group, f^{abc}, turn out to be real and completely antisymmetric in the exchange of their three indexes. The λ^a matrices are said generators of the $SU(N)$ group.

Therefore, the generic element $U \in SU(N)$ can be rewritten as

$$U = e^{ig^a \lambda^a}, \qquad \text{where} \qquad g^a \in \mathbb{R},\ \lambda^a \in H, \qquad \text{with} \qquad \mathrm{Tr}\,\lambda^a = 0\,. \tag{3.5}$$

Let us now consider the action of a $SU(N)$ operator infinitesimally close to the identity on the representation space. This can be parametrized by the first order expansion of the general parametrization (3.5), in symbols we have: $e^{ig^a\lambda^a} = I + ig^a\lambda^a$. The elements of the group in the connected component of the identity generate the algebra of $SU(N)$, which useful for the construction of gauge theories. In this respect, we want to anticipate here that the canonical constraints we are going to calculate below are the generators of small gauge transformations only, i.e. they generate those gauge transformations in the connected component of the identity. In other words, the behavior of the states of the theory under large gauge transformations, i.e. those generated by

[8] Two representations are said equivalent if it exists a unitary map $W : h_1(g) \to h_2(g)$, such that $W h_1(g) W^{\dagger} = h_2(g), \forall g$.

the elements of the group characterized by a non-vanishing winding number (see the Appendix B), cannot be deduced by the theory. This leads to an extremely interesting issue. The observables of a gauge theory are, in fact, invariant under the full group, so they can be used to super-select states of the theory belonging to gauge sectors characterized by different winding numbers. From this perspective the global (and in general non-trivial) structure of the gauge group enters in the physical outcomes of the theory.[9]

By considering the direct product between the fundamental and the conjugate representations, we can construct the following representation, acting on $N \times N$ matrices as follows:

$$v'_{ij} = U_{ik} U^*_{jl} v_{kl} \qquad \text{where} \qquad U \in SU(N), \quad v_{kl} \in V \otimes V\,. \tag{3.6}$$

The adjoint representation is reducible, to decompose it in its irreducible components, let us identify the generic element v_{kl} of the vector space as the "kl" entry of the matrix W, thus we have

$$W' = UWU^+\,. \tag{3.7}$$

It is worth noting that the hermitian matrices form an invariant subspace,[10] as one can easily verify. The trace is an invariant too, so that we can reduce the representation considering the subspace formed by the hermitian matrices with null trace.

The representation acting on the $N \times N$ hermitian matrices with null trace is called *adjoint representation* and has dimension $N^2 - 1$.

Let us now study the algebraic structure of the adjoint representation. We consider an infinitesimal transformation, which acts on the hermitian traceless matrix V, we have:

$$V' = UVU^+ = (I + ig^a \lambda^a) V \left(I - ig^b \lambda^b\right) = V + ig^b \left[\lambda^b, V\right]\,. \tag{3.8}$$

Now, by expanding V on the basis, we obtain the transformation law of the component V^a, namely:

$$V'^a = V^a - f^{abc} g^b V^c = \left(\delta^{ac} - f^{abc} g^b\right) V^c\,. \tag{3.9}$$

The above expression (3.9) allows us to identify the generators of the adjoint representation, they are, in fact, given by the matrices of components "bc" identified with $(Q^a)_{bc} = f^{abc}$. It is straightforward to verify the universality of the commutation rules of the generators, in other words starting from the Jacobi identity for the λ^a matrices and using the commutation relations (3.4) one can demonstrate that

$$\left[Q^a, Q^b\right] = f^{abc} Q^c\,. \tag{3.10}$$

In other words, the structure constants characterize the group independently from the representation, as their name suggests. It should appear obvious that the structure constants characterize

[9] Interestingly enough, this fact is not only of mathematical interest, in this respect that the solution of the so-called $U(1)_A$ puzzle in QCD is directly correlated to the topological aspects of the gauge group. For the sake of completeness, we also stress that the extension of the theory to contain topologically non-trivial terms in the action has led to an interesting "beyond the Standard Model" prediction, the existence of the axion, which is considered as a natural solution to the so-called *strong CP problem*.

[10] Considering only the hermitian matrices is not restrictive because a generic matrix A can be rewritten as $A = A_1 + iA_2$, with $A_1 = A_1^\dagger$ and $A_2 = A_2^\dagger$

only the algebraic properties of the group, not the global ones, which cannot be deduced from the algebra.

We are now ready to introduce $SU(N)$ Yang–Mills gauge theories. In this respect, let

$$\Psi(x) = \begin{pmatrix} \psi_1(x) \\ \vdots \\ \psi_N(x) \end{pmatrix} \tag{3.11}$$

be a collection of N Dirac spinor fields, the dynamics of which is described by the Dirac Lagrangian

$$\mathscr{L}[\Psi, \overline{\Psi}] = \overline{\Psi}(x) \left(i\gamma^\mu \partial_\mu - M \right) \Psi(x) , \tag{3.12}$$

where M is the $N \times N$ mass matrix.[11] The Lagrangian above is symmetric under the action of the $SU(N)$ group, acting on the spinor field Ψ and its conjugate $\overline{\Psi}$ according to the following rules:

$$\Psi(x) \to \Psi'(x) = U[g]\Psi(x) = e^{ig^a \lambda^a} \Psi(x) , \tag{3.13}$$

$$\overline{\Psi}(x) \to \overline{\Psi}'(x) = \overline{\Psi}(x) U^\dagger[g] = \overline{\Psi}(x)\, e^{-ig^b \lambda^b} . \tag{3.14}$$

where g^a are the $N^2 - 1$ constant parameters of the transformation. In other words, the collective spinor Ψ transforms in the vectorial representation of the group $SU(N)$.

The gauge principle states that the minimal coupling interaction between fermions and boson gauge fields can be obtained by requiring the local gauge invariance of the Dirac Lagrangian. In other words, we can directly extract the right minimal coupling by simply requiring that the Lagrangian (3.12) is invariant under the following local transformations

$$\Psi(x) \to \Psi'(x) = U[g(x)]\Psi(x) = e^{ig^a(x) \lambda^a} \Psi(x) , \tag{3.15}$$

$$\overline{\Psi}(x) \to \overline{\Psi}'(x) = \overline{\Psi}(x) U^\dagger[g(x)] = \overline{\Psi}(x)\, e^{-ig^b(x) \lambda^b} . \tag{3.16}$$

It is worth noting that, in fact, the kinetic term in the Lagrangian (3.12) cannot be invariant under such a transformation for a very simple reason: The derivative operator is defined through a limiting procedure of objects transforming differently under the local $SU(N)$ group. More practically, we can say that the ordinary derivative of a spinor do not transform in the vectorial representation of the $SU(N)$ group. Even more straightforwardly one can note that the action of the derivative operator on the transformed spinor field $\Psi'(x)$ generates a term which cannot be reabsorbed unless we introduce a counter term.

So the invariance requirement induces to modify the Lagrangian by defining a new derivative operator, D_μ, which transforms in the adjoint representation of the $SU(N)$ group, so that

$$D_\mu \psi \to \left(D_\mu \psi \right)' = U[g(x)] D_\mu U^\dagger[g(x)] U[g(x)] \Psi = U[g(x)] D_\mu \Psi . \tag{3.17}$$

This can be easily achieved by introducing a connection gauge field $A_\mu = A^a_\mu \lambda^a$, valued on the group $SU(N)$ and transforming according to the following equation

$$\begin{aligned} A_\mu \to A'_\mu &= U[g(x)] A_\mu U^\dagger[g(x)] - iU[g(x)] \partial_\mu U^\dagger[g(x)] \\ &= e^{ig^a(x)\lambda^a} A^c_\mu \lambda^c e^{-ig^a(x)\lambda^a} - ie^{ig^a(x)\lambda^a} \partial_\mu e^{-ig^a(x)\lambda^a} . \end{aligned} \tag{3.18}$$

[11] The mass matrix M can be in general non-diagonal. It fact, it can be easily diagonalized by a chiral transformation. This procedure, even though completely safe in the classical theory, can produce striking effects in the quantum theory, because of the chiral anomaly.

Consequently, for an infinitesimal gauge transformation we have:

$$A^a_\mu \rightarrow A'^a_\mu = A^a_\mu + if^{abc}g^b A^c_\mu - \partial_\mu g^a = A^a_\mu - \mathscr{D}_\mu g^a\,, \tag{3.19}$$

where the covariant derivative of the adjoint representation, $\mathscr{D}_\mu$, has been defined. In order to rigorously introduce the mathematical concept of connections and study their properties, we should digress on the geometrical theory of fiber bundles, but this is far from the scope of this paper. So let us just remark that the replacement

$$\partial_\mu \Psi(x) \rightarrow D_\mu \Psi(x) = \partial_\mu \Psi(x) + iA_\mu \Psi\,, \tag{3.20}$$

is motivated by the fact that the global symmetry group has been promoted to be a local group and this introduces a fiber bundle structure which motivates the introduction of connections.

Now, it is a simple exercise writing a Dirac Lagrangian which features the required properties of symmetry, i.e.

$$\begin{aligned}\mathscr{L}[\Psi,\overline{\Psi},A] &= \overline{\Psi}(x)\left(i\gamma^\mu D_\mu - M\right)\Psi(x)\\ &= \overline{\Psi}(x)\left(i\gamma^\mu \partial_\mu - M\right)\Psi(x) - \overline{\Psi}(x)A^i_\mu \lambda^i \Psi = \mathscr{L}[\Psi,\overline{\Psi}] + \mathscr{L}_{\text{int}}[\Psi,\overline{\Psi},A]\,,\end{aligned} \tag{3.21}$$

Remarkably, a new interaction term denoted by $\mathscr{L}_{\text{int}}$ has appeared in the full Lagrangian, representing the minimal interaction of the $SU(N)$ gauge connection and the spinor matter fields.

So, by fulfilling the requirement of the gauge principle we have been able to automatically extract the correct minimal interaction between fermions and boson gauge fields. But, the gauge principle does not give us any hint on how to construct a kinetic term for the new gauge field A_μ. From the general theory of fiber bundles we know that once a connection has been defined, a natural curvature tensor can be constructed. Usually, in physics books, the curvature tensor is defined as follows:

$$\left[D_\mu, D_\mu\right]\Psi = iF_{\mu\nu}\Psi\,, \tag{3.22}$$

where $F_{\mu\nu} = F^a_{\mu\nu}\lambda^a$ is said curvature or strength tensor. By the above definition it is easy to extract the explicit expression of the curvature tensor as function of the connection, i.e.

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + \left[A_\mu, A_\mu\right]\,. \tag{3.23}$$

The curvature tensor satisfies the so-called Bianchi identity,

$$D_{[\mu}F_{\nu\rho]} = 0\,, \tag{3.24}$$

which, remembering the definition (3.22), can be considered as a consequence of the Jacobi identity associated to the covariant derivative operator D_μ.

It is worth noting that if the gauge symmetry is represented by the $U(1)$ group, the structure constants vanish as a consequence of the vanishing of the commutator between the group generators (Abelian group) and, in this specific case, the connection physically describes the electromagnetic potential, while the curvature tensor is the field strength.

The curvature tensor transforms as a proper element of the adjoint representation (in contrast with the connection), specifically we have:

$$F_{\mu\nu} \rightarrow F'_{\mu\nu} = U[g(x)]F_{\mu\nu}U^\dagger[g(x)]\,. \tag{3.25}$$

This fact and the existing analogy with the electromagnetic field suggest the expression of the action for the field A_μ. Specifically, requiring that the dynamics be described by second order partial differential equations of motion, a good action is:

$$S[A] = -\frac{1}{2}\int d^4x \mathrm{Tr} F_{\mu\nu}F^{\mu\nu} = -\frac{1}{4}\int d^4x \sum_a F^a_{\mu\nu}F^{a\mu\nu}\,, \tag{3.26}$$

where in the second equality we used the orthogonality condition (3.3). The trace acts on the N^2-1 gauge internal indexes and makes the action invariant under the action of the $SU(N)$ group as can be easily demonstrated.

As a final remark, note that we can add a term to the above action without affecting the classical equations of motion, i.e. the action

$$S_\theta[A] = -\frac{1}{4}\int d^4x \sum_a F^a_{\mu\nu}F^{a\mu\nu} + \frac{\theta}{64\pi^2}\int d^4x \sum_b F^b_{\mu\nu}F^b_{\rho\sigma} \tag{3.27}$$

is dynamically equivalent to $S[A]$, since the θ-term contribution to the classical equations of motion vanishes identically according to the Bianchi identity (3.24). Even though such a term does not affect the classical theory, it produces striking effects in the non-perturbative quantum theory (further details are given in Appendix B). Keep in mind the form of the action $S_\theta[A]$, since an analog situation will characterize the Ashtekar–Barbero formulation of canonical gravity.

3.3 Gauge Symmetries and Constraints

Let us now start describing the canonical formulation of theories with gauge symmetries. This argument is particularly important for us, since it is the classical starting point of the canonical quantization of gauge systems. At the end of the section we will calculate the canonical Hamiltonian of the electromagnetic theory, which will serve as a description of the formalism in a simple practical example.

3.3.1 General formalism

Let us consider a physical system described by a Lagrangian

$$L = L(q_i, \dot{q}_k)\,, \tag{3.28}$$

where Latin indexes denotes the different generalized coordinates and velocities determining the classical motion of the system on the configuration space. The Lagrangian equations of motion will be

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \frac{\partial L}{\partial q_i}\,. \tag{3.29}$$

It is possible to rewrite the equations of motion of a physical system as referred to the generalized coordinates and their conjugate momenta, by suitably define the momenta and performing the so-called Legendre transformation. In this respect, we define the momentum p_k as

$$p_k = \frac{\partial L}{\partial \dot{q}_k}\,. \tag{3.30}$$

The couple of canonical variables (q_i, p_k) define the so-called phase space. Usually, in classical mechanics, we introduce the hypothesis that the momenta are independent functions of the velocities. Even though this hypothesis is fulfilled in many interesting classical macroscopic systems, it is too restrictive to be applied to more fundamental physical theories based on the gauge principle.

In general, in fact, the Lagrangian could result singular, namely

$$\det\left[\frac{\partial^2 L}{\partial\dot{q}_i\partial\dot{q}_k}\right]=0\,. \tag{3.31}$$

If this is the case, the velocities cannot be all inverted as functions of the generalized coordinates and momenta. By the way, it is easy to show starting from the Lagrangian equations of motion that the vanishing of the above determinant implies also that the velocities cannot be uniquely expressed as functions of the positions and velocities. As a consequence, the momenta are not all independent, rather some relations resulting from the definition of the momenta crop up. Specifically, we have

$$\phi_m(q_i,p_k)=0\,,\qquad m=1,2,\cdots,M\,. \tag{3.32}$$

These relations are called *primary constraints*, emphasizing that they result from the very definition of the momenta. The submanifold determined in the phase space by the conditions (3.32) is called primary constraint surface.

At this point, let us define the following function

$$H=\sum_k p_k\dot{q}_k-L(q_i,\dot{q}_k(p_j))\,, \tag{3.33}$$

and calculate its variation:

$$\begin{aligned}\delta H&=\sum_i\delta p_i\dot{q}_i+\sum_k p_k\delta\dot{q}_k-\sum_j\left(\frac{\delta L(q_i,\dot{q}_k)}{\delta q_j}\right)\delta q_j-\sum_l\left(\frac{\delta L(q_i,\dot{q}_k)}{\delta\dot{q}_l}\right)\delta\dot{q}_l\\&=\sum_k\delta p_k\dot{q}_k-\sum_j\left(\frac{\delta L(q_i,\dot{q}_k)}{\delta q_j}\right)\delta q_j\,,\end{aligned} \tag{3.34}$$

this is called canonical Hamiltonian and its variation depends only on the positions and canonical momenta. However, the Hamiltonian defined above is not uniquely determined because we can add to it any linear combination of the primary constraints, which are to be zero. In other words, the thory cannot distinguish between the Hamiltonian H defined above and the new Hamiltonian $H^\star = H+\sum_m c_m\phi_m$. Nevertheless, since the above equation has to hold for any variation, provided that the variation preserves the conditions (3.32), we can obtain the following Hamiltonian equations of motion:

$$\dot{q}_i=\frac{\partial H}{\partial p_i}+\sum_m u_m\frac{\partial\phi_m}{\partial p_i}\,, \tag{3.35a}$$

$$-\dot{p}_k=\frac{\partial H}{\partial q_k}+\sum_m u_m\frac{\partial\phi_m}{\partial q_k}\,, \tag{3.35b}$$

which are in accordance with the general method of the calculus of variations applied to a system with constraints. We stress that the symbol u_m denotes a completely arbitrary set of functions. Before going on, it is convenient to introduce a formalism that allows to write the canonical equations

of motion in a compact way. We are referring to the Poisson brackets. Let f and g be two generic functions of the canonical variables, then we define:

$$[f,g] = \sum_i \left(\frac{\partial f}{\partial q_i}\frac{\partial f}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial f}{\partial q_i} \right) . \tag{3.36}$$

It immediately follows that

$$[q_i, p_k] = \delta_{ik} \, , \tag{3.37}$$

where δ_{ik} is the Kronecker symbol. Moreover, it is easy to demonstrate that the equations of motion (3.35a) can be easily rewritten as

$$\dot{q}_i = [q_i, H] + \sum_m u_m \, [q_i, \phi_m] \, , \tag{3.38}$$

$$\dot{p}_k = [p_k, H] + \sum_m u_m \, [p_k, \phi_m] \, . \tag{3.39}$$

more generically, we have

$$\dot{g} = [g, H] + \sum_m u_m \, [g, \phi_m] \, , \tag{3.40}$$

where g is a generic functions of the canonical variables. The Poisson brackets are defined only for those quantities that are functions of the canonical variables, nonetheless, the above definition for the time derivative of the generic function g can be rewritten more concisely as:

$$\dot{g} = [g, H_T] \, , \tag{3.41}$$

where the total Hamiltonian H_T is defined as

$$H_T = H + \sum_m u_m \phi_m \, . \tag{3.42}$$

One may immediately wonder about the acceptability of this definition, because one of the term that come out from Poisson bracket (3.41) is $\sum_m [g, u_m] \, \phi_m$. This is badly defined because the u_m are arbitrary functions not depending on the canonical variables. But, the correctness of this definition stems from the fact that the Poisson brackets multiply the vanishing functions ϕ_m. So, we have to carefully consider the fact that in dealing with the Poisson formulation of classical mechanics in theories with constraints, these have to be imposed only after having calculated all the Poisson brackets, otherwise we would affect the consistency of the construction. That is the reason why the weakly vanishing symbol "$\approx$" is widely used; it emphasizes the fact that the constraints have to be imposed at the end of the canonical analysis, limiting the evolution of the system to a restricted region of the phase space.

As a consistency requirement, we have to impose another condition on the dynamics. Namely, we have to require that the primary constraints surface be preserved by the Hamiltonian flow, i.e.

$$\dot{\phi}_n = [\phi_n, H] + \sum_m u_m \, [\phi_n, \phi_m] \approx 0 \, . \tag{3.43}$$

If the above equation is automatically satisfied, namely the time derivative of the primary constraints vanishes on the primary constraints surface, then no other consistency check is necessary.

But, in general, other different cases may occur. In particular, if $\dot{\phi}_n \neq 0$ for certain values of n, then some *secondary constraints* $\chi(q_i, p_k) \approx 0$ have to be imposed, further restricting the available region of the phase space. Obviously, the same consistency condition in (3.43) has to be satisfied by the secondary constraints, otherwise tertiary constraints have to be imposed and so on.

Since there is no fundamental difference between primary and secondary constraints, we collect all of them in the same symbol, i.e.

$$\phi_m(q_i, p_k) \approx 0\,, \qquad m = 1, 2, \cdots, M+K\,, \tag{3.44}$$

where K is the number of secondary constraints. A closer look at the consistency equation is now in order, because it, in fact, contains some interesting informations about the arbitrary functions entering in the Hamiltonian. We have:

$$[\phi_n, H] + \sum_m u_m\,[\phi_n, \phi_m] \approx 0\,, \tag{3.45}$$

generally, if $\det([\phi_n, \phi_m]) \neq 0$, we can extract a solution for the u_m by inverting the matrix $C_{nm} = [\phi_n, \phi_m]$; for convinience let us indicate this set of solutions as $U_m(q_i, p_k)$. Moreover, in order to write the general solution of the equation above, to the particular solutions $U_m(q_i, p_k)$, we must add the solution of the homogeneous equation, namely

$$\sum_m V_m\,([\phi_n, \phi_m]) \approx 0\,. \tag{3.46}$$

In general, there can be A independent solution of the equation above, which we indicate as V_{am}. So that the most general solution of the consistency condition is $u_n = U_n + \sum_a v_a V_{an}$, where the components of the functions u_m which can be fixed by the consistency conditions have been separated from that which remains arbitrary. In terms of this new expression for u_m, the total Hamiltonian can be rewritten as

$$H_T = H + \sum_m U_m \phi_m + \sum_a v_a \phi_a = H' + \sum_a v_a \phi_a\,, \tag{3.47}$$

where we have defined $H' = H + \sum_m U_m \phi_m$ and $\phi_a = V_{am}\phi_m$. The total Hamiltonian can be now used to calculate the canonical equations of motion through the Poisson brackets formalism. Obviously, even though some of the arbitrariness has been eliminated via the consistency condition, they still contain arbitrary functions, v_a. Nevertheless, the canonical equations of motion are equivalent, by construction, to the Lagrangian ones.

At this point the introduction of some terminology is in order. We call *first class* all those functions of the canonical variables which have weakly vanishing Poisson brackets with all the constraints. By remembering that the ϕ's are the only independent quantities which are weakly zero, then the function $S = S(q_i, p_k)$ is first class if

$$[S, \phi_m] = \sum_n c_{mn}\phi_n\,, \qquad \forall m = 1, \cdots, M+K\,. \tag{3.48}$$

All the other functions of the canonical variables are said to be *second class*. It is easy to demonstrate that the Poisson bracket of two first class functions is first class as well as the Hamiltonian

H' defined in (3.47). This terminology applies to constraints as well: We call *first class constraints* the set of ϕ_m with $m \leq K+M$ such that

$$[\phi_m, \phi_n] = \sum_p c_{mnp}\phi_p\,, \qquad \forall m = 1, \cdots, M+K\,, \tag{3.49}$$

while the others are referred as *second class constraints*.

Finally, we want to digress on the kind of transformation induced by the first class constraints. In order to do that, let us stress that in a theory with constraints the evolution of the generalized coordinates and momenta are not uniquely determined by the initial state because of the presence of the arbitrary functions v_a. This means that there are many choices of the fundamental variables that characterize the same physical state. In this respect, it is interesting to consider particular values of phase space variables (let us call them g) at an initial time, e.g. $g_0 = g(t=0)$, and look at their values after an infinitesimal temporal lapse δt. By using the Poisson brackets we obtain

$$g(\delta t) = g_0 + \dot{g}\delta t = g_0 + [g, H_T]\,\delta t = g_0 + \delta t\left([g, H'] + \sum_a v_a\,[g, \phi_a]\right). \tag{3.50}$$

Imagine now that we had taken different functions v'_a, then we would have obtained

$$g'(\delta t) = g_0 + \dot{g}\delta t = g_0 + [g, H_T]\,\delta t = g_0 + \delta t\left([g, H'] + \sum_a v'_a\,[g, \phi_a]\right). \tag{3.51}$$

In other words, during the infinitesimal time δt, the difference Δv_a of the two functions v_a and v'_a (i.e. $\Delta v_a = v_a - v'_a$) generates a difference between g and g' given by

$$\Delta g(\delta t) = \delta t \Delta v_a\,[g, \phi_a] = \varepsilon_a\,[g, \phi_a]\,, \tag{3.52}$$

where ε_a is a small quantity, being proportional to δt. So, according to the above rule, the variables describing a particular physical configuration of the system can be arbitrarily changed without affecting the state of the system. In other words, many different sets of canonical variables related each other by the above transformation equivalently describe the same physical state. Hence, the functions ϕ_a result to be the generators of gauge transformations.

Concluding, we have found that first class primary constraints generate gauge transformations, but it is, in general, expectable that also secondary first class constraints can generate gauge transformations, and, in fact, this happens in many mechanical systems. It is commonly believed that all the first class constraints are generators of gauge transformations, even though, we have to say, this belief is not supported by a rigorous proof and is sometimes referred as Dirac's conjecture.

3.3.2 Electromagnetic canonical theory

The electromagnetic theory provides a very simple, but non-trivial and practically interesting example to operatively use the procedure just described. In order to be definite, we start from the Lagrangian for the Maxwell's theory, i.e.

$$L(A, \partial A) = -\frac{1}{4}\int d^3x F_{\mu\nu}F^{\mu\nu}\,. \tag{3.53}$$

Here the fundamental variable is the electromagnetic potential A_μ, which, as is well known, is the connection field associated to the $U(1)$ gauge symmetry; namely, acting with a gauge transformation on A_μ, we have that

$$A_\mu \to A'_\mu = A_\mu + \partial_\mu \lambda \,, \tag{3.54}$$

where λ is a generic function of the space-time points. According to what said above, the presence of a gauge symmetry will reveal in the canonical theory through the appearance of first class constraints, associated with the generators of the gauge transformation. Let us check if this is the case, starting the canonical analysis of the electromagnetic theory by calculating the conjugate momenta to the variables A_μ. Namely,

$$E^\alpha = \frac{\partial L}{\partial A_\alpha} = -\partial^0 A^\alpha + \partial^\alpha A^0 = F^{\alpha 0}\,, \qquad \text{and} \qquad P^0 = \frac{\partial L}{\partial A_0} = 0\,, \tag{3.55}$$

are respectively the momenta conjugate to A_α and A_0 (the Greek indexes indicate purely spatial components, while the index 0 indicates the time component). The resulting phase space is 8-dimensional with coordinates $(A_\alpha, A_0, E^\alpha, P^0)$, and can be equipped with the following symplectic structure:

$$\Big[A_\alpha(t,x), E^\beta(t,x')\Big] = \delta^\beta_\alpha \delta(x,x')\,, \qquad \big[A_0(t,x), P^0(t,x')\big] = \delta(x,x')\,, \tag{3.56}$$

the other brackets vanishing.

The fact that the right hand side of the momentum P^0 does not contain any velocity implies that the Lagrangian is singular, according to definition (3.28); moreover it generates a primary constraint, i.e.

$$\phi := P^0 \approx 0\,. \tag{3.57}$$

Now, the canonical Hamiltonian can be calculated by performing the Legendre transformation, obtaining:[12]

$$H = \int d^3x \left[\frac{1}{2}E^\alpha E_\alpha + \frac{1}{4}F_{\alpha\beta}F^{\alpha\beta} - E^\alpha \partial_\alpha A_0\right]. \tag{3.58}$$

So, the primary total Hamiltonian results to be:

$$H_T = \int d^3x \left[\frac{1}{2}E^\alpha E_\alpha + \frac{1}{4}F_{\alpha\beta}F^{\alpha\beta} + A_0 \partial_\alpha E^\alpha + u\phi\right], \tag{3.59}$$

where we integrated by parts the third term. As a consistency check, we calculate the time derivative of the primary constraint P^0, we have:

$$\dot\phi = [\phi, H] = -\partial_\alpha E^\alpha\,, \tag{3.60}$$

so that we obtain a secondary constraint, namely

$$\chi := \partial_\alpha E^\alpha \approx 0\,. \tag{3.61}$$

The next consistency check does not generate any tertiary constraint, indeed

$$\dot\chi = [\chi, H] = 0\,. \tag{3.62}$$

[12] It is worth noting that the velocity $\dot A_\alpha$ can be rewritten as function of the momentum as $\dot A_\alpha = E_\alpha + \partial_\alpha A_0$, where we have taken into account the signature of the metric, i.e. $-,+,+,+$.

So, the theory generates one primary and one secondary constraint which form a set of first class constraints, i.e.

$$[\phi,\phi]=0\,,\quad [\phi,\chi]=0\,,\quad [\chi,\chi]=0\,. \tag{3.63}$$

Now, in order to get the most general physically possible motion, we write the extended Hamiltonian containing the secondary constraint as well, we have:

$$H_E=\int d^3x\left[\frac{1}{2}E^\alpha E_\alpha+\frac{1}{4}F_{\alpha\beta}F^{\alpha\beta}+A_0\partial_\alpha E^\alpha+u\phi+v\chi\right]. \tag{3.64}$$

The equations of motion can be easily obtained by calculating the Poisson bracket of the canonical variables with the extended Hamiltonian above; but, interestingly enough, given the structure of H_E, some simplifications are possible. In this respect, let us firstly, calculate the equation of motion of A_0, obtaining

$$\dot{A_0}=[A_0,H_E]=u\,, \tag{3.65}$$

revealing the nature of one of the ambiguities of the theory, which turns out to be the time derivative of A_0. More importantly, the above equations states that the evolution of A_0 is completely arbitrary. So, remembering the expression of the secondary constraint χ, we can reabsorb the variable A_0 in the definition of the arbitrary function v. Moreover, the momentum P^0 is constrained to vanish along all the evolution and its presence in the Hamiltonian only ensures that the variable A_0 is an arbitrary function. So, in order to simplify the expression of the extended Hamiltonian and reduce the number of unphysical degrees of freedom we can drop both the variables A_0 and P^0. It is worth stressing that the dynamics of the physically relevant degrees of freedom is not affected by this reduction. Finally, the total Hamiltonian turns out to be:

$$H_T=\int d^3x\left[\frac{1}{2}E^\alpha E_\alpha+\frac{1}{4}F_{\alpha\beta}F^{\alpha\beta}+v\partial_\alpha E^\alpha\right], \tag{3.66}$$

which can be used to calculate the canonical equations of motion, or to quantize the system by implementing the Dirac procedure described in § 6.1. We remark that the only survived constrain is the first class Gauss constraint $\partial_\alpha E^\alpha\approx 0$. A Gauss constraint appears in Yang–Mills gauge theories of non-abelian groups as well, in that case the ordinary derivative is replaced by a covariant derivative, so the connection field enters in the Gauss constraint, complicating its mathematical structure (see, e.g. (B.1)).

It is interesting to note that the Gauss constraint, in fact, generates gauge transformations. Let us, for example, consider the action of the smeared Gauss constraint, i.e.

$$G(f)=\int d^3x\, f(t,x)\partial_\alpha E^\alpha(t,x)\,, \tag{3.67}$$

on the fundamental variable A_α, we have:

$$\delta A_\alpha(t,x)=[A_\alpha(t,x),G(f)]=\partial_\alpha f(t,x)\,, \tag{3.68}$$

where the generic smearing function f plays the role of λ in (3.54).

3.4 Initial Value Formulation

We conclude this section giving an account of the fundamental theorems for establishing a well posed initial value formulation for gauge theories and GR. The results expressed by these theorems are pretty general, but as we will show below, they can easily applied to specific cases. As an example we will address the simple problem regarding the initial value formulation of Maxwell's electromagnetism. The case of electromagnetism is interesting because, being a gauge theory, shares with GR the presence of first class constraints and, as a consequence, the necessity to make a proper "gauge choice" in order to write the equations in a suitably form to face the initial value problem.

3.4.1 Some important theorems

First of all we enunciate the following theorem without demonstrating it, addressing the reader to [10] for a complete proof (a partial proof can be found in Wald's book [9]):

Theorem 3.6. *Let $(M, g_{\mu\nu})$ be a globally hyperbolic space-time (or a globally hyperbolic region of an arbitrary space-time) and let ∇_μ be any derivative operator. Let Σ be a smooth, space-like Cauchy surface. Consider the system of n linear equations for n unknown functions $\phi_1, \dots, \phi_n$ of the form*

$$g^{\mu\nu}\nabla_\mu\nabla_\nu\phi_i + \sum_j A^\mu{}_{ij}\nabla_\mu\phi_j + \sum_j B_{ij}\phi_j + C_i = 0\,, \tag{3.69}$$

namely a linear, diagonal, second order hyperbolic system. Then equation (3.69) has a well posed initial value formulation on Σ. More precisely, given arbitrary smooth initial data, $(\phi_i, n^\mu\nabla_\mu\phi_i)$ for $i = 1, \dots, n$ on Σ, there exists a unique solution of equation (3.69) throughout M. Furthermore, the solutions depend continuously on the initial data. Finally, a variation of the initial data outside of a closed subset, S, of Σ does not affect the solution in $D(S)$.

It is worth noting that the theorem explicitly refers to linear systems of equations, moreover, although there are few results concerning the initial value formulation for non-linear systems of equations, an important result exists concerning the so called quasi-linear, diagonal second order hyperbolic equations due to Leray in 1952 and is contained in the following theorem:

Theorem 3.7. *Let $(\phi_0)_1, \dots, (\phi_0)_n$ be any solution of a quasi-linear hyperbolic system of equations below*

$$g^{\mu\nu}(x; \phi_j, \nabla_\mu\phi_j)\nabla_\mu\nabla_\nu\phi_i = F_i(x; \phi_j, \nabla_\mu\phi_j)\,, \tag{3.70}$$

on a manifold M and let $(g_0)^{\mu\nu} = g^{\mu\nu}(x; (\phi_0)_j, \nabla_\mu(\phi_0)_j)$. Suppose $(M, (g_0)_{\mu\nu})$ is globally hyperbolic (or alternatively consider a globally hyperbolic region of this space-time). Let Σ be a smooth space-like Cauchy surface for $(M, (g_0)_{\mu\nu})$. Then, the initial value formulation of equation (3.70) is well posed on Σ in the following sense: for initial data on Σ sufficiently close to the initial data for $(\phi_0)_1, \dots, (\phi_0)_n$, there exists an open neighborhood O of Σ such that equation (3.70) has a solution $\phi_1, \dots, \phi_n$, in O and $(O, g_{\mu\nu}(x; \phi_j, \nabla_\mu\phi_j)$ is globally hyperbolic. The solution is unique in O and propagates causally in the sense that if the initial data for $\phi'_1, \dots, \phi'_n$ agree with that of $(\phi_0)_1, \dots, (\phi_0)_n$ on a subset S of Σ, then the solution agree on $O \cap D^+(S)$. Finally the solutions depend continuously on the initial data.

As before we do not give the demonstration of the theorem, which can be find in [10], together with some other interesting properties of the solutions. Let us remark that equation (3.70) differs from (3.69), because the Lorentzian metric field $g_{\mu\nu}$ is now allowed to depend on the unknown variables and their first derivatives, while the smooth functions F_i may have a non-linear dependence on these variables. An interesting recent application of Leray's theorem is the demonstration that a well posed initial value formulation can be formulated for the scalar-gravity coupled system as showed in [48].

3.4.2 Initial value formulation for the electromagnetic field

The strategy we want to follow in order to show that the Maxwell system of equations has a well posed initial value formulation should be now clear: It basically consists in recasting the equations in a form which can be traced back to those in line (3.69).

So let us start recalling the expression of the Maxwell system of equations in Minkowski space-time:

$$\partial_\mu \left(\partial^\mu A^\nu - \partial^\nu A^\mu\right) = 0. \tag{3.71}$$

Now we can easily split the background, fixing a one parameter family of hypersurfaces Σ_t parametrized by constant values of the inertial time t, in particular let us assume that $\Sigma_0 = \Sigma_{t=t_0}$ be our initial hypersurface. This procedure allows to emphasize a central feature of the Maxwell system of equations: the appearance of the so called Gauss constraint, due to the fact that the time component of the equation above does not contain any second time derivative term. Namely:

$$\nabla_i E^i = 0, \tag{3.72}$$

where

$$E^i = \delta^{ij}\nabla_j A^0 - \frac{\partial A^i}{\partial t}, \tag{3.73}$$

is the electric field. Thus equation (3.72) represents a constraint for the initial data $(A^\mu, \partial A^\nu/\partial t)$ on Σ_0. In other words, the choice of the initial data is not free, they must, in fact, satisfy the Gauss constraint, otherwise they cannot generate solutions of the Maxwell's equations. One could expect that differentiating the Gauss constraint with respect to time, an equation containing a second time derivative of the scalar potential A_0 can be obtained. But, as can be easily verified, the Bianchi identity $dF = 0$ (where $F = dA$ is the curvature 2-form associated to the electromagnetic field A and d is the exterior derivative operator) prevents from generating second order time derivatives of A_0. As a side remark, we stress that in the opposite case, the initial value problem would have a simple solution, at least in the sense expressed by the Cauchy-Kowalewski theorem:

Theorem 3.8. (Cauchy-Kowalewski theorem): *Let $(t, x^1, \ldots, x^m - 1)$ be coordinates of $\mathbb{R}^m$. Consider a system of n partial differential equations for n unknown functions $\phi_1, \ldots, \phi_n$ in $\mathbb{R}^m$, having the following form*

$$\frac{\partial^2 \phi_i}{\partial t^2} = F_i\left(t, x^a; \phi_j; \frac{\partial \phi_j}{\partial t}; \frac{\partial \phi_j}{\partial x^a}; \frac{\partial^2 \phi_j}{\partial t \partial x^a}; \frac{\partial^2 \phi_i}{\partial x^a \partial x^b}\right), \tag{3.74}$$

where each F_i is an analytic function of its variables. Let $f_i(x^a)$ and $g_i(x^a)$ be analytic functions. Then there is an open neighborhood O of the initial hypersurface $t = t_0$ such that within O there exists a unique analytic solution of equation (3.74) such that $\phi_i(t_0, x^a) = f_i(x^a)$ and $\frac{\partial \phi_i}{\partial t}(t_0, x^a) = g_i(x^a)$.

Unfortunately, it is not the case: the Maxwell system of equations (3.71) is under-determined, being equivalent to the dynamical equations for the spatial components of the electromagnetic field, which contain second time derivatives, plus a constraint equation for four unknown functions. But this feature is absolutely not unexpected. It simply reflects the presence of the $U(1)$ gauge freedom. In other words, the existing gauge freedom prevents the Maxwell equations from completely determining the potential A_μ, namely we should expect that they uniquely determine the potential up to a gauge transformation. From this perspective, it is easy to understand that the system of Maxwell equations admit a well posed initial value formulation only for the *physical* states of the theory. In fact, once realized how the gauge transformations enter into the game, the solution becomes simple.

In order to clarify this point, let us fix the Lorentz gauge, i.e.

$$\partial_\mu A^\mu = 0\,, \tag{3.75a}$$

which is particularly useful to treat this problem. The gauge choice allows to simplify the structure of Eq. (3.71), i.e.

$$\partial_\nu \partial^\nu A^\mu = 0\,. \tag{3.75b}$$

The system of equations are physically equivalent to (3.71). More precisely, solutions of the Eq. (3.71) can differ from those obtainable by solving the system of equations in (3.75a) and (3.75b) only by a gauge transformation. So, from a physical perspective, the dynamics is well described by Eqs. (3.75a) and (3.75b), with the remarkable advantage that now we can use the result of theorem (3.6).

Specifically, let us suppose that the initial data are chosen in such a way that they satisfy the Lorentz gauge condition in Eq. (3.75a) on the initial hypersurface Σ_0 (if they don't we can operate on them by a suitable gauge transformation), then, using equation (3.75b) and the Schwarz theorem, we can write:

$$\partial_\mu \partial^\mu \left(\partial_\nu A^\nu\right) = \partial_\nu \left(\partial_\mu \partial^\mu A^\nu\right) = 0\,. \tag{3.76}$$

Hence, provided that equation $\partial_\mu \partial^\mu A^\nu = 0$ is satisfied everywhere, as a consequence of theorem (3.6), also the gauge condition will be satisfied everywhere if and only if $\partial \partial_\mu A^\mu / \partial t = 0$ on Σ_0. It is worth explaining the role played here by the Gauss constraint. At a first glance the Gauss constraint seems to have disappeared; actually, it has only been written in a different form. In fact, we have required that $\partial \partial_\mu A^\mu / \partial t = 0$ on Σ_0, namely

$$\begin{aligned} 0 &= \frac{\partial}{\partial t}\, \partial_\mu A^\mu = \partial_t^2 A^0 + \partial_i \partial_t A^i \\ &= \partial_t^2 A^0 + \partial_i \left(E^i - \partial^i A^0\right) = \partial_i E^i\,, \end{aligned} \tag{3.77}$$

where we used equation (3.75b). It is not surprising at all that the Gauss constraint turns out to be encapsulated in the initial conditions, because the initial field configuration must satisfy not only the gauge condition, but also the constraint. In particular, its role is to assure that, if equation (3.75b) holds everywhere and provided that $\partial_i E^i = 0$ on Σ_0, the Lorentz gauge condition remains valid throughout all the evolution for the gauge transformed initial data.

It remains only to solve the dynamical equations (3.75b), with the given, suitably chosen, initial data. But now the problem is simple, the equations have in fact the desired form, i.e. that

required to apply the result of theorem (3.6), which indeed establishes the existence of a well posed initial data formulation. The last question is: can we conclude that the solutions are unique? We can briefly answer to this question supposing that the original systems of Maxwell equations (3.71) provides two solutions, S_1 and S_2, with the same initial conditions. By making a gauge transformation, it is possible to recast them into the solution of the equation (3.75b) with the same initial condition. But, since the solution of the system (3.75b) is unique, once assigned suitably initial conditions, then we conclude that S_1 and S_2 can differ at most by a gauge transformation. In other words they represent the same physical field configuration. Concluding, physically the solution is, in fact, unique.

4. Canonical General Relativity

As we have clarified above, the canonical constraints play a crucial role in the initial value formulation of a theory with gauge freedom. The same consideration is valid as far as we regard GR, which features a gauge symmetry correlated with the invariance under diffeomorphisms. Moreover, as will be clarified later, the canonical quantization of a system with gauge symmetry requires that the classical constraints are promoted to operators, acting as supplementary conditions on the quantum states of the theory. So, in view of discussing the problem of quantum gravity, in this section we calculate the constraints of GR.

Specifically, starting from the description of the so-called 3+1 splitting procedure, we finally arrive at writing the canonical equations of motion of GR, for the first time studied by Dirac [49] and then by Arnowitt, Deser and Misner [50, 51, 52, 53]. A brief digression on the initial value formulation in gravity will conclude the section.

4.1 3+1 splitting of space-time

Specifically, the splitting procedure allows to sort an evolution parameter out of the covariant general relativistic space-time. It is worth stressing that covariance is not lost in this formulation, even though it is no longer manifest as in the Lagrangian approach. As suggested by the Geroch theorem, by using a gauge transformation, which we refer to as embedding diffeomorphism, it is possible to "slice" a globally hyperbolic space-time, representing it as the evolution in "time" of 3-dimensional space. As any gauge transformation, the embedding diffeomorphisms does not affect the dynamical content of the theory, in other words the canonical Hamiltonian equations plus constraints (obviously!) are completely equivalent to the usual Einstein equations.

Technically, what we are going to do is the following: Starting from the Hilbert-Einstein action,

$$S_{H-E}(g) = \frac{1}{2}\int_M d^4x\sqrt{-g}R\,, \tag{4.1}$$

where $R = g_{\mu\rho}g_{\nu\sigma}R^{\mu\nu\rho\sigma}$ is the Ricci scalar curvature ($R_{\mu\nu\rho}{}^{\sigma}v_\sigma = \left[\nabla_\mu,\nabla_\nu\right]v_\rho$ being the Riemann curvature tensor), we restrict M to be a globally hyperbolic space-time, $\mathscr{M}$, so that, according to theorem (3.5), $\mathscr{M} = \mathbb{R}\times\Sigma^3$; Σ^3 being a compact three-dimensional manifold without boundary.

It is interesting to note that the restriction to globally hyperbolic space-times is a quite strong requirement, especially in view of the formalization of a quantum theory of gravity. Being aware

of this strong hypothesis, necessary to canonically quantize the theory, we expect that it could be relaxed once a rigorous formulation of the quantum theory will be at hand, as suggested in [3]. Unfortunately, so far and up to the author's knowledge, no rigorous prescription exists that allows to get rid of this hypothesis or consider different topologies directly in the quantum theory.

However, once the hypothesis of the Geroch theorem are satisfied, we can foliate $\mathscr{M}$ by Cauchy hypersurfaces $\Sigma^3_t \overset{def}{=} y_t(\sigma)$, in other words $\forall t \in \mathbb{R}, \exists$ a globally injective immersion (embeddings) $y_t : \sigma \to M$, defined by $y_t(x^i) = y(t,x^i)$, where $x^i \in \sigma$ are the coordinates over the hypersurfaces Σ^3. Hence, Σ^3_t represents a foliation of the manifold M parametrized by the continuous function t. Now let us denote n^μ the unit normal vector to the hypersurfaces Σ^3_t, so that the four-dimensional metric $g_{\mu\nu}$ induces a three-dimensional Riemannian metric $h_{\mu\nu}$ on each hypersurface:

$$h_{\mu\nu} = g_{\mu\nu} + n_\mu n_\nu \,. \tag{4.2}$$

The above relation is often referred as *first fundamental form* of Σ. Now, consider a vector field t^μ, called "deformation vector", satisfying the the following relation $t^\mu \nabla_\mu t = 1$. It generates a 1-parameter family of diffeomorphisms, $\phi_t : \mathbb{R} \times \sigma \to M$, defined as $(t,x) \to y^\mu(t,x) := y^\mu_t(x)$, called embedding diffeomorphisms. Geometrically, the deformation vector represents the "flow of time" throughout space-time, in other words it is the tangent vector to the "time line", namely the directional derivative it generates corresponds to an increment in label time t. Remarkably, the label time t does not correspond to physical time, the measure of which would imply the knowledge of the space-time metric, rather it is a mere label denoting the different Cauchy hypersurfaces.[13]

So, the embedding diffeomorphisms is completely arbitrary and can be usefully parametrized by decomposing the deformation vector in its normal and tangential components with respect to Σ_t. Specifically, by defining the "Lapse function" N and the "Shift vector" N^μ as

$$N = -n_\mu t^\mu = \left(n^\mu \nabla_\mu t\right)^{-1} , \tag{4.3a}$$

$$N_\mu = h_{\mu\nu} t^\nu , \tag{4.3b}$$

we have:

$$t^\mu(y) = \left(\frac{\partial y^\mu(t,x)}{\partial t}\right)|_{y(t,x)=y_t(x)} = N(y) n^\mu(y) + N^\mu(y) \,. \tag{4.4}$$

It is important to note that in order to generate a consistent foliation the Lapse function has to be monotonic.

At this point a brief digression is in order. It should be easy for the reader familiar with the canonical formalism to imagine that the lapse function and shift vector, being two completely arbitrary functions which parametrize the time flow, will turn out to be Lagrange multipliers. As regarding the 4-metric, four of its entries directly depend on them. So, only six out of the ten component of the metric are dynamical variables. This fact nicely reflects in the canonical theory, where the 3-metric, with its six components, turns out to be the fundamental dynamical variable. In other words, a globally hyperbolic space-time represents the time evolution of a Riemannian 3-metric field on a 3-dimensional abstract manifold [9], while the other four components express

[13] Note that if we do not fix the metric the lapse of time, say Δt, dividing two different spatial hypersurfaces, Σ^3_t and $\Sigma^3_{t+\Delta t}$, is completely general and correlated with the integral curve of t^μ, which generates the embedding diffeomorphism. So it is not referable to any real physical measurement, being correlated to gauge transformations.

only the arbitrariness we have in choosing the reference system [54]. It is worth recalling that the Einstein equations in vacuum

$$R_{\mu\nu} = 0 \tag{4.5}$$

are, in fact, six equations (not ten!) because of the Bianchi identity $\nabla_\mu R^\mu_{\ \nu} = \frac{1}{2}\partial_\nu R$), which relates some of the components of the Ricci tensor.

Let us now enter in the technical details of the canonical formulation of gravity, starting by enunciating the following Lemma, which represents an important "working tool".

Lemma 4.1. *Let $(M, g_{\mu\nu})$ be a space-time and let Σ be a smooth space-like hypersurface in M. Let $h_{\mu\nu}$ denote the induced metric on Σ and let D_μ denote the covariant derivative operator associated with the metric $h_{\mu\nu}$. Then the action of D_μ is given by the formula*

$$D_\rho T^{\nu_1\cdots\nu_n}{}_{\mu_1\cdots\mu_m} = h^{\nu_1}_{\ \alpha_1}\cdots h^{\nu_n}_{\ \alpha_n} h_{\mu_1}^{\ \beta_1}\cdots h_{\mu_m}^{\ \beta_m} h_\rho^{\ \sigma}\nabla_\sigma T^{\alpha_1\cdots\alpha_n}{}_{\beta_1\cdots\beta_m}\,, \tag{4.6}$$

where ∇_σ is the derivative operator associated with $g_{\mu\nu}$.

Proof. *It is simple to verify that D_μ satisfies the following properties: Linearity, Leibniz rule, Commutativity with contraction, Torsion free ($\forall f \in \mathscr{F}, \left[D_\mu, D_\nu\right] f = 0$) and acts as a directional derivative on scalar functions f. Moreover we have:*

$$D_\mu h_{\rho\sigma} = h_\rho^{\ \alpha} h_\sigma^{\ \beta} h_\mu^{\ \gamma}\nabla_\gamma\left(g_{\alpha\beta} + n_\alpha n_\beta\right)\,, \tag{4.7}$$

because $\nabla_\mu g_{\rho\sigma} = 0$ and $h_{\mu\nu}n^\nu = 0$. Thus D_μ is the unique derivative operator associated with $h_{\mu\nu}$. □

Let us define the *second fundamental form* of Σ^3, called extrinsic curvature:

$$K_{\mu\nu} = h_\mu^{\ \rho} h_\nu^{\ \sigma}\nabla_{(\rho} n_{\sigma)} = \frac{1}{2} h_\mu^{\ \rho} h_\nu^{\ \sigma}\pounds_n h_{\rho\sigma} = \frac{1}{2}\left(\pounds_n h\right)_{\mu\nu}\,, \tag{4.8}$$

where the symbol $\pounds_v$ denotes the Lie derivative with respect to the vector field v^μ. The extrinsic curvature is a spatial vector by definition and represents the parallel transport of the normal vector along the hypersurface Σ or the variation of the three-metric along the integral line of n^μ. We can easily rewrite the extrinsic curvature in order to make explicit the Lie time derivative of the 3-metric as we are going to show:

$$\begin{aligned} K_{\mu\nu} &= \frac{1}{2}\left(\pounds_n h\right)_{\mu\nu} = \frac{1}{2} h_\mu^{\ \rho} h_\nu^{\ \sigma}\left(n^\alpha\nabla_\alpha h_{\rho\sigma} + h_{\rho\alpha}\nabla_\sigma n^\alpha + h_{\alpha\sigma}\nabla_\rho n^\alpha\right) \\ &= \frac{1}{2N} h_\mu^{\ \rho} h_\nu^{\ \sigma}\left(Nn^\alpha\nabla_\alpha h_{\rho\sigma} + h_{\rho\alpha}\nabla_\sigma(Nn^\alpha) + h_{\alpha\sigma}\nabla_\rho(Nn^\alpha)\right) \\ &= \frac{1}{2N} h_\mu^{\ \rho} h_\nu^{\ \sigma}\pounds_{(t-N)} h_{\rho\sigma} = \frac{1}{2N}\left(\dot{h}_{\mu\nu} - 2D_{(\mu}N_{\nu)}\right)\,, \end{aligned} \tag{4.9}$$

where we defined $\dot{h}_{\mu\nu} = h_\mu^{\ \rho} h_\nu^{\ \sigma}\pounds_t h_{\rho\sigma}$ and we used equation (4.7).

Once defined the covariant derivative operator (4.7), we can define the curvature tensor of the Cauchy surface Σ as usual:

$${}^{(3)}R_{\mu\nu\rho}^{\ \ \ \sigma}\omega_\sigma = D_\mu D_\nu\omega_\rho - D_\nu D_\mu\omega_\rho\,. \tag{4.10}$$

Now using the prescription to express the covariant derivative on the 3-dimensional manifold as projection of the 4-dimensional derivative operator we have:

$$\begin{aligned} D_\mu D_\nu \omega_\rho &= D_\mu \left(h_\nu{}^\sigma h_\rho{}^\tau \nabla_\sigma \omega_\tau \right) = h_\nu{}^\alpha h_\rho{}^\beta h_\mu{}^\gamma \nabla_\gamma \left(h_\alpha{}^\sigma h_\beta{}^\tau \nabla_\sigma \omega_\tau \right) \\ &= h_\nu{}^\sigma h_\rho{}^\tau h_\mu{}^\gamma \nabla_\gamma \nabla_\sigma \omega_\tau + h_\nu{}^\alpha h_\rho{}^\beta h_\mu{}^\gamma \nabla_\gamma \left(h_\alpha{}^\sigma h_\beta{}^\tau \right) \nabla_\sigma \omega_\tau \\ &= h_\nu{}^\sigma h_\rho{}^\tau h_\mu{}^\gamma \nabla_\gamma \nabla_\sigma \omega_\tau + h_\rho{}^\tau K_{\nu\mu} n^\sigma \nabla_\sigma \omega_\tau + h_\nu{}^\sigma K_{\mu\rho} n^\tau \nabla_\sigma \omega_\tau \,, \end{aligned} \tag{4.11}$$

where we used the expression of the 3-metric as function of the 4-metric and the normal vector (4.2).

At last, we have all the elements to write down the relation between the 3-dimensional curvature tensor as function of the 4-dimensional Riemann tensor and extrinsic curvature:

$$\begin{aligned} {}^{(3)}R_{\mu\nu\rho}{}^\sigma \omega_\sigma &= D_\mu D_\nu \omega_\rho - D_\nu D_\mu \omega_\rho = 2 D_{[\mu} D_{\nu]} \omega_\rho \\ &= 2 h_{[\mu}{}^\gamma h_{\nu]}{}^\sigma h_\rho{}^\tau \nabla_\gamma \nabla_\sigma \omega_\tau + 2 h_\rho{}^\tau K_{[\nu\mu]} n^\sigma \nabla_\sigma \omega_\tau - 2 K_{[\mu|\rho|} K_{\nu]}{}^\sigma \omega_\sigma \\ &= h_\mu{}^\gamma h_\nu{}^\beta h_\rho{}^\tau h^\sigma{}_\alpha R_{\gamma\beta\tau}{}^\alpha \omega_\sigma - K_{\mu\rho} K_\nu{}^\sigma \omega_\sigma + K_{\nu\rho} K_\mu{}^\sigma \omega_\sigma \,, \end{aligned} \tag{4.12}$$

where, in the second line, we used $h_\nu{}^\sigma K_{\mu\rho} n^\tau \nabla_\sigma \omega_\tau = h_\nu{}^\sigma K_{\mu\rho} \nabla_\sigma \left(n^\tau \omega_\tau \right) - h_\nu{}^\sigma K_{\mu\rho} \omega_\tau \nabla_\sigma n^\tau = K_{\mu\rho} K_\nu{}^\tau \omega_\tau$ and $K_{[\mu\nu]} = 0$. Now, considering that ω_μ is a common factor, we can write down the *first Gauss-Codacci relation*:

$${}^{(3)}R_{\mu\nu\rho}{}^\sigma = h_\mu{}^\gamma h_\nu{}^\beta h_\rho{}^\tau h^\sigma{}_\alpha R_{\gamma\beta\tau}{}^\alpha - K_{\mu\rho} K_\nu{}^\sigma + K_{\nu\rho} K_\mu{}^\sigma \,. \tag{4.13}$$

With an analogous procedure, we can also easily obtain the *second Gauss-Codacci relation*:

$$D_\mu K^\mu{}_\nu - D_\nu K^\mu{}_\mu = R_{\mu\rho} n^\rho h^\mu{}_\nu \,. \tag{4.14}$$

At this point, once we realize that

$$R_{\mu\nu\rho\sigma} h^{\mu\rho} h^{\nu\sigma} = R_{\mu\nu\rho\sigma} \left(g^{\mu\rho} + n^\mu n^\rho \right) \left(g^{\nu\sigma} + n^\nu n^\sigma \right) = R + 2 R_{\nu\sigma} n^\nu n^\sigma \,, \tag{4.15}$$

we can write the Ricci scalar as

$$R = R_{\mu\nu\rho\sigma} h^{\mu\rho} h^{\nu\sigma} - 2 R_{\nu\sigma} n^\nu n^\sigma \tag{4.16}$$

and from the first Gauss-Codacci relation we have:

$$R = {}^{(3)}R + (K^\mu{}_\mu)^2 - K_{\mu\nu} K^{\mu\nu} - 2 R_{\nu\sigma} n^\nu n^\sigma \,. \tag{4.17}$$

Moreover,

$$\begin{aligned} R_{\nu\sigma} n^\nu n^\sigma &= R_{\nu\rho\sigma}{}^\rho n^\nu n^\sigma = -n^\nu \left(\nabla_\nu \nabla_\rho - \nabla_\rho \nabla_\nu \right) n^\rho \\ &= (K^\mu{}_\mu)^2 - K_{\mu\nu} K^{\mu\nu} - \nabla_\mu \left(n^\mu \nabla_\rho n^\rho \right) + \nabla_\rho \left(n^\mu \nabla_\mu n^\rho \right) \,. \end{aligned} \tag{4.18}$$

Therefore, being $\sqrt{-g} = N\sqrt{h}$, from equations (4.17) and (4.18) we finally obtain the following action for the gravitational field (having dropped the last term in the last line above, which being a total divergence, does not affect the equations of motion):

$$S_{3+1}(N, N^\mu, h_{\mu\nu}) = \frac{1}{2} \int_{\mathbb{R}\times\sigma} dt d^3x N \sqrt{h} \left({}^{(3)}R + K_{\mu\nu} K^{\mu\nu} - (K^\mu{}_\mu)^2 \right) \,. \tag{4.19}$$

It is particularly useful to pull back spatial tensors to the hypersurface Σ^3. This can be easily done by suitably defining the projectors $X^\mu_\alpha = (\partial y^\mu(t,x)/\partial x^\alpha)_{|y_t(x)=y(t,x)}$, where x^α are spatial coordinates on Σ^3. In this respect, let us define

$$h_{\alpha\beta} = h_{\mu\nu} X^\mu_\alpha X^\nu_\beta = g_{\mu\nu} X^\mu_\alpha X^\nu_\beta \,, \tag{4.20a}$$

$$K_{\alpha\beta} = K_{\mu\nu} X^\mu_\alpha X^\nu_\beta \,, \tag{4.20b}$$

so that we can rewrite the extrinsic curvature as

$$K_{\alpha\beta} = \frac{1}{2N}\left(\dot{h}_{\alpha\beta} - 2D_{(\alpha}N_{\beta)}\right) , \tag{4.21}$$

where only spatial indexes appear.

The Lagrangian contains the time derivatives of the 3-metric field through the terms depending on the extrinsic curvature, while no time derivatives of the lapse function and the shift vectors appear. Hence, the Lagrangian is singular and we expect that the theory generates four primary constraints as, in fact, we are going to demonstrate.

The next step in the canonical analysis is the definition of the conjugate momenta. Once the space-time has been split, we recall that the fundamental variables are N, N^α and $h_{\beta\gamma}$, the conjugate momenta of which respectively are

$$p^{(N)} = \frac{\partial \mathscr{L}_{3+1}}{\partial \dot{N}} = 0 \,, \tag{4.22a}$$

$$p^{(N)}_\alpha = \frac{\partial \mathscr{L}_{3+1}}{\partial \dot{N}^\alpha} = 0 \,, \tag{4.22b}$$

$$p^{\alpha\beta} = \frac{\partial \mathscr{L}_{3+1}}{\partial \dot{h}_{\alpha\beta}} = \frac{\sqrt{h}}{2}\left(K^{\alpha\beta} - h^{\alpha\beta}K\right) . \tag{4.22c}$$

So, the phase space is twenty dimensional and coordinatized by the set $N, N^\alpha, h_{\beta\gamma}, p^{(N)}, p^{(N)}_\alpha, p^{\beta\gamma}$ and equipped with the following symplectic structure:

$$\left\{N(t,x), p^{(N)}(t,x')\right\} = \delta(x,x') \,, \tag{4.23a}$$

$$\left\{N^\alpha(t,x), p^{(N)}_\beta(t,x')\right\} = \delta^\alpha_\beta \delta(x,x') \,, \tag{4.23b}$$

$$\left\{h_{\alpha\beta}(t,x), p^{\gamma\delta}(t,x')\right\} = \delta^{\gamma\delta}_{\alpha\beta}\delta(x,x') \,, \tag{4.23c}$$

where the symbol $\{\cdots,\cdots\}$ denotes the Poisson brackets, while $\delta^{\gamma\delta}_{\alpha\beta} = \frac{1}{2}\left(\delta^\gamma_\alpha\delta^\delta_\beta - \delta^\gamma_\beta\delta^\delta_\alpha\right)$.

From the definition of conjugate momenta, we immediately obtain four primary constraints as expected, i.e.

$$\mathscr{C}^{(N)} := p^{(N)} \approx 0 \,, \tag{4.24a}$$

$$\mathscr{C}^{(N)}_\alpha := p^{(N)}_\alpha \approx 0 \,. \tag{4.24b}$$

Now, we can perform the Legendre transformation. Because of the presence of primary constraints, we cannot re-express all the velocities as functions of the fundamental variables and their conjugate momenta. This implies that the Hamiltonian, usually defined as

$$H = \sum_i p_i \dot{q}_i - \mathscr{L} \,, \tag{4.25}$$

where q_i and p_k are the generalized coordinates on the phase space, is not uniquely determined as function of the fundamental variables and momenta. In other words, because of the presence of the primary constraints, the Hamiltonian is well defined only on a restricted region of the phase space determined by the primary constraints. So that, in order to take into account the restriction of the phase space implied by the four constraints (4.24), we have to introduce four Lagrange multipliers λ and λ^α, which have to be varied independently in the action. We have,

$$S_{3+1}(N,p_N,N^\alpha,p^{(N)}_\beta,h_{\gamma\delta},p^{\alpha\beta}) = \frac{1}{2}\int\limits_{\mathbb{R}\times\sigma} dtd^3x\Big[p^{\alpha\beta}\pounds_t h_{\alpha\beta} + p_N\dot{N} + p^{(N)}_\gamma\dot{N}^\gamma \\ -NH - N^\alpha H_\alpha - \lambda\mathscr{C}^{(N)} - \lambda^\beta\mathscr{C}^{(N)}_\beta\Big]\,, \tag{4.26}$$

where the super-Hamiltonian H and super-momentum H^α are defined as:

$$H = \frac{1}{\sqrt{h}}\left[p_{\alpha\beta}p^{\alpha\beta} - \frac{1}{2}\left(h_{\alpha\beta}p^{\alpha\beta}\right)^2\right] - \sqrt{h}^{(3)}R = \frac{1}{2}G_{\alpha\beta\gamma\delta}p^{\alpha\beta}p^{\gamma\delta} - \sqrt{h}^{(3)}R\,, \tag{4.27a}$$

$$H^\gamma = -2D_\delta p^{\gamma\delta}\,. \tag{4.27b}$$

Above we have introduced the so called super-metric $G_{\alpha\beta\gamma\delta} = \frac{1}{\sqrt{h}}\left(h_{\alpha\gamma}h_{\beta\delta} + h_{\alpha\delta}h_{\beta\gamma} - h_{\alpha\beta}h_{\gamma\delta}\right)$ and dropped the total divergence $2D_\alpha(h^{-1/2}N_\beta p^{\alpha\beta})$.

Finally, we write down the canonical Hamiltonian for the gravitational field,

$$\mathscr{H}_{\text{can}} = \int\limits_{\sigma^3} d^3x\left[NH + N^\alpha H_\alpha + \lambda\mathscr{C}_{(N)} + \lambda^\beta\mathscr{C}^{(N)}_\beta\right]\,, \tag{4.28}$$

and go on to discuss the dynamics.

4.2 Canonical constrained dynamics

This Section is devoted to the study of the constrained dynamics of the gravitational field. Namely, starting from the split action (4.26), we are going to calculate the Hamiltonian equations of motion. At the end, a discussion about the formulation of a well posed initial value problem will follow, in correlation also with a well known issue of canonical quantum gravity, the so-called problem of time. In GR, in fact, once suitable initial conditions have been assigned and the gauge fixed, the time evolution is uniquely determined and depends continuously on the initial data; but, as remarked more than once the t parameter appearing in the equations of motion does not represent physical time, rather it is a label denoting the different Cauchy hypersurfaces. The fact that "time" evolution is a symmetry of the theory induces to say that in GR "time does not exist". From the canonical point of view this reflects on the fact that the Hamiltonian is constrained to weakly vanish, in other words it does not generate time translations, but gauge transformations (on the classical trajectories). The concept of time evolution can be reintroduced in such a "frozen" formalism through a physical procedure, consisting in associating the time evolution to the relational dynamics of other dynamical fields coupled to gravity. We consider this aspect of GR extremely natural, because the theory is describing the dynamics of the space-time itself and not the dynamics of a field on a background. Moreover, from a quantum perspective, the frozen formalism does not affect the interpretation of the outcomes of the theory, which describes the actual quantum state of

space-time and, eventually, the transition from one quantum configuration to another. From this perspective, the classical idea of evolution in time has to be completely abandoned, as the concept of classical trajectories has to be abandoned in describing the quantum transitions of an electron in an atom.

Having in mind the above premise, let me start the canonical analysis by varying action (4.26) with respect to the Lagrange multipliers λ and λ^k, thus obtaining the primary constraints (4.24). In order to guarantee that the dynamics of the system is consistent we have to require that the constraints be preserved during the evolution, namely that the Poisson brackets of the constraints with the Hamiltonian vanish. We have,

$$\dot{\mathscr{C}}^{(N)}(t,x) = \left\{ \mathscr{H}_{\text{can}}, \mathscr{C}^{(N)}(t,x) \right\} = H(t,x)\,, \tag{4.29a}$$

$$\dot{\mathscr{C}}^{(N)}_{\alpha}(t,x) = \left\{ \mathscr{H}_{\text{can}}, \mathscr{C}^{(N)}_{\alpha}(t,x) \right\} = H_{\alpha}(t,x)\,, \tag{4.29b}$$

hence a set of secondary constraints have to be imposed, i.e.

$$H(t,x) \approx 0\,, \tag{4.30a}$$

$$H_{\alpha}(t,x) \approx 0\,, \tag{4.30b}$$

for all $x \in \Sigma$. The above weak equations are called super-Hamiltonian and super-momentum constraints and generate the following algebra

$$\left\{ H_{\alpha}(t,x), H_{\beta}\left(t,x'\right) \right\} = H_{\beta}(t,x)\partial_{\alpha}\delta\left(x,x'\right) - H_{\alpha}(t,x)\partial_{\beta}\delta\left(x',x\right)\,, \tag{4.31a}$$

$$\left\{ H_{\alpha}(t,x), H\left(t,x'\right) \right\} = H(t,x)\,\partial_{\alpha}\delta\left(x,x'\right)\,, \tag{4.31b}$$

$$\left\{ H(t,x), H\left(t,x'\right) \right\} = H^{\beta}(t,x)\,\partial_{\beta}\delta\left(x,x'\right) - H^{\beta}\left(t,x'\right)\partial_{\beta}\delta\left(x',x\right)\,. \tag{4.31c}$$

The algebra above reveals that the super-Hamiltonian and super-momentum constraints are the generators of diffeomorphisms: It is worth noting that differently from the usual Yang-Mills gauge theories, the algebra of the constraints has structure functions instead of structure constants. Furthermore, the above relations prevent from the emergence of tertiary constraints.

The fact that the Poisson brackets between the whole set of constraints weakly vanish indicate that the super-Hamiltonian and super-momentum together with the primary constraints form a set of first class constraints. Interestingly enough, this set can be easily reduced, by taking into account that the primary constraints can be strongly satisfied by considering the lapse function N and the shift vector N^{α} as Lagrange multipliers. In other words, let us consider the dynamical equations for N and N^{α}, i.e.

$$\dot{N}(t,x) = \lambda(t,x)\,, \tag{4.32a}$$

$$\dot{N}^{\alpha}(t,x) = \lambda^{\alpha}(t,x)\,, \tag{4.32b}$$

so, as we can immediately understand, the evolution of the Lapse function and the Shift vector is completely arbitrary, being their time derivatives related to the Lagrange multipliers λ and λ^{α}, which are unspecified functions. Henceforth, we can consider the Lapse function and the Shift vector themselves as Lagrange multipliers, so that their momenta identically vanish satisfying the primary constraints (4.24) and, simultaneously, reducing the dimensions of the phase space, which

is now coordinatized by the remaining variables $h_{\alpha\beta}$ and $p^{\gamma\delta}$. The dynamical equations for $h_{\alpha\beta}$ and $p^{\gamma\delta}$ can be directly calculated from the reduced Hamiltonian

$$\mathscr{H} = \int_{\Sigma^3} d^3x \left[NH + N^\alpha H_\alpha \right], \tag{4.33}$$

obtaining:

$$\dot{h}_{\alpha\beta} = 2N G_{\alpha\beta\gamma\delta} p^{\gamma\delta} + 2\nabla_{(\alpha} N_{\beta)}, \tag{4.34a}$$

$$\begin{aligned} \dot{p}^{\alpha\beta} =& \frac{1}{2\sqrt{h}} N h^{\alpha\beta} \left(p^{\gamma\delta} p_{\gamma\delta} - \frac{1}{2} p^2 \right) - \frac{2N}{\sqrt{h}} \left(p^{\alpha\gamma} p_\gamma{}^\beta - \frac{1}{2} p p^{\alpha\beta} \right) - N\sqrt{h} \left({}^{(3)}R^{\alpha\beta} - \frac{1}{2} {}^{(3)}R h^{\alpha\beta} \right) \\ &+ \sqrt{h} \left(\nabla^\alpha \nabla^\beta N - h^{\alpha\beta} \nabla^\gamma \nabla_\gamma N \right) - 2 p^{\gamma(\alpha} \nabla_\gamma N^{\beta)} + \nabla_\gamma \left(N^\gamma p^{\alpha\beta} \right), \end{aligned} \tag{4.34b}$$

where we ignored all the boundary terms and used equation (4.30b). The system of equations (4.30a), (4.30b), (4.34a) and (4.34b) is equivalent to the vacuum Einstein equations $R_{\mu\nu} = 0$. As remarked above, the super-Hamiltonian and super-momentum constraints are first class (see Eqs. (4.31)) and reflect the gauge invariance of the theory. They, in fact, provided that the spatial equations of motion are satisfied, generate a diffeomorphisms flow on the phase space according to the following identifications:

$$\{ H(N), \dots \} = \pounds_{Nn^\mu} (\dots), \tag{4.35a}$$

$$\left\{ \vec{H}(\vec{N}), \dots \right\} = \pounds_{N^\mu} (\dots), \tag{4.35b}$$

where we used the following notation:

$$H(N) = \int_\sigma d^3x N(t,x) H(t,x) \quad \text{and} \quad \vec{H}(\vec{N}) = \int_\sigma d^3x N^i(t,x) H_i(t,x). \tag{4.36}$$

In particular, the super-momentum or vectorial constraint is clearly correlated with spatial diffeomorphisms and can be satisfied by introducing the Wheeler superspace.[14] The super-Hamiltonian or scalar constraint, instead, represents a serious obstacle toward the canonical quantization of the gravitational field. As we remarked before, it generates diffeomorphisms along the normal vector to the Cauchy hypersurfaces, provided that the spatial equations of motion are satisfied. Interestingly enough, as noted by Wald and Kuchař, the scalar constraint is strictly analogous to the constraint coming out when one tries to parametrize an original non-constrained theory on fixed background. More specifically, an analogous constraint crops up when one introduces within the Lagrangian a time function which labels the hypersurfaces Σ_t, starting from an initial hypersurface Σ_0 and then treats this "time function" as a dynamical variable [9, 55]. But, as one can easily verify in the case of the point particle in flat space-time [56], the parametrized theory is linear in the momentum conjugate to the time function, thus the theory can be easily deparametrized by solving the constraint

[14] The Wheeler superspace is the space of the 3-metrics modulo 3-diffeomorphisms. Namely, two metric fields related by a spatial diffeomorphisms represent the same point on the Wheeler superspace).

with respect to this momentum. The scalar constraint of GR is, in stead, quadratic in the momenta, therefore such a deparametrization seems not to be possible, at least in pure gravity.[15]

In order to understand how many physical degrees of freedom are described by the theory, let us count the number of independent variables. Obviously, the presence of first class constraints associated with the gauge freedom of the theory, indicates that there is more than one set of canonical variables which correspond to a particular physical state. In other words, a physical state is well described by two or more different sets of canonical variables if they are correlated by a gauge transformation. Or, we can rephrase saying, a physical state is represented by the class of equivalence of canonical variables under the symmetries of the theory. Usually, in order to eliminate such an ambiguity in the description of the physical system, a set of gauge condition are imposed on the canonical variables. The gauge conditions are chosen *ad hoc* and are, sometimes, suggested by the mathematical or physical structure of the theory, but they are not a consequence of the theory. They are external conditions, nevertheless, removing only the unphysical degrees of freedom of the theory, are completely admissible, on condition that they satisfy some consistency requirements. Firstly, they have to be accessible, i.e. it must exist a transformation which maps the original set of variables to the set satisfying the gauge condition. Secondly, we have to require that the gauge condition be preserved by the symmetry flow. The two requirements above imply that the number of gauge conditions has to be equal to the number of first class constraints in order to completely fix the gauge. Moreover, the determinant of the square matrix constructed with the Poisson brackets between the gauge conditions and the first class constraints has to be different from zero. Remarkably, this is exactly the definition of a set of second class constraints, which can be in principle solved through the Dirac procedure [49]. At this point the count of the number of physical degrees of freedom is easy. We have to subtract to the number of canonical variables the total number of second class constraints, or, according to what said before, the number of first class constraints plus the number of gauge conditions plus the second class constraints not coming from the gauge fixing. In other words, we have to subtract to the number of canonical variables twice the number of first class constraints plus independent second class constraints. It is worth noting that in the present case the number of physical degrees of freedom in the phase space is four, which correspond to the two polarization of the graviton in the configuration space.

Concluding, we can say that the symmetry group of GR is well implemented in the canonical formalism, which is in this sense generally covariant, even though the covariance of the theory is not manifest as in the Lagrangian formulation. We want also to stress the importance of the invariance under diffeomorphisms, pointing out that every possible observable for this theory must be invariant under this group of symmetries of the action. But the meaning of this statement goes over the usual meaning it has in Yang-Mills gauge theories, because the request of 4-diffeomorphisms invariance involves also the dynamics, therefore the definition of an observable is not only a kinematical problem, it necessarily implies to solve the dynamics. In other words in GR kinematics and dynamics are inextricably bound. Quoting a famous sentence by J. Stachel, we can say: "*There is no kinematics without dynamics*".

[15] In the case matter fields are introduced in the dynamics, a deparametrization is, in fact, possible [57] (see also [27, 28, 30]). In this framework, the evolution of the physical system can be interpreted in terms of relational variables [14, 25] and the so-called problem of time can be solved, by referring the evolution to the dynamical "relations" between distinguished fields.

4.3 Initial value formulation for gravity

As we have shown above, the Hamiltonian of GR is not a true Hamiltonian, but a linear combination of constraints. In particular, if we assume that the spatial components of the Einstein equations are satisfied, then it generates a flow along the integral curve of the deformation vector, rather than time translations. Nevertheless, the Cauchy problem can be well posed until the appearance of a singularity affects the consistency of the theory itself.

In order to construct a parallel with the case of electromagnetism, it is worth recalling that the Gauss constraint comes out from the time component of the Maxwell equations, which does not contain second time derivatives. The same consideration can is valid for Einstein equations in vacuum. In fact, the equations $G_{\mu\nu}n^\nu = 0$ do not contain second time derivatives of anyone of the metric component, namely, as in the electromagnetic case, these equations depend only on the initial data: In other words, they represent a restriction on the possible acceptable initial data set. So, we expect that the canonical constraints be contained in the equations $G_{\mu\nu}n^\nu = 0$, indeed we have:

$$G_{\mu\nu}n^\mu n^\nu = -\frac{H}{2\sqrt{h}} = 0\,, \tag{4.37a}$$

$$G_{\mu\nu}n^\mu e^\nu{}_i = \frac{H_i}{2\sqrt{h}} = 0\,. \tag{4.37b}$$

Therefore, the constraints equations are actually equivalent to four of the Einstein dynamical equations; furthermore the Bianchi identity $\nabla_\mu G^{\mu\nu} = \nabla_\mu R^{\mu\nu} - 1/2\partial^\nu R = 0$, together with the equations of motion for the spatial components, implies that the constraints are *involutive*. Namely, provided that the super-Hamiltonian and super-momentum constraints are satisfied on the initial Cauchy surface and the equations of motion are satisfied everywhere, then also the constraints are satisfied along the evolution. A very simple argument allows to show what just claimed. Assuming that we have already solved the equations for the spatial components of the gravitational field, then the Bianchi identity represents a relation between the time derivative of the normal components of the Einstein tensor $G_{\mu\nu}n^\nu$ and the non-time differentiated components of $G_{\mu\nu}$ and their spatial derivatives. Now, by pulling back equation $\nabla_\mu G^{\mu\nu} = 0$ on the solution for the spatial components of the Einstein equations and realizing that the spatial part of $G_{\mu\nu}$ vanishes, the Bianchi identity becomes a linear homogeneous system of four first order partial equations for the four unknown functions $G_{\mu\nu}n^\nu$. Then, it follows that if $G_{\mu\nu}n^\nu$ vanish on the initial slice, they must vanish on any slice.

From the Lagrangian point of view the vanishing of four of the Einstein vacuum equations could appear as an under-determination of the components of the gravitational field, instead it just reflect the invariance of the theory under the group of diffeomorphisms, as we have already explained above. In other words, the apparent under-determination is not a physical one, exactly as in the electrodynamics case. In fact, if $\phi : M \to M$ is a diffeomorphisms and $(M, g_{\mu\nu})$ is a solution of the Einstein equations, then $(M, \phi^* g_{\mu\nu})$ is a solution too. So, the number of the "non-gauge" components of the gravitational field is exactly six, as the number of the dynamical equations.

Now, in analogy with the electromagnetic case, we fix the gauge: since the gauge freedom of GR regards the general coordinates transformation, then choosing a gauge means to fix a particular system of coordinates. A suitable choice is the "harmonic" gauge, characterized by a system of

coordinates satisfying the following equations:

$$\nabla_\mu \nabla^\mu y^\rho = 0\,. \tag{4.38}$$

This choice does not affect the generality of the procedure. Specifically, in a neighborhood of a portion of space-time covered by an original set of coordinates, say x^μ, we can proceed as follows: firstly, let us note that equation (4.38) has the form of equation (3.69), therefore, assuming as initial data the old set x^μ and its derivative $\nabla_\mu x^\nu$, we can uniquely solve equation (4.38) in a neighborhood of the portion of space-time covered by the old set of coordinates. Since $\nabla_\mu x^\nu$ are linearly independent, then also $\nabla_\mu y^\rho$ are linearly independent and consequently the set y^ρ will provide a local coordinates system. The choice of the harmonic coordinates will result in the following equation:

$$0 = \nabla_\mu \nabla^\mu y^a = \frac{1}{\sqrt{-g}}\partial_\mu\left(\sqrt{-g}g^{\mu\nu}\partial_\nu y^a\right) = \frac{1}{\sqrt{-g}}\left(\partial_\mu\sqrt{-g}g^{\mu a}\right) = \partial_\mu g^{\mu a} + \frac{1}{2}g^{\mu a}g^{\rho\sigma}\partial_a g_{\rho\sigma}\,. \tag{4.39}$$

Furthermore, we have that the Ricci tensor $R_{\mu\nu}$ can be written as

$$R_{\mu\nu} = -\frac{1}{2}g^{\rho\sigma}\left[-2\partial_\sigma\partial_{(\nu}g_{\mu)\rho} + \partial_\sigma\partial_\rho g_{\mu\nu} + \partial_\mu\partial_\nu g_{\rho\sigma}\right] + F_{\mu\nu}\left(g,\partial g\right)\,. \tag{4.40}$$

the above expression emphasizes that the Ricci tensor is, in fact, linear in the second derivatives of the metric tensor, where $F_{\mu\nu}\left(g,\partial g\right)$ contains the non-linear dependence on the metric and its derivatives. So that, considering both Eq. (4.39) and (4.40), we can isolate, in the Einstein equations, the non-linear dependence on the metric and its derivatives, obtaining [58]:

$${}^{(h)}R_{\mu\nu} = R_{\mu\nu} + g_{\rho(\mu}\partial_{\nu)}\nabla_\sigma\nabla^\sigma y^\rho = -\frac{1}{2}g^{\rho\sigma}\partial_\rho\partial_\sigma g_{\mu\nu} + \tilde{F}_{\mu\nu}(g,\partial g) = 0\,. \tag{4.41}$$

Therefore, the Einstein equations in vacuum are equivalent to the system of equations (4.41) (generally referred as reduced Einstein equations) and (4.38). The reduced Einstein equations are suitable to apply the result of the Leray's theorem (3.7).

In this respect, let $h_{\alpha\beta}$ and $K_{\gamma\delta}$ be the metric and extrinsic curvature of the hypersurface Σ^3 and let us assume that they satisfy the constraints (4.37). Then, after having chosen a suitable system of coordinates over a portion of the hypersurface Σ^3, let me assign the following initial data for the metric and its time derivative $(h^0_{\alpha\beta}, \dot{h}^0_{\alpha\beta})$, such that the extrinsic curvature $K_{\alpha\beta}$ on Σ^3 results from this choice via equation (4.21). Since the Einstein equations involve all the ten components of the metric field, we have to give the initial values of g_{00} and $g_{0\alpha}$ too. A very simple choice could be $g_{00} = -1$, $g_{0\alpha} = 0$ and, as a consequence $K_{\alpha\beta} = \frac{1}{2}\dot{h}_{\alpha\beta} = \frac{1}{2}\dot{g}_{\alpha\beta}$. Now the time derivative of the "0" components of the metric tensor $\partial g_{0\mu}/\partial t$ remains undetermined by this choice, but they can be fixed via the gauge fixing condition $\nabla_\mu\nabla^\mu y^a = 0$ on Σ^3. In the canonical formalism this means that we initially choose the value of the Lapse function $N = -1$ and Shift vector $N^i = 0$, then via the gauge fixing we assign the time evolution of these two geometrical objects. Interestingly enough, by assigning the functional form of the Lapse and Shift one can completely fix the gauge, i.e. the reference system.

Now, let us suppose that the chosen initial conditions are sufficiently near that of flat space-time, then, according to theorem (3.7), we can solve equations (4.41) in a neighborhood of that

portion of Σ^3 covered by the original set of coordinates. Moreover, Σ^3 will be a Cauchy surface of the globally hyperbolic space-time generated by the solution.

Furthermore, it is possible to demonstrate that a solution of equations (4.41) will be a solution of the vacuum Einstein equations in a neighborhood of a portion of space-time where the condition (4.38) holds. This concludes this brief digression about the initial value problem in gravity. The interested reader can find the demonstration of the last statement in [9].

Concluding we can say that we can give a prescription to demonstrate that it exists, at least locally, a solution of the Einstein equations, moreover the solution depends continuously on the initial data and the space-time it generates is globally hyperbolic. This demonstration is based on the assumption that the set of initial data is near to that of flat space, but this requirement can be relaxed using a trick. The idea is that any curved space if observed from a sufficiently small scale appears nearly flat, so the trick consists simply in rescaling the initial data metric function via a coordinate transformation if they did not appear sufficiently flat (for details see [9]).

5. Ashtekar Canonical Gravity

As is well known, the program of canonical quantization is not a rigid algorithm and can be slightly adapted to the theory one is going to construct. In fact, the general program of quantization of classical systems requires to make choices in different steps of the quantization procedure, as is briefly described in the next section. Generally speaking, one could say that the achievement of the desired result, namely the construction of a rigorous quantum theory, depends on the choices made in the different steps of the canonical procedure. The structure of the theory could result remarkably simplified if a smart choice of variables were done, allowing to consistently reduce the difficulties one has to face in the quantization procedure. In other words, a smart choice of fundamental variables could make the theory manageable in view of quantization.

To be more specific, the introduction of the Ashtekar self-dual $SL(2,C)$ connections [59] allows to reduce the phase space of GR to that of a Yang–Mills gauge theory, which can be non-perturbatively quantized by formulating the theory using holonomies and fluxes as fundamental variables, à la Wilson.[16] But, let me follow the natural order of things, clarifying one thing at the time.

It is possible to demonstrate, in fact, that by introducing the Ashtekar self-dual $SL(2,\mathbb{C})$ connections in the framework of canonical GR, a Gauss constraint, which incorporates the generators of the local Lorentz boosts and rotations in a complex combination, appears besides the vectorial and scalar constraints, both connected with the diffeomorphisms gauge invariance of the theory. Simultaneously, the high non-linearity of the Arnowitt–Deser–Misner (ADM) canonical formulation of GR disappears: the new canonical constraints depend polynomially on the fundamental variables, both in vacuum and in the presence of matter [60].[17] By using the Ashtekar formulation

[16] The loop formalism does not work properly in Yang–Mills gauge theories, remarkably, the reason of this failure can be traced back to the basic assumption of a continuum space-time. This fact suggests that it may work well in QG, where a discrete space-time naturally emerges.

[17] It is worth noting that the standard ADM formulation of GR requires that the metric field is non-degenerate, since it contains the three-dimensional Ricci scalar, which is constructed by the Ricci tensor saturating the indexes with the inverse metric field. In the Ashtekar formulation the same requirement is not mandatory, since the constraints

of GR, a background independent quantum theory of gravity was later formulated [61]. But the use of complex fundamental variables generates a serious difficulty connected with the implementation of the *reality conditions* in the quantum theory, which are strictly necessary to ensure that the evolution of the system is real. This difficulty has not been overcome so far and, basically, it can be considered the technical motivation which led to the adoption of the real Ashtekar-Barbero (AB) connections [62, 63] as fundamental variables, instead of the complex ones. The link existing between real and complex variables can be clarified by observing that both are obtainable from the ADM canonical pair via a contact transformation. In particular, a suitable canonical transformation allows to introduce a finite complex number, $\beta \neq 0$, namely the BI parameter, in the definition of the new variables, so that they correspond to the (anti)self-dual ones when $\beta = \pm i$ and to the real ones for any real value of β.

Geometrically, the main difference between these two sets of possible new variables for GR is the following: while the complex connections are the projection over the 3-space of the self-dual part of the Ricci spin connections, the real ones are non-trivially related to them, complicating their reconstruction [64]. In fact, the real $SU(2)$ valued connections contain only half of the necessary information for reconstructing the Lorentz valued connections of GR [3], motivating also the necessity of fixing the temporal gauge in order to avoid the appearance of second class constraints.[18] By fixing the temporal gauge, the accessible part of the phase space is determined by first class constraints only [65] and the system can be quantized through the Dirac procedure. The result is a non-perturbative background independent quantum theory of gravity called *Loop Quantum Gravity* (LQG) [1, 2, 3, 66].[19]

Since the BI parameter has been introduced via a canonical transformation, one can naively believe that different values of β correspond to unitary equivalent quantum theories. Strangely enough, this is not the case. In fact, β enters in the spectrum of the main geometrical observables of the theory, e.g. the spectra of the area and volume operators, revealing that a one parameter family of non-equivalent quantum theories exists. As argued by Rovelli and Thiemann [75], two dynamically equivalent $SO(3)$-valued connections exist and, as a consequence, an ambiguity appears in the theory, which is essentially expressed by the presence of the BI parameter.

Immirzi suggested that the appearance of the BI parameter in the quantum theory was a consequence of the temporal gauge fixing [76], so that it would have disappeared in a fully Lorentz covariant theory. But this expectation was not completely confirmed by the so-called *Covariant Loop Quantum Gravity* (CLQG), which is a fully Lorentz covariant quantum theory of gravity,

are polynomial. So, we can say that the Ashtekar self-dual formulation of Gravity represents a possible extension of GR allowing the presence of degenerate metrics. Whether or not this extension has any physical relevance, up to my knowledge, is not completely understood yet.

[18] The temporal gauge fixing consists in rotating the local basis by using a suitable Wigner boost so that, at every instant of "time", its zeroth component is parallel to the normal vector to the instantaneous Cauchy hypersurface Σ^3_t. This condition reduces the local $SO(3,1)$ gauge group to the subgroup of spatial rotations, $SO(3)$, by fixing the boost component of the Lorentz symmetry.

[19] LQG besides providing interesting physical predictions as the quantization of areas and volumes [2] (see also [67, 68]), has been able to cure the inevitable singular behavior of classical GR in symmetric spacetimes [69, 70, 71]. Furthermore, the recently obtained results about the graviton propagator have strengthened the physical content of the theory, providing new insights into its non-singular behavior [72, 73, 74].

constructed *à la* Dirac relaxing the time gauge condition [77].[20] This approach, in fact, revealed a correlation between the choice of the fundamental variables and the appearance of the BI ambiguity in the quantum theory. In other words, in CLQG different choices of the fundamental variables are possible. In particular, for a geometrically well motivated specific choice of variables the resulting area spectrum no longer depends on the BI parameter [78]. But, choosing different fundamental variables considered as a direct generalization of the AB connections, the resulting area spectrum turns out to depend on the BI ambiguity [79], reproducing the result of the gauge fixed theory (see also the interesting paper [80]).

Recently, it has been proposed the idea that the BI parameter is, in fact, analogous to the θ-angle of the topological sector of Yang–Mills gauge theories [81, 82] (for a brief description of the topological sector of Yang–Mills gauge theories see Appendix B). This idea, initially proposed by Gambini, Obregon and Pullin [83], has been lately reconsidered in relation to the proposal to generalize the action for gravity to contain a topological term [82, 84, 85]. This argument will be better described below in 5.3, but it is worth anticipating that the presence of a topological term, called Nieh–Yan density [86], which further generalize the so-called Holst modification [87], allows, in fact, to construct a precise analogy between the BI parameter and the θ-angle. Furthermore, by clarifying the large structure of the gauge group involved in gravity through the Nieh-Yan density, it is possible to demonstrate its supposed topological origin and, as a consequence, the existence of non-unitary equivalent quantum theories associated to different values of β.

Having briefly outlined the AB formulation of GR and some recent aspects concerning the interpretation of the BI parameter, let me now enter in more technical details, starting from the tetrad formulation of gravity and the consequent generalization of the 3+1 splitting procedure.

5.1 3+1 splitting again

Let us introduce a one to one map $e : M^4 \to TM^4_x$, which sends tensor fields on M^4 in tensor fields in the Minkowskian tangent space TM^4_x. The fields $e_\mu{}^a$ are commonly called tetrads or vierbein (or, more physically, gravitational field! [2]) and represent a local reference system for space-time. They satisfy the following relations with the metric field:

$$g_{\mu\nu} = \eta_{ab} e_\mu{}^a e_\nu{}^b \,, \qquad e_\mu{}^a e^\mu{}_b = \delta^a_b \,, \qquad e_\mu{}^a e^\nu{}_a = \delta^\nu_\mu \,, \tag{5.1}$$

where Greek and Latin indexes run from 0 to 3, and transform respectively under general coordinates transformations and local Lorentz transformations. The symbol η_{ab} denotes the metric tensor in the local Minkowski frame. So, the tetrad fields incorporate all the metric properties of M. It is worth noting that the converse is not true. In fact, there are infinitely many choices of the local basis which reproduce the same metric tensor: This is clearly a consequence of the local Lorentz gauge invariance, manifestly present in this formalism. This is also the reason why there are more components in $e_\mu{}^a$ than in the metric $g_{\mu\nu}$, the difference being exactly six, that is the number of degrees of freedom of the group $SO(3,1)$ representing the number of independent parameters of a Lorentz rotation in the tangent space-time.

[20] It is worth remarking that the complicated form of the Dirac brackets, necessary to solve the second class constraints (see [49, 108]), prevents the fully Lorentz covariant theory from being rigorously formalized.

As we briefly described in § 3.2, the presence of a local gauge freedom requires the introduction of a covariant derivative D_μ transforming in the adjoint representation of the gauge group. This implies the introduction of a Lorentz valued connection (often referred as spin connection), here denoted as $\omega^{ab}(x)$ and satisfying the following property $\omega^{ab}(x) = -\omega^{ba}(x)$. The covariant derivative operator D_μ acts on Lorentz valued tensor fields as follows:

$$\begin{aligned} D_\mu T_{\nu_1\dots\nu_n}{}^{a_1\dots a_m} &= \partial_\mu T_{\nu_1\dots\nu_n}{}^{a_1\dots a_m} - \sum_{k=1}^{n} \Gamma^\rho_{\mu\nu_k} T_{\nu_1\dots\nu_{k-1}\rho\nu_{k+1}\dots\nu_n}{}^{a_1\dots a_m} + \sum_{l=1}^{m} \omega^{a_l}{}_{b\mu} T_{\nu_1\dots\nu_n}{}^{a_1\dots a_{l-1}ba_{l+1}\dots a_m} \\ &= \nabla_\mu T_{\nu_1\dots\nu_n}{}^{a_1\dots a_m} + \sum_{l=1}^{m} \omega^{a_l}{}_{b\mu} T_{\nu_1\dots\nu_n}{}^{a_1\dots a_{l-1}ba_{l+1}\dots a_m}, \end{aligned} \tag{5.2}$$

where $\Gamma^\rho_{\mu\nu}$ denotes the affine connection. Now, requiring the compatibility of the above defined covariant derivative operator with the tetrad basis, we can extrapolate the expression of the spin connection as function of the local basis vectors:

$$D_\mu e_\nu{}^a = 0 \quad \Longrightarrow \quad \omega^a{}_{b\mu} = e_\rho{}^a \nabla_\mu e^\rho{}_b, \tag{5.3}$$

where ∇_μ as usual satisfies the metric compatibility condition, i.e. $\nabla_\mu g_{\rho\sigma} = 0$.

As we are going to show, the same conclusion can be derived from the solution of the second Cartan structure equation, which will be extensively used below. In this respect, from the expression given above for the four dimensional spin connection as function of the tetrad fields, it is easy to derive the following equation:

$$\nabla_{[\mu} e_{\rho]}{}^a = -\omega_{[\mu}{}^{ab} e_{\rho]b}\,. \tag{5.4}$$

If the torsion-less condition holds, namely

$$\nabla_\mu \nabla_\nu f = \nabla_\nu \nabla_\mu f\,, \tag{5.5}$$

and remembering that the ∇ operator is compatible with the metric, we can easily deduce

$$\partial_{[\mu} e_{\rho]}{}^a = -\omega_{[\mu}{}^{ab} e_{\rho]b}\,, \tag{5.6}$$

which recalling the definition of exterior derivative of n-forms can be rewritten as:

$$de^a + \omega^a{}_b \wedge e^b = 0\,. \tag{5.7}$$

The above equation is called homogeneous second Cartan structure equation and completely determines the spin connection as function of the gravitational field.[21] It is worth noting that in the case of non-vanishing torsion, the affine connection $\Gamma^\rho_{\mu\nu}$ has also an antisymmetric part, then the second Cartan structure equation generalizes to:

$$de^a + \omega^a{}_b \wedge e^b = T^a\,, \tag{5.8}$$

where the torsion 2-form T^a is defined as

$$T^a_{\mu\nu} = e_\rho{}^a \left(\Gamma^\rho_{\mu\nu} - \Gamma^\rho_{\nu\mu}\right). \tag{5.9}$$

[21] For the reader's convenience, we collected a description of the forms formalism in Appendix A, where he/she can also find the Hilbert-Palatini and matter actions translated in the forms language.

In this more general case, the solution of the second Cartan structure equation provides the full spin connection as a sum of its torsion-less component $\omega^{ab}[e]$ plus a contortion term K^{ab}, namely

$$\Omega^{ab}[e,\dots] = \omega^{ab}[e] + K^{ab}[e,\dots]\,, \tag{5.10}$$

where the dots indicate that the contortion component can depend on matter fields as in the case of spinors coupled to gravity.

The spin connection 1-form generates, in the usual way, the curvature 2-form $R^a{}_b$ through the following identification (in what follows let us use the compact notation $d^{(\omega)}(\dots) \equiv d(\dots) + \omega \wedge (\dots)$):

$$d^{(\omega)} \circ d^{(\omega)} v^a = R^a{}_b \wedge v^b\,. \tag{5.11}$$

Explicitly we have

$$R^a{}_b = d\omega^a{}_b + \omega^a{}_c \wedge \omega^c{}_b\,, \tag{5.12}$$

which is called first Cartan structure equation. The curvature tensor satisfies the Bianchi identity

$$d^{(\omega)} R^a{}_b = 0\,, \tag{5.13}$$

which, as we noted before, is a consequence of the Jacobi identity applied to the covariant exterior derivatives operator, i.e. $d^{(\omega)} \circ d^{(\omega)} \circ d^{(\omega)} = 0$. Another identity can be obtained applying the exterior covariant derivative operator on the left and right hand sides of the second Cartan structure equation (5.8), i.e.

$$R^a{}_b \wedge e^b = dT^a + \omega^a{}_b \wedge T^b\,, \tag{5.14}$$

which in the torsion-less case reduces to the Bianchi cyclic identity:

$$R^a{}_b \wedge e^b = 0\,. \tag{5.15}$$

It is possible to rewrite the action for gravity in the tetrad formalism by using the relations given in (5.1), specifically we have

$$S[e] = \frac{1}{2}\int d^4x \det[e] e^\mu{}_a e^\nu{}_b R_{\mu\nu}{}^{ab}\,, \tag{5.16}$$

which can be used as starting point to canonically reformulate the gravitational theory. Actually, the canonical formulation of tetrad gravity can be straightforwardly deduced from the canonical theory described in Sections 4.1 and 4.2, by taking into account the presence of an additional local symmetry in this new framework.

So, in order to construct the canonical theory, let us assume that the space-time is a globally hyperbolic metric manifold ($M = \mathbb{R}\times\Sigma^3, g_{\mu\nu} = \eta_{ab} e^a_\mu e^b_\nu$). Let $h_{\mu\nu} = g_{\mu\nu} + n_\mu n_\nu$ and $K_{\mu\nu} = 1/2\pounds_n h_{\mu\nu}$ be respectively the first and second fundamental form of the Cauchy hypersurface Σ^3, so that, by using the tetrad fields, we can write:

$$h_{\mu\nu} = \eta_{ab} e_\mu{}^a e_\nu{}^b + n_\mu n_\nu = -e_\mu{}^0 e_\nu{}^0 + \delta_{ij} e_\mu{}^i e_\nu{}^j + n_\mu n_\nu\,. \tag{5.17}$$

The expression of the first fundamental form above suggests a very simple gauge choice to eliminate *ab initio* some non-dynamical degrees of freedom. Specifically, by operating a Wigner boost

it is possible to rotate the local basis in such a way that the component $e_\mu{}^0$ results to coincide with the normal vector n_μ. This gauge choice, called temporal gauge, reduces the local gauge symmetry to the group of spatial rotations, since the boost component of the original Lorentz group is obliged to vanish by having chosen the direction of one of the component of the local basis.[22]

By implementing the temporal gauge and projecting the 3-metric on the hypersurface, we have

$$h_{\alpha\beta} = \delta_{ij} e_\alpha{}^i e_\beta{}^j \,, \tag{5.18}$$

where Greek indexes $\alpha, \beta, \gamma \ldots$ run over $1,2,3$, while Latin indexes $i, j, k, \ldots$ are related to the $SO(3)$ (or via an isomorphism $SU(2)$) local symmetry. Note that the orthogonality condition $n^\mu h_{\mu\nu} = 0$ is automatically fulfilled according to the properties of the tetrad basis. This allows us to easily write the components of the tetrad fields and their inverse in the coordinates system (t,x), we have

$$e_\mu{}^a = \begin{pmatrix} N & N^\alpha e_\alpha{}^i \\ 0 & e_\alpha{}^i \end{pmatrix} \qquad \text{and} \qquad e^\nu{}_b = \begin{pmatrix} \dfrac{1}{N} & 0 \\ -\dfrac{N^\beta}{N} & e^\beta{}_j \end{pmatrix}, \tag{5.19}$$

as before N denotes the Lapse function, while N^α the Shift vector. This identification of some of the components of the local basis with the Lapse and Shift is possible by considering that the line element

$$ds^2 = e_\mu{}^a e_{a\nu} dx^\mu dx^\nu = N^2 dt^2 - h_{\alpha\beta} \left(N^\alpha dt + dx^\alpha\right)\left(N^\beta dt + dx^\beta\right), \tag{5.20}$$

corresponds to the ADM decomposition of the 4-metric. It is important to note that the 3-metric (5.18) is invariant under a local $SO(3)$ rotations of the tetrad basis, namely the dreibein carries three degrees of freedom more with respect to the three-metric $h_{\alpha\beta}$. As a consequence also the number of constraints of the canonical theory written in tetradic formalism must increase. In particular, we should have three (first class) constraints more in order to reabsorb the local gauge degrees of freedom connected with the $SO(3)$ symmetry.

In this respect, consider the second fundamental form. Define the 1-form $K_\alpha{}^i$ on Σ^3 and contract the internal index with the tetrads, i.e.

$$\mathscr{K}_{\alpha\beta} = \delta_{ik} K_\alpha{}^i e_\beta^k \,. \tag{5.21}$$

Now, it is easy to realize that the symmetric part of the tensor $\mathscr{K}_{\alpha\beta}$ is the extrinsic curvature, namely

$$\mathscr{K}_{(\alpha\beta)} = K_{\alpha\beta} \,, \tag{5.22}$$

while its antisymmetric part has to vanish corresponding to the antisymmetric part of the extrinsic curvature, which is naturally symmetric. Then the 1-form $K_\alpha{}^i$ must satisfies the following con-

[22] It is worth stressing that the fixation of the temporal gauge remarkably simplifies the canonical theory. Nevertheless, we have to say that, the temporal gauge is not a mandatory choice to construct the canonical theory. In fact, we can canonically formulate the theory as well without fixing the gauge, but the technical difficulties one would have to face in solving the second class constraints are far from the scope of this paper.

straint:[23]

$$\mathscr{K}_{[\alpha\beta]} = K_{[\alpha}{}^{i} e_{\beta]i} \approx 0\,. \tag{5.23}$$

$\mathscr{K}_{[\alpha\beta]}$ is a 3×3 antisymmetric matrix constrained to vanish, so it reabsorbs exactly the three degrees of freedom that we have introduced with the choice of the tetrads as elementary variables.[24]

As a next step, let me introduce the weighted triad fields

$$E^{\alpha}_{i} = e\, e^{\alpha}{}_{i}\,, \tag{5.24}$$

where $e = \det[e^{i}_{\alpha}]$.[25] By using them, we can rewrite the constraint (5.23) as

$$G_{ij} = K_{\alpha[i} E^{\alpha}{}_{j]} \approx 0\,. \tag{5.25}$$

At this point we are ready for changing variables, specifically we can easily rewrite the canonical ADM variables as:

$$h_{\alpha\beta} = \delta_{ij} E_{\alpha}{}^{i} E_{\beta}{}^{j}\,, \tag{5.26a}$$

$$p^{\alpha\beta} = \delta^{(\alpha}_{\delta} E^{\beta)i} E_{\gamma}{}^{k} E^{[\gamma}{}_{i} K^{\delta]}_{k}\,. \tag{5.26b}$$

The new couple of fundamental variables can be introduced in the scalar and vectorial constraints too, we obtain

$$H_{\alpha} = -D_{\beta}\left[K_{\alpha}{}^{i} E^{\beta}_{i} - \delta^{\beta}_{\alpha} K_{\gamma}{}^{i} E^{\gamma}_{i}\right]\,, \tag{5.27a}$$

$$H = \frac{1}{(\det[E^{i}_{\alpha}])^{1/2}} E^{\alpha}_{i} E^{\beta}_{j}\left(K_{\alpha}{}^{j} K_{\beta}{}^{i} - K_{\alpha}{}^{i} K_{\beta}{}^{j}\right) - \left(\det[E^{i}_{\alpha}]\right)^{1/2}\, {}^{(3)}R(E)\,, \tag{5.27b}$$

where the Ricci scalar curvature ${}^{(3)}R(E)$ is considered as function of the weighted tetrads. At this point, it is important to demonstrate that the canonical dynamics, described by the new variables on the extended phase space,[26] is equivalent to that described by the usual ADM variables. In this respect, we can easily demonstrate that once the extended phase space is equipped with the natural symplectic structure

$$\left\{E_{\gamma}{}^{i}(t,x), K^{\delta}{}_{j}(t,y)\right\} = \delta^{i}_{j}\delta^{\delta}_{\gamma}\,\delta(x-y)\,, \tag{5.28a}$$

$$\left\{E_{\gamma}{}^{i}(t,x), E_{\delta}{}^{j}(t,y)\right\} = \left\{K^{\gamma}{}_{i}(t,x), K^{\delta}{}_{j}(t,y)\right\} = 0\,, \tag{5.28b}$$

[23] It is important to note that here we are working in the second order formalism, namely torsion vanishes *ab initio*. In first order formalism, the same result can be obtained for pure gravity, but the situation changes when, e.g., spinor matter fields are considered in the dynamics. Spinors, in fact, are a matter source for torsion and, in general, their presence prevents the antisymmetric part of the tensor $\mathscr{K}_{\alpha\beta}$ from vanishing, so that a source appears in the rotational constraint.

[24] The reader may wonder about the legitimacy of the weakly vanishing symbol in (5.23). In this respect, we have to say that K^{i}_{α} has been generically defined, but only when it satisfies the condition expressed by (5.23), it can be safely related to the extrinsic curvature. Since the final goal will be to change variables and describe the canonical dynamics through the canonical couple $(K^{i}_{\alpha}, E^{\beta}_{k})$, the second being defined below, such a canonical system describes ordinary gravity when (5.23) is satisfied. In other words, condition (5.23) plays exactly the role of a constraint (weak equation) limiting the physically relevant evolution to a restricted region of the enlarged phase space.

[25] Notice that the inverse of the weighted triad is divided by the determinant of e^{i}_{α}.

[26] The adjective "extended" refers to the fact that the phase space associated to the new couple of variables has six dimensions more with respect to the ADM one.

the constraint

$$G(\alpha) = \int_\sigma d^3x \alpha^{ik} K_{\alpha i} E^\alpha_{\ j}, \tag{5.29}$$

where α^{ik} is an arbitrary antisymmetric matrix function, generates $SO(3)$ rotations. Moreover, as soon as $G(\alpha) = 0$, namely the rotational constraint is satisfied, the ADM canonical variables, written as functions of the extended phase space variables, generate the usual Poisson brackets (4.23c). In fact, calculating the Poisson algebra generated by the rotational constraint, we obtain:

$$\{G(\alpha), G(\alpha')\} = G\left([\alpha, \alpha']\right), \tag{5.30}$$

which is exactly the algebra of spatial $SO(3)$ rotations. Furthermore, any Poisson bracket between the rotational constraint and the ADM canonical variables vanishes, simply because the latter are manifestly rotations invariant. The 3-metric tensor $h_{\alpha\beta}$, being a function only of the weighted tetrad fields (see relation (5.26a)), simply satisfies the following relation:

$$\{h_{\alpha\beta}(t,x), h_{\gamma\delta}(t,y)\} = 0. \tag{5.31}$$

The case of the canonical momentum is more complicated, but finally we have [3]:

$$\left\{p^{\alpha\beta}(t,x), p^{\gamma\delta}(t,y)\right\} = -\frac{\sqrt{h}}{8}\left[h^{\alpha\gamma}G^{\beta\delta} + h^{\alpha\delta}G^{\beta\gamma} + h^{\beta\gamma}G^{\alpha\delta} + h^{\beta\delta}G^{\alpha\gamma}\right](t,x)\ \delta(x-y). \tag{5.32}$$

Namely, the above brackets vanish as soon as the rotation constraint is satisfied, as anticipated above. At last, with few algebraic passages we also obtain that

$$\left\{p^{\alpha\beta}(t,x), h_{\gamma\delta}(t,y)\right\} = \delta^\alpha_{(\gamma}\delta^\beta_{\delta)}\,\delta(x-y). \tag{5.33}$$

Summarizing, the new extended phase space elementary variables $E_\alpha^{\ i}$ and K_i^α reduces through the definitions in lines (5.26a) and (5.26b) to the ADM ones, moreover their Poisson brackets mimic the ADM ones as soon as the rotational constraint is satisfied. So we can conclude that the Hamiltonian system described by the action

$$S_{3+1}(E,K) = \frac{1}{2}\int_{\mathbb{R}\times\sigma} dt d^3x \left(K_i^\alpha \dot{E}_\alpha^{\ i} - NH - N^\alpha H_\alpha + O^{ik}G_{ik}\right),$$

once solved the rotational constraint, $G_{ik} = 0$, is equivalent to that described by the ADM action (4.33).

The constraints (5.23), (5.27a) and (5.27b) are first class, as can be demonstrated with some algebra. They reflect the gauge structure of the theory, indeed they are correlated to the automorphisms of the tangent bundle, namely the $SO(3)$ rotational symmetry, and to the space-time diffeomorphisms.[27]

[27] It should be clear from what stated above that this statement does not mean that the flow generated by the "new" super-Hamiltonian and super-momentum constraints on the extended phase space is a diffeomorphisms. Rather, once the rotational constraint is satisfied, i.e. on the constraint surface determined by the condition $G_{ij} = 0$, it still exists a representation of the diffeomorphisms group. There, in fact, the Hamiltonian flow of the rotational invariant variables $h_{\alpha\beta}$ and $p^{\alpha\beta}$ is equivalent to that generated on the ADM phase space.

5.2 Canonical transformations and new variables for gravity

Let us begin noting that, given the symplectic structure (5.28), the transformations $E^{\alpha}_{\ j} \rightarrow E^{\alpha}_{\ j}/\beta$ and $K_{\alpha}^{\ j} \rightarrow \beta K_{\alpha}^{\ j}$ are canonical, in fact they do not change the symplectic structure. The parameter β is called Barbero–Immirzi (BI) parameter and is, in general, a complex number. It is worth noting that the rotational constraint remains invariant under this rescaling, so we can rewrite it as

$$G_k = \varepsilon_{kij} G^{ij} = \varepsilon_{kij}\, {}^{(\beta)}K_{\alpha}^{\ i}\, {}^{(\beta)}E^{\alpha j} \approx 0, \tag{5.34}$$

where we used the properties of the total antisymmetric symbol ε_{kij} and we indicated as ${}^{(\beta)}K_{\alpha}^{\ i}$ and ${}^{(\beta)}E^{\alpha}_{\ i}$ the canonical variables rescaled by the BI parameter.

Let us now going on discussing the canonical formalism of tetrad gravity by introducing the connection associated with the $SO(3)$ symmetry (or equivalently $SU(2)$). So, we introduce the connection 1-form $\Gamma^{ij}_{\ \ \alpha}$, in order to covariantly derive $SO(3)$ valued tensors. In particular, let us define the connection via its action on the generic tensor $T^{i_1\ldots i_m}_{\alpha_1\ldots\alpha_n}$ containing both vectorial and $SO(3)$ indexes, i.e.

$$D_\beta T_{\alpha_1\ldots\alpha_n}^{\ \ \ \ \ i_1\ldots i_m} = \partial_\beta T_{\alpha_1\ldots\alpha_n}^{\ \ \ \ \ i_1\ldots i_m} - \sum_{k=1}^{n} \Gamma^{\gamma}_{\beta\alpha_k} T_{\alpha_1\ldots\alpha_{k-1}\gamma\alpha_{k+1}\ldots\alpha_n}^{\ \ \ \ \ \ \ \ \ \ \ i_1\ldots i_m} + \sum_{l=1}^{m} \Gamma^{i_l}_{\ \ j\beta} T_{\alpha_1\ldots\alpha_n}^{\ \ \ \ \ i_1\ldots i_{l-1} j i_{l+1}\ldots i_m}. \tag{5.35}$$

It can be verified by direct analysis of the formula above that the generalized covariant derivative D_β on Σ sends each smooth $SO(3)$ valued tensor field of type (p,q) to a smooth $SO(3)$ valued tensor of type $(p,q+1)$. As usual, we require that the covariant derivative operator is compatible with the tetrad basis, i.e. it annihilates the field $e_{\alpha}^{\ i}$, namely we have

$$D_\beta e_{\alpha}^{\ i} = 0 \quad \Longrightarrow \quad \Gamma^{ij}_{\ \ \alpha} = e^{\beta i} \nabla_\alpha e_{\beta}^{\ j}. \tag{5.36}$$

As usual we can associate to the connection the curvature 2-form $R_{\alpha\beta}^{\ \ ij}$ defined via the action on a $SO(3)$ scalar

$$R_{\alpha\beta}^{\ \ i}{}_{j} v^j = \left[D_\alpha, D_\beta\right] v^i \longrightarrow R_{\alpha\beta}^{\ \ ik} = 2\partial_{[\alpha}\Gamma^{ik}_{\ \ \beta]} + \Gamma^{i}_{\ j[\alpha}\Gamma^{jk}_{\ \ \beta]}. \tag{5.37}$$

Let us now note that the following relation holds:

$$D_\alpha E^{\beta}_{\ j} = \nabla_\alpha \left(e\, e^{\beta}_{\ j}\right) - \Gamma^{k}_{j\alpha} e\, e^{\beta}_{\ k} = e D_\alpha e^{\beta}_{\ j} = 0, \tag{5.38}$$

where we taken into account the compatibility condition (5.36). Consequently we have,

$$D_\alpha E^{\alpha}_{\ j} = 0, \tag{5.39}$$

moreover, being $\partial_\alpha E^{\alpha}_{\ j} = \partial_\alpha \left(e\, e^{\alpha}_{\ j}\right) = \nabla_\alpha E^{\alpha}_{\ j}$, we finally obtain the following important relation:

$$D_\alpha E^{\alpha}_{j} = \partial_\alpha E^{\alpha}_{j} - \Gamma^{k}_{\ j\alpha} E^{\alpha}_{k} = \partial_\alpha E^{\alpha}_{j} + \varepsilon_{jk}^{\ \ l} \Gamma^{k}_{\alpha} E^{\alpha}_{\ j} = 0, \tag{5.40}$$

where we defined $\Gamma^{k}_{\alpha} \overset{def}{=} -1/2\, \varepsilon^{k}_{\ ij} \Gamma^{ij}_{\ \ \alpha}$. It is worth noting that the above defined Γ^{i}_{α} as function of the dreibeins fields can be recasted as function of the weighted triads $E^{\alpha}_{\ j}$, we give below the

expression:[28]

$$\Gamma^i_\alpha = \frac{1}{2}\varepsilon^i{}_{jk}E^{\beta k}\left[\partial_\beta E_\alpha{}^j - \partial_\alpha E_\beta{}^j + E^{\gamma j}E_{\alpha l}\partial_\beta E_\gamma{}^l\right] + \frac{1}{4}\varepsilon^i{}_{jk}E^{\beta k}\left[2E_\alpha{}^j\frac{\partial_\beta\left(\det[E_\gamma{}^i]\right)}{\det[E_\gamma{}^i]} - E_\beta{}^j\frac{\partial_\alpha\left(\det[E_\gamma{}^i]\right)}{\det[E_\gamma{}^i]}\right], \tag{5.41}$$

which can be calculated simply by substituting the definition (5.24) in the expression (5.36) for the spatial spin connection. We also note that Γ^k_α is not affected by the rescaling $E^\alpha_{\ j} \to {}^{(\beta)}E^\alpha_{\ j} = E^\alpha_{\ j}/\beta$ of the weighted tetrad fields. By using the strong equation (5.40), we can now replace the rotational constraint (5.25) with a Gauss constraint of an $SO(3)$ or $SU(2)$ Yang-Mills gauge theory, as follows

$$\begin{aligned} G_k &= \varepsilon_{ki}{}^{j\,(\beta)}K_\alpha{}^{i\,(\beta)}E^\alpha_{\ j} = \partial_\alpha{}^{(\beta)}E^\alpha_{\ k} + \varepsilon_{ki}{}^j\Gamma^i_\alpha{}^{(\beta)}E^\alpha_{\ j} + \varepsilon_{ki}{}^{j\,(\beta)}K_\alpha{}^{i\,(\beta)}E^\alpha_{\ j} \\ &= \partial_\alpha{}^{(\beta)}E^\alpha_{\ k} + \varepsilon_{ki}{}^{j\,(\beta)}\mathscr{A}^i_\alpha{}^{(\beta)}E^\alpha_{\ j} \approx 0\,, \end{aligned} \tag{5.42}$$

where the Ashtekar-Barbero connection, defined as ${}^{(\beta)}\mathscr{A}^i_\alpha = \Gamma^i_\alpha + \beta K_\alpha{}^i$ has been introduced. For $\beta = \pm i$ we obtain the original definition of the Ashtekar self-dual variables for gravity.

Again a change of variables is in order. Specifically, we can replace the canonical couple (K^i_α, E^β_k) with the new couple $({}^{(\beta)}\mathscr{A}^i_\alpha, {}^{(\beta)}E^\beta_k)$ and, once checked that the replacement does not affect the symplectic structure (canonical transformation),[29] i.e.

$$\left\{{}^{(\beta)}E_\gamma{}^i(t,x),\, {}^{(\beta)}A^\delta{}_j(t,y)\right\} = \delta^i_j\delta^\delta_\gamma\,\delta(x,y)\,, \tag{5.43a}$$

$$\left\{{}^{(\beta)}E_\gamma{}^i(t,x),\, {}^{(\beta)}E_\delta{}^j(t,y)\right\} = \left\{{}^{(\beta)}A^\gamma{}_i(t,x),\, {}^{(\beta)}A^\delta{}_j(t,y)\right\} = 0\,, \tag{5.43b}$$

we can go on to rewrite the canonical constraints as functions of the new variables.

So, let us define the curvature of the new connections

$$\mathscr{F}^k_{\alpha\beta} = 2\partial_{[\alpha}{}^{(\beta)}\mathscr{A}^k_{\beta]} + \varepsilon^k_{\ ij}{}^{(\beta)}\mathscr{A}^i_\alpha{}^{(\beta)}\mathscr{A}^j_\beta\,, \tag{5.44}$$

satisfying the following relation

$$\mathscr{F}^k_{\alpha\beta} = \varepsilon^k_{\ ij}R_{\alpha\beta}{}^{ij} + 2D_{[\alpha}{}^{(\beta)}K_{\beta]}{}^k + \varepsilon^k_{\ ij}{}^{(\beta)}K_\alpha{}^{i\,(\beta)}K_\beta{}^j\,. \tag{5.45}$$

As a useful formula, we rewrite the 3-dimensional Bianchi cyclic identity, $R_{[\alpha\beta}{}^{ij}e_{\gamma]i} = 0$, as

$$\varepsilon^i{}_{jk}R_{[\alpha\beta}{}^{jk}E_{\gamma]i} = 0\,. \tag{5.46}$$

By using the definition (5.44) and the equations (5.45) and (5.46), the canonical constraints can be rewrite as follows:

$$H_\alpha = {}^{(\beta)}E^\gamma_i\mathscr{F}^i_{\alpha\gamma} - {}^{(\beta)}K^i_\alpha G_i\,, \tag{5.47a}$$

$$H = \frac{\beta^2}{2\left(\det[E_\alpha{}^i]\right)^{1/2}}{}^{(\beta)}E^\alpha_i\,{}^{(\beta)}E^\gamma_j\left[\varepsilon^{ij}{}_k\mathscr{F}^k_{\alpha\gamma} - 2\left(\beta^2+1\right)K^i_{[\alpha}K^j_{\gamma]}\right] + {}^{(\beta)}E^\gamma_j D_\gamma G^j\,, \tag{5.47b}$$

[28] The expression of the 3-dimensional spin connection as function of the densitized triad is the one that we should use in (5.27b) to rewrite the Ricci scalar as function of the E^α_i. This can be easily done considering that $R(E) = \frac{E^\alpha_i E^\beta_k}{\det[E]}R_{\alpha\beta}{}^{ik}[\Gamma(E)]$, where $R_{\alpha\beta}{}^{ik}[\Gamma(E)]$ is the curvature tensor associated with the connection Γ and defined in (5.37).

[29] The only non-trivial check is that the Poisson bracket $\left\{{}^{(\beta)}A^\gamma{}_i(t,x),\, {}^{(\beta)}A^\delta{}_j(t,y)\right\}$ in fact vanishes. This is a non-trivial result, mainly based on the fact that $\left\{\Gamma^i_\alpha(E), K^j_\beta\right\} = 0$.

Where we have taken into account of the formula given in note 24. It is interesting to note that both the scalar and vectorial constraints (5.47) involve the Gauss constraint. Obviously, since the new expressions of the canonical constraints (5.42) and (5.47) are the consequence of a well defined canonical transformation, the structure of their Poisson brackets is unaffected. In other words, they still represent a set of first class constraints, correlated with the gauge symmetry of the system and oblige the system to evolve in a restricted region of the extended phase space. The same dynamics can be obtained directly working with the following set of first class constraints:

$$G_i = \mathscr{D}_\alpha E^\alpha_i = \partial_\alpha E^\alpha_i + \varepsilon_{ij}{}^{k\ (\beta)}\mathscr{A}^j_\alpha E^\alpha_k \approx 0\,, \tag{5.48a}$$

$$C_\alpha = E^\beta_i\, \mathscr{F}^i_{\alpha\beta} \approx 0\,, \tag{5.48b}$$

$$C = \frac{1}{2\left(\det[E^i_\alpha]\right)^{1/2}} E^\alpha_i E^\beta_j \left[\varepsilon^{ij}{}_k \mathscr{F}^k_{\alpha\beta} - 2\left(\beta^2+1\right) K^i_{[\alpha} K^j_{\beta]}\right] \approx 0\,, \tag{5.48c}$$

which are dynamically equivalent to the previous ones.

It is worth noting that the Gauss law (5.48a) and the vector constraint (5.48b) do not depend on the BI parameter β, while the scalar constraint (5.48c) is β-dependent, implying that the physical predictions of the quantum theory will in general depend on the Immirzi parameter. A weird fact is that even physical quantities not directly depending on the Hamiltonian, for example the area operator come out to be β-dependent.

As a final remark, the values $\beta = \pm i$ corresponding to the complex Ashtekar (anti)self-dual variables are pretty special: The constraints become polynomial, provided that we can reabsorb the determinant of the densitized triad in the denominator of the scalar constraint, e.g. reabsorbing it in the Lapse function [60].

5.3 Holst action as Lagrangian formulation of Ashtekar gravity

So far, we have never introduced the action which correspond to the AB constraints calculated previously. We dedicate this last part of the section to this argument, describing also a recent proposal for a possible generalization of the so-called Holst action.

The Holst action represents an important contribute in understanding the geometrical content of the Ashtekar-Barbero formalism. In [87], Holst showed that the AB canonical constraints of GR [62, 63] can be derived by splitting a generalized Hilbert-Palatini action. The Hilbert-Palatini action in tetrad formalism is:[30]

$$S[e,\omega] = \frac{1}{2}\int e_a \wedge e_b \wedge \star R^{ab}(\omega)\,. \tag{5.49}$$

The Riemann tensor is a function of the spin-connection, ω, which is considered as a separate distinct variables with respect to the gravitational 1-form $e^a = e^a_\mu dx^\mu$. In other words, the action as to be varied with respect to both the gravitational field and the spin connections to write down the full set of equation of motion. Specifically, by varying the action with respect to ω and e respectively, we obtain

$$de^a + \omega^a_b \wedge e^b = 0\,, \tag{5.50a}$$

$$\varepsilon^a_{bcd} e^b \wedge R^{cd}(\omega) = 0\,. \tag{5.50b}$$

[30] Let us use differential forms in order to be more concise in formulating this argument. Some details are given in Appendix A.

The first one of the above equations is the second Cartan structure equation (5.8), containing the information about the relation between the spin connection and the gravitational field. This can be easily solved and the solution can be put in (5.50b), which are the Einstein equations in vacuum.

It is worth noting that the presence of matter, in general, affects the Einstein equations by generating a source in the right hand side of (5.50b). Remarkably spinor fields, which interact with gravity through both the gravitational field and the spin connection, generate also a source for the second Cartan structure equation, namely they represent a source for torsion [84].

We claim that the Holst action

$$S(e,\omega) = S_{\text{HP}}(e,\omega) + S_{\text{Hol}}(e,\omega) = \frac{1}{2}\int e_a \wedge e_b \wedge \left(\star R^{ab} - \frac{1}{\beta} R^{ab}\right), \tag{5.51}$$

where β is the BI parameter, is the starting point to formulate canonical gravity in the AB variables. The Holst action is made up of two parts, the first one is the usual Hilbert-Einstein action, while the second one is an "on (half-)shell" vanishing term. Classically, the Holst action is dynamically equivalent to the HP one, indeed, by varying it with respect to the spin connection we get the unmodified second Cartan structure equation, while the variation with respect to the gravitational field gives

$$\varepsilon^{a}{}_{bcd}\, e^b \wedge R^{cd}(\omega) - \frac{1}{\beta}\, e_b \wedge R^{ab}(\omega) = 0\,. \tag{5.52}$$

But as previously demonstrated, the homogeneous second Cartan structure equation implies the cyclic Bianchi identity (5.15), which ensures that the Einstein dynamics is preserved by the Holst modification.

We note that in the Holst formulation the BI parameter turns out to be a multiplicative constant in front of a on (half-)shell vanishing term, this clarifies why it does not affect the classical dynamics, while it has important effects on the quantum dynamics as remarked previously. This behavior is reminiscent of the parameter characterizing the topological sector of Yang–Mills gauge theories (see action (3.27) and the comment below; see also Appendix B). If the θ-angle and the BI parameter β have an analogous origin, then it must exist a classical framework where the analogy between the two parameters can be made manifest. In the pure gravitational case, in fact, the argument proposed fails to be completely convincing. The Holst modification, in fact, is not a topological density. It does not reduce to a total divergence, rather it is an on-shell identically vanishing term. But the action (5.51) can be further generalized to include in the picture also the interesting case of torsional space-times. In particular, in [84, 85], by introducing spinor matter fields, an interesting hint was given to complete the Holst picture; specifically, the presence of spinors can generate the necessary torsion contribution to generalize the Holst modification and construct a topological term. In other words, by using a non-minimal coupling between spinors and gravity,[31] it has been, indirectly, demonstrated that the EC action can be generalized without modifying the classical dynamics by adding the Nieh–Yan topological density [86], i.e.

$$S_{\text{Grav}} = S_{\text{HP}}[e,\omega] + S_{\text{NY}}[e,\omega] = \frac{1}{2}\int e_a \wedge e_b \wedge \star R^{ab} + \frac{1}{2\beta}\int \left(T^a \wedge T_a - e_a \wedge e_b \wedge R^{ab}\right).$$

[31] See [88] for the extension to supergravity theories.

By remembering the definition of the torsion 2-form $T^a = de^a + \omega^a_{\ b} \wedge e^b$, the NY term can be easily rewritten as a total divergence, i.e.

$$\int \left(T^a \wedge T_a - e_a \wedge e_b \wedge R^{ab} \right) = \int d\left(e_a \wedge T^a \right) . \tag{5.53}$$

The modification is now a true topological term related to the Chern–Pontryagin classes [89], so that new interesting insights can be provided on the physical origin of the BI parameter. Moreover, this generalization is quite natural [90, 91] and motivated another approach similar to the discussion in the Appendix B, presented in [81], as well as other recent works [92, 93]. The structure of the large gauge group, which is supposed to be at the base of the proposed interpretation of the BI parameter is quite subtle; a possible framework is described in [82].

Finally, we want to briefly digress on an interesting possibility which has lately attracted much interest. The analogy existing between the θ-angle of Yang–Mills gauge theories and the BI parameter in gravity suggest a further generalization, namely the idea that the BI parameter is actually a field [94, 92, 95]. Initially, this idea was considered just as a possible generalization of the theory, but recently it has been demonstrated that promoting the BI parameter to be a field could be necessary in order to reabsorb a divergence coming from the chiral anomaly on space-time with torsion [91]. This proposal has an interesting outcome, indeed, once the BI field is coupled to gravity via the Nieh–Yan density, this ceases to be topological and the presence of the new field generates a torsion contribution in the second Cartan structure equation. In the case of pure gravity interacting with the BI field, the net effect of the presence of torsion is the appearance of a kinetic term for the BI field, which becomes a decoupled pseudo-scalar field. A more interesting dynamics appears as soon as we consider the presence of fermion fields. Indeed, the BI field turns out to interact with fermions and, through the chiral anomaly, with boson fields as well [91, 93, 96]. As it happens for the axion, also in this case instantonic effects can give an extremely small mass to the BI field. The mass of the BI field can be easily evaluated [91, 93], as well as some cosmological implications [97].

We conclude this section saying that the nature of the BI parameter is still unclear and is argument of active discussion from both the purely classical and quantum perspective.

6. Quantization Program

Previously in this paper, specifically in § 2.1, we discussed some general and very well known arguments which motivates the attempts to formulate a consistent QG. Remarkably, the necessity of a quantum theory of gravity was pointed out by Einstein himself in 1916: More than ninety years later a fully consistent and complete formulation of a quantum gravity theory still lacks.

On the one hand, it appeared immediately obvious from the pioneering works of Dirac, Wheeler, and DeWitt, that the problem of quantizing gravity was much more conceptually involved than other analogous problems regarding the other interactions. This led to the idea that the problem of QG could not be solved separately from the other interactions, namely, that it was inextricably bound to the issue of unification. So, for many years, the problem of quantizing pure gravity marginally interested physicists, more attracted by the attempt of unifying the other interactions or, more recently, by the idea of supersymmetry and extra-dimensions. Another interesting aspect that

is worth mentioning is the belief that the main question to answer to construct a consistent quantum gravity was the disappearance of time. In other words, initially it seemed that the so called *frozen formalism* was the main obstacle to obtain a physically consistent formulation of QG. But the peculiar role played by time in ordinary quantum mechanics is mainly correlated with the, let's say, evolutionary interpretation of physical theory we have been used to by classical mechanics. The presence of an evolutionary parameter is neither a fundamental request of the quantization procedure, nor a fundamental ingredient for the physical interpretation of the theory. A quantum theory without time can be, in fact, perfectly consistent.

On the other hand, the physical situation a theoretical physicist is called to face in constructing a quantum theory of gravity should be the best he/she can imagine. At present, in fact, there is no strong experimental constraint on the quantum gravity regimes. Naively, one could expect that a rich variety of different consistent theories has been formulated so far, on the contrary, we do not have anyone. Most likely, the reason is the double nature of GR, namely as a field theory describing gravity and as a theory of space-time. Any quantum gravity theory, in fact, has to put together three fundamental dynamical elements, i.e. geometry, gravity and quantum laws. In this perspective, the ordinary quantum theory of field cannot provide any insightful hint as one of the fundamental ingredients lacks, i.e. the dynamical nature of the space-time geometry.

We know, in fact, that as soon as we treat quantum gravity perturbatively, namely neglecting the full dynamics of space-time, as one would do following the prescriptions of quantum field theory, the result we get is a non-renormalizable theory. So, it seems pretty natural to incorporate the full dynamics of space-time in the theory through a non-perturbative approach. But one may wonder if a non-perturbative quantum theory of gravity can actually be a consistent theory. This question is often suggested by a naive analogy based on the behavior at high energy of the Fermi theory for the weak interaction. As is well known, the Fermi model contains a point-like four fermions interaction, which is non-renormalizable. Fermi's model works well at low energy, but it is doomed to fail at high energies. A striking progress was done in this sense by completely reformulating the theory through the introduction of the massive bosons $W^{\pm}$ and Z^0 carrying the weak interaction.

It is often argued that an analogous procedure has to be applied to GR, since its perturbative non-renormalizability points in a direction similar to that of the Fermi model of weak interaction. Nevertheless, this argumentation completely fails in getting an essential difference between the weak and the gravitational interaction, namely the fact that perturbative expansions presuppose that the space-time is a smooth continuum at all the energy scales. But, there is no reason to believe that the classical concept of continuum space-time has to survive at scales of the order of the Plank energy. That is why a non-perturbative approach, able to incorporate the complete dynamics of the geometry of space-time, may safely describe quantum general relativity.

Furthermore, the failure of the standard perturbation expansion in gravity may well reflect the fact that GR is characterized by a non-trivial fixed point of the renormalization group flow. This extremely fascinating aspect of perturbative QG has been well described in this school by Roberto Percacci, who, in his two lectures, has pointed out that there is a growing evidence that this is exactly the case. Furthermore the requirement that the fixed point should continue to exist also in presence of matter fields constrains the possible couplings in an interesting physical way [98].

In general, it is expectable that a consistent quantum theory of gravity is able to remove sin-

gularities, replacing them with a well defined quantum state of the gravitational field. Initially, this was just a hope, but now encouraging results exist. They are mainly due to the general quantization program of Loop Quantum Gravity, which faces the problem of quantum gravity merging together the three main ingredients said above. They are still partial results, in the sense that it is possible to remove singularities at least in symmetric space-times, both of cosmological origin, as the Big Bang singularity [70, 99, 100, 101, 102, 103, 104, 105] and those resulting from a complete gravitational collapse [71, 106, 107]. Of course, a theorem establishing a general result about the possibility to avoid singularities in LQG still lacks.

However, the existing results make us confident that a suitable background independent quantization of the gravitational field can solve the problem of classical singularities and the program of canonical quantization of gravity has exactly this task. But, it is a matter of fact, that so far the complicated structure of the canonical constraints has prevented from making progresses in the full theory.

Below, we describe the main features of canonical quantization, starting from a brief account of the prescriptions of the Dirac quantization procedure, then we digress on the old Wheeler-DeWitt quantum gravity and finally we go on to briefly introduce some aspects of Loop Quantum Gravity.

6.1 Dirac quantization procedure

The Dirac quantization procedure is a set of prescriptions aiming to consistently face the quantization of constrained physical system. One useful example to understand how the Dirac procedure works is the quantization of the electromagnetic field, which, as remarked in § 3.2 is a gauge theory of the compact $U(1)$ group.

Canonical quantum electrodynamics is usually constructed by imposing a gauge condition on the electromagnetic potential, as initially suggested by Fermi. Specifically, the quantization usually chosen is the so-called Lorentz gauge $\partial_\mu A^\mu = 0$. By imposing the gauge in the action, the definition of the momenta conjugated to the electromagnetic potential do not generate any primary constraint. So, the theory can be quantized and the gauge condition weakly imposed on the Fock quantum states, reducing the physical degrees of freedom to the two polarization of the photon. Another possibility is to consider the full canonical classical theory, as we have calculated in § 3.3 a first class Gauss constraint appears, and can be solved by a gauge fixing. A possible choice is the Coloumb gauge $\partial_\alpha A^\alpha = 0$. The dynamical variables (A_α, E^β) have to satisfy both the gauge fixing condition and the Gauss constraint. As we said in Sec. 4.2, the general conditions that a consistent gauge fixing has to satisfy imply that the Gauss constraint and the gauge condition form a set of second class constraints. Thus, the degrees of freedom in configuration space are exactly reduced to those corresponding to the two polarizations of the photon field. Finally, once the non-physical degrees of freedom have been eliminated, the system can be quantized by promoting the canonical variables to operators, satisfying the relations derived from promoting the Dirac brackets to quantum commutators.

This procedure works well if applied to linear physical systems, but it presents some complicated issues when applied to non-linear systems as, for example, gravity or Yang-Mills $SU(N)$ gauge theories for $N \geq 2$. The reason is connected with the existence of the so called Gribov ambiguities, which are produced by the complicated geometrical aspects of the complete phase space associated with the dynamics of gauge theories. In fact, in general, the geometry of the constraints

surface and gauge orbits could be such that the gauge fixing surface cuts some of the orbits more than once and it does not intersect some others at all [108]. So the gauge fixing surface works properly only locally, in general, it is impossible to find a global suitable gauge condition. This fact is generally referred as the Gribov obstruction and represents a shared characteristic of all the non-abelian gauge theories. Also gravity is affected by this problem, indeed, studying the classical canonical aspects of the theory, we pointed out that the attempts to solve the Cauchy problem in GR reveals the non-existence of global gauge conditions. In fact, it is possible to find a solution of the Cauchy problem at most locally.

For these reasons it is important to develop a theory of first class constraints without having the necessity to fix the gauge by using the procedure above. The way, suggested by Dirac, is to impose the first class constraints after the quantization, namely directly on the quantum states. In other words, the idea is to set up a Schrödinger-like equation by promoting the first class Hamiltonian to a quantum operator acting on the states of the theory. The classical first class constraints, in stead, are imposed on the state functional as a supplementary conditions, i.e.

$$\widehat{C}_I\,|\psi\rangle = 0\,. \tag{6.1}$$

The action of the first class Hamiltonian and the first class constraints on the state functional is dictated by the upgrade of the classical Dirac brackets to quantum commutators.

In this way, every quantum state remains unchanged under a transformation generated by the constraints, namely we are reintroducing the gauge invariance at a quantum level. In fact the condition above implies, as a consequence, that the quantum states are invariant under finite gauge transformations in the sector connected with the identity,[32] i.e.

$$\exp\left\{i\alpha_I\widehat{C}_I\right\}|\psi\rangle = |\psi\rangle\,. \tag{6.2}$$

We are assuming that the set of classical constraints C_I is first class, i.e.

$$\{C_I, C_J\} = f_{IJ}{}^K C_K\,, \tag{6.3}$$

if this relation is preserved by the quantization we have

$$\left[\widehat{C}_I, \widehat{C}_J\right] = i\hbar\widehat{f}_{IJ}{}^K\widehat{C}_K\,, \tag{6.4}$$

but in general it is possible that the first class conditions above show the presence of additional terms of quantum mechanical origin, i.e. we could have

$$\left[\widehat{C}_I, \widehat{C}_J\right] = i\hbar\widehat{f}_{IJ}{}^K\widehat{C}_K + \hbar^2\widehat{A}_{IJ}\,. \tag{6.5}$$

If this is the case the physical states, namely the states invariant under finite gauge transformation connected with the identity, must satisfy the additional condition

$$\widehat{A}_{IJ}\,|\psi\rangle = 0\,, \tag{6.6}$$

[32] The transformation properties of the physical states under "large gauge transformation", that is those not connected with the identity, are not contained in the action principle, indeed no constraint is generated by them, so requiring the invariance of the physical states under this larger class of transformation would be an extra assumption. See Appendix B for more details.

which has not a classical analogue and in general restrict the phase space too much. In particular if the operator $\widehat{A}_{IJ}$ is invertible, it would imply that the space of the physical states is empty. So, on the one hand, we cannot impose such a condition without drastically affecting the content of the theory. But, on the other hand, if we do not pose that condition the operator $\widehat{C}_I$ is not first class any longer, so it no longer generates gauge transformations. In other words, the gauge invariance is broken at a quantum level, i.e. the quantization of the system has produced a gauge anomaly. Summarizing, if quantum effects break down gauge invariance, then it is meaningless to search for gauge invariant physical states, i.e. we cannot impose equation (6.1). We can finally say that if a gauge anomaly is present the Dirac quantization method cannot be applied and a different quantization procedure, e.g. BRST, must be considered, with the hope that it could improve the situation in view of a consistent quantum theory.

6.2 Wheeler-DeWitt equation

The Wheeler-DeWitt equation is essentially the result of the Dirac quantization procedure to gravity. There is one peculiar aspect that makes this argument interesting, namely the fact that the first class Hamiltonian of GR is a combination of constraints, so that it does not exist the equivalent of the Schrödinger equation in QG. This aspect is well known and, as remarked previously, it is often referred as the problem of time. Nevertheless, the quantization can be formally performed, by following the standard procedure.

Firstly, let us define the smeared ADM variables,

$$Q(h) = \int_\Sigma d^3x\, h_{\alpha\beta} Q^{\alpha\beta}\,, \tag{6.7a}$$

$$P(f) = \int_\Sigma d^3x\, p^{\alpha\beta} f_{\alpha\beta}\,, \tag{6.7b}$$

where $Q^{\alpha\beta}$ and $f_{\gamma\delta}$ are smooth tensor valued function, while $h_{\alpha\beta}$ and $p^{\gamma\delta}$ are the canonical variables defined in § 4.1. As is well known the wave function depends only on half of the elementary variables, since in this case a natural separation between configuration variables and momenta exists, then the "polarization of the symplectic manifold" is pretty natural, i.e. the wave function will depend on $Q(h)$. Now, once a suitable quantum configuration space $\mathscr{C}$ has been introduced, it has to be equipped with the structure of an Hilbert space. This consists in choosing a suitable measure $d\mu_0$, in such a way that $\mathscr{C}$ becomes naturally an L_2 space. Obviously, the present Hilbert space does not know about the dynamics, so it will be referred as kinematic Hilbert space. Now the quantization proceeds in the usual way, namely requiring that the operator representation of the elementary variables, i.e. $\widehat{Q}(h)$ and $\widehat{P}(f)$, acting as linear operator on a common dense domain of the kinematic Hilbert space $\mathscr{H}(L_2, d\mu_0)$ generate an irreducible representation of the canonical commutation relation. In other words, we require that

$$\left[\widehat{Q}(h), \widehat{P}(f)\right] = i\hbar\widehat{\{Q(h), P(f)\}}\,. \tag{6.8}$$

As usual, $\widehat{Q}(h)$ operates by multiplication when valuated on the quantum configuration space, i.e.

$$\langle h|\, \widehat{Q}(h)\, |\psi\rangle = Q(h)\, \langle h|\psi\rangle = Q(h)\Psi(h)\,, \tag{6.9}$$

while $\widehat{P}(f)$ acts as a functional derivative operator, i.e.

$$\langle h|\,\widehat{P}(f)\,|\psi\rangle = \frac{\hbar}{i}\int d^3x\, f_{\alpha\beta}\frac{\delta}{\delta h_{\alpha\beta}}\langle h|\psi\rangle = \frac{\hbar}{i}\int d^3x\, f_{\alpha\beta}\frac{\delta}{\delta h_{\alpha\beta}}\Psi(h)\,. \tag{6.10}$$

Now, the naive quantization of the system follows from the translation of the classical constraints to quantum operators, in accordance with the prescription above, i.e.

$$\vec{H}(\vec{N})\approx 0 \qquad \Longrightarrow \qquad \widehat{H}(\vec{N})\Psi[Q(h)] = 0\,, \tag{6.11}$$

$$H(N)\approx 0 \qquad \Longrightarrow \qquad \widehat{H}(N)\Psi[Q(h)] = 0\,. \tag{6.12}$$

But this procedure presents a lot of shortcomings, some of them are of a general nature, while others are specific for the gravitational case [3]; we summarize in what follows the main ones.

- We know that in the construction of the quantum phase space functions we can arbitrarily add to the elementary variables terms proportional to the constant $\hbar$ without affecting the classical limit of the theory. This ambiguity in the choice of the phase space function is known as *factor ordering ambiguity*. Divergences can arise in gauge theories where a bad factor ordering is fixed, a simple example is provided by QED, in fact only after the choice of a suitable factor ordering the Hamiltonian operator results well defined, being otherwise divergent and nowhere defined.

- In general the divergences of an operator are of a worse kind and can be reabsorbed only after a regularization and renormalization procedure. It is worth noting that the renormalization is connected with the free possibility of adding localized terms to the quantum operators.

- It is important choose the factor ordering in such a way the quantum operators be self-adjoint, this is a crucial question in theories with true Hamiltonian, in order to guarantee that the eigenvalues of the operator be real and could be neglected in theories like gravity with a constrained Hamiltonian on condition that the eigenvalue zero is contained in the spectrum. Working with self-adjoint operators is however advantageous.

- The Hamiltonian quantum operator of GR depends neither polynomially nor analytically on the metric, this fact poses a serious problem because in general operator valued distributions multiplied at the same point gives divergent results, so a regularization procedure is required; even worse the presence of distribution in the denominator poses a more difficult problem of formal definition. Anyway we can try to seek formal solutions of the Wheeler-DeWitt operator, being aware of the fact that a regularization procedure is however required.

- Let us suppose to neglect initially the amount of technical issues and bad definitions of the Wheeler-DeWitt approach, hopefully solvable after a suitable regularization and renormalization procedure. Let us suppose that we succeed in finding a solution of this equation, then a conceptual and interpretative issue arises, indeed, we do not know, in general, how to interpret the result. This fact is strictly connected with the problem of time and, even though a quantum mechanics without time can be constructed, then a further amount of work is required to correctly interpret the result of quantization. A possible solution to this interpretative problem is relational evolution, namely, a Schrödinger-like equation can be constructed

in place of the Wheeler-DeWitt one, showing that the presence of timein the quantum equations reflect on the classical theory generating matter fields [28]. The other way around is to couple matter to gravity, e.g. a free scalar field or a dust of particles and extract a relational time variables related to the momentum of the scalar field, by using the Brown-Kuchař procedure [57, 30]. It is worth noting that in this kind o approaches the evolution parameter is not external with respect to the physical system as in background dependent theories, here the evolution is referred to an "internal time" as one would expect in a background independent theory.

- Finally we stress that a central issue should be faced and regards the presence of gauge anomalies. In GR, indeed, the problem is particularly complicate, because the group structure constants are replaced by structure function depending on the metric field, so in the computation of equation (6.4) it is possible that an anomalous factor comes out.

All the above described issues have led to seek for a better formalization of the problem. In particular since the choice of the elementary variables is dictated only by the convenience and simplicity of the resulting constraints, the use of Ashtekar-Barbero variables turned out to be extremely useful in facing some of these problems. In fact, a consistent anomaly free quantization is possible and has given interesting results on the spectrum of regularized self-adjoint operators exactly in line with the results one would expect from a quantum General Relativity theory [1].

6.3 The program of Loop Quantum Gravity

As we showed in Section 5, canonical GR in the Ashtekar–Barbero formulation is characterized by the following set of first class constraints:

$$G_i = \partial_\alpha P^\alpha_{\ k} + \varepsilon_{ki}{}^{j} \mathscr{A}^i_\alpha P^\alpha_{\ j} \approx 0\,, \tag{6.13}$$

$$C_\alpha = P^\gamma_i \mathscr{F}^i_{\alpha\gamma} \approx 0\,, \tag{6.14}$$

$$C = \frac{1}{2} P^\alpha_i P^\gamma_j \left[\varepsilon^{ij}{}_k \mathscr{F}^k_{\alpha\gamma} - 2\left(\beta^2+1\right) K^i_{[\alpha} K^j_{\gamma]}\right] \approx 0\,, \tag{6.15}$$

where, to simplify the notation, we defined $P^\gamma_i = {}^{(\beta)}E^\gamma_i$ and we dropped the upper left β in the connection, i.e. $\mathscr{A}^i_\alpha = \beta K^i_\alpha + \Gamma^i_\alpha$. For convenience we rewrite here the symplectic structure as well

$$\left\{\mathscr{A}^i_\alpha(t,x), P^\gamma_k(t,x')\right\} = \delta^i_k \delta^\gamma_\alpha \delta(x,x')\,, \quad \left\{\mathscr{A}^i_\alpha(t,x), \mathscr{A}^k_\gamma(t,x')\right\} = 0\,, \quad \left\{P^\alpha_i(t,x), P^\gamma_k(t,x')\right\} = 0\,. \tag{6.16}$$

The above formulation of the classical canonical GR is the starting point of LQG. Remarkably, the AB formulation, allowing to rewrite GR as a theory of connections, provides a sort of *kinematical unification* with the other forces, successfully described by Yang–Mills gauge theories, namely as theories of connections valued on compact groups of the $SU(N)$ family. Nevertheless, it has to be emphasized that, from a dynamical perspective, a profound different can be immediately recognized: In Yang–Mills theories the metric of space-time plays a central role, while in QG no background metric is assumed *a priori*, better to say that in QG there is no space-time at all.

According to the Dirac prescriptions previously described, the theory can be quantized by suitably define a quantum representation of the canonical algebra and then imposing the operator

translation of the canonical constraints on the state functional, $\Psi(\mathscr{A})$, representing the states of the theory. A possible choice for the representation is the one suggested by the old Wheeler–DeWitt approach described before, which, even though formally correct, cannot be made rigorous.

The program of Loop Quantum Gravity goes, in fact, in a different direction. The idea is to use a different set of fundamental variables,which are more suitable for quantization. In this respect, let us introduce the holonomies, $h_\gamma[\mathscr{A}]$, of the connection $\mathscr{A}^i_\alpha$ and the fluxes, $P[\Sigma, f]$, of the momentum P^γ_k respectively as

$$h_\gamma[\mathscr{A}] = \mathscr{P}\exp\left\{-\int_\gamma \mathscr{A}^i_\alpha \tau_i \frac{dx^\alpha}{ds} ds\right\}, \tag{6.17}$$

and

$$P[S, f] = \int_S P^\alpha_i f^i \varepsilon_{\alpha\beta\gamma} \frac{dx^\beta}{ds_1}\frac{dx^\gamma}{ds_2} ds_1 ds_2\,, \tag{6.18}$$

where $\tau_i = \frac{1}{2}\sigma_i$ are the generators of the $SU(2)$ group, σ_i being the Pauli matrices, while f^i is an $SU(2)$ valued smearing function. The symbol γ is the parametric oriented curve on which the holonomy is valued, while S represent a 2-dimensional surface in Σ^3. It is easy to demonstrate that the holonomy has the following properties:

$$h_{\gamma_1\circ\gamma_2}[\mathscr{A}] = h_{\gamma_1}[\mathscr{A}]\, h_{\gamma_2}[\mathscr{A}] \quad \text{and} \quad h_{\gamma^{-1}}[\mathscr{A}] = h^{-1}_\gamma[\mathscr{A}]\,, \tag{6.19}$$

where $\gamma_1 \circ \gamma_2$ corresponds to join together the end point of γ_1 and the initial point of γ_2, while γ^{-1} denotes a change in the orientation of the curve. Notice that the holonomy is an element of the group $SU(2)$.

Our purpose is to describe the canonical dynamics of the gravitational system by using the new variables defined above. The first step in this program is the evaluation of the Poisson brackets between the new configuration observables $h_\gamma[\mathscr{A}]$ and momenta $P[S, f]$. Specifically, the Poisson brackets between two configuration variables vanish, while, considering that any edge γ with $\gamma \cap S \neq \emptyset$ can be trivially written as the union of elementary edges which either lie in S, or intersect S in exactly one of their end-points, then, for each of these elementary edges γ which intersect S at a point p, we have:

$$\{h_e[\mathscr{A}], P(\Sigma, f)\} = -\kappa(S,\gamma) \times \begin{cases} h_\gamma[\mathscr{A}]\,\tau_i f^i(p) & \text{if } p \text{ is the source of } \gamma \\ -f^i(p)\tau_i h_\gamma[\mathscr{A}] & \text{if } p \text{ is the target of } \gamma \end{cases}, \tag{6.20}$$

where

$$\kappa(S,\gamma) = \begin{cases} 0 & \text{if } \gamma\cap\Sigma = \emptyset \text{ or } \gamma\cap\Sigma = \gamma \\ \pm 1 & \text{if } \gamma\cap\Sigma \neq \emptyset \end{cases}. \tag{6.21}$$

Notice that the Poisson brackets between two momenta is non-trivial; the reason being related to the fact that new variables are still distributional quantities even though they are smeared, since the smearing is made respectively on one and two dimensional functions for holonomies and fluxes, so that a particular care has to be taken in handling with Poisson brackets involving two momenta.

According to the usual procedure, the quantum fundamental operators in the auxiliary Hilbert space will be required to satisfy the algebra originating from the commutation relations. So, one

of the main issue one has to face is to find a consistent representation for the quantum algebra. It is worth saying that the properties of the resulting quantum geometry can be extracted through the momentum operators $\widehat{P}[\Sigma, f]$ in that representation, it is, in fact, related to the classical orthonormal basis via the definitions given above. Surprisingly enough, the diffeomorphisms invariance requirement sort out a unique representation of the quantum algebra. This result is often referred as LOST theorem by the acronym of Lewandowski, Okolow, Sahlmann, and Thiemann [109] (an independent result with the same physical content was given by Fleischhack [110])

After this brief introduction, a schematic description of how the Dirac procedure applies to gravity in the LQG program is in order.

- Holonomies of the connection are chosen as configuration variables. In particular, the auxiliary Hilbert space can be constructed and consists of a set of functionals of the holonomies, square integrable in the Ashtekar–Lewandowski measure.

- The Gauss and vectorial constraints have a natural action on the states of the theory. In particular, the space of solutions of the Gauss and vectorial constraints is well understood.

- The situation becomes much more involved as far as the scalar constraint is considered. The main problem is that the scalar constraint is highly non-linear. Some strategies have been developed to deal with the scalar constraint, particularly through the Thiemann's "master constraint program" [111], but many unsolved issues are still present in the theory. Nevertheless, well defined version of the scalar constraint (in symmetric systems) have been constructed, leading to striking results, which answer some long-standing question about (quantum) gravity.

- At the present stage of the development of the theory, physical observables are known only in some special cases.

In order to be more specific, let us give a brief account of these steps by introducing the so-called spin-networks representation.

In order to be as clear as possible, let us start by defining an abstract graph, Γ, which is intended as a collection of paths $\gamma \in \Sigma$ meeting at most at their end-points. Given a graph Γ, we denote by $1, 2, \cdots, N$ its edges γ, i.e. $\Gamma = \bigcup_{k=1}^{N} \gamma_k$. We call cylindrical functions of generalized connections a functional of the holonomies valued on the edges of the graph to complex numbers,

$$F : SU(2)^N \to \mathbb{C}\,, \tag{6.22}$$

defined as

$$\psi_{\Gamma,F}[\mathscr{A}] = F\left(h_{\gamma_1}[\mathscr{A}], \cdots, h_{\gamma_N}[\mathscr{A}]\right). \tag{6.23}$$

As a simple example, consider a closed loop γ and the functional

$$W_\gamma[\mathscr{A}] = \psi_{\gamma,\mathrm{tr}}[\mathscr{A}] = \mathrm{tr}\left\{h_\gamma[\mathscr{A}]\right\}, \tag{6.24}$$

this is often referred as Wilson loop and belongs to the space of cylindrical functions, i.e. $W_\gamma[\mathscr{A}] \in \mathrm{Cyl}_\Gamma$.

Let us now denote as $\mathscr{F}$ the linear space of all functionals $\psi_{\Gamma,f}[A]$ for all Γ and f. The space $\mathscr{F}$ can be equipped with a scalar product through the following procedure. Define a new state μ_{AL} as

$$\mu_{AL}\left(\psi_{\Gamma,F}\right) = \int \prod_{\gamma\subset\Gamma} dh_\gamma F\left(h_{\gamma_1},\ldots,h_{\gamma_N}\right), \tag{6.25}$$

where dh_e is the normalized Haar measure of $SU(2)$. The state $\mu_{AL}\left(\psi_{\Gamma,F}\right)$ is normalized, i.e., $\mu_{AL}(1)=1$, because the Haar measure is normalized, and positive, i.e.

$$\mu_{AL}\left(\overline{\psi}_{\Gamma,F}\psi_{\Gamma,F}\right) = \int \prod_{\gamma\subset\Gamma} dh_\gamma F^*\left(h_{\gamma_1},\ldots,h_{\gamma_N}\right) F\left(h_{\gamma_1},\ldots,h_{\gamma_N}\right) \geq 0. \tag{6.26}$$

As a consequence, a scalar product on $\mathscr{F}$ can be defined as

$$\left\langle \psi_{\Gamma,F} | \psi_{\Gamma',F'} \right\rangle = \mu_{AL}\left(\overline{\psi_{\Gamma,F}}\psi_{\Gamma',F'}\right) = \int \prod_{e\subset\Gamma\cup\Gamma'} dh_e F^*\left(h_{e_1},\ldots,h_{e_N}\right) F'\left(h_{e_1},\ldots,h_{e_N}\right). \tag{6.27}$$

Usually μ_{AL} is called Ashtekar-Lewandowski measure. The above scalar product gives to the kinematical state space the structure of an auxiliary Hilbert space. Furthermore, the kinematical scalar product is invariant under the automorphisms of the local bundle and 3-diffeomorphisms, so that the kinematical state space carries a *unitary* representation of local $SU(2)$ and 3-diffeomorphisms.

At this point, the states of the theory, represented by the functionals $\psi_{\Gamma,F}\in\mathscr{F}$, have to be restricted by imposing the constraints. In particular, by imposing the Gauss and vectorial constraints, the state of the theory will be invariant under the local $SU(2)$ symmetry, correlated with the (double cover of the) group of spatial rotations, and under the 3-diffs. But, in order to rigorously implement in the quantum theory the following formal equations

$$\widehat{G}_i\psi_{\Gamma,F}[\mathscr{A}] = 0, \tag{6.28}$$

$$\widehat{C}_\alpha\psi_{\Gamma,F}[\mathscr{A}] = 0, \tag{6.29}$$

it is necessary to find a quantum operator representation of the classical elementary variables. In this respect, the introduction of a suitable basis for the kinematical states space is particularly useful. Without entering in the details, we can just use the result of the Peter–Weyl theorem, stating that a basis on the Hilbert space of L_2 functions on $SU(2)$ is given by the matrix elements of the irreducible representations of the group. We indicate the matrix elements in the j-representation as $R^{(j)m}{}_n$, where $m,n,\ldots$ denote the matrix elements of the specific representation. Therefore, a basis for each graph Γ is simply obtained by "tensoring" the basis above, i.e.

$$\psi_{\Gamma,F}[\mathscr{A}] = \sum_{j_1\cdots j_N} f^{m_1\cdots m_N,n_1\cdots n_N}_{j_1\cdots j_N} R^{(j_1)}_{m_1n_1}\left(h_{\gamma_1}[\mathscr{A}]\right)\cdots R^{(j_N)}_{m_Nn_N}\left(h_{\gamma_N}[\mathscr{A}]\right). \tag{6.30}$$

The symbol j labels the irreducible representation of $SU(2)$, which can be characterized by half-integer spins pictorially associated to each edge of the graph Γ. The coefficients of the expansion $f_{j_1\cdots j_N}$ are restricted by the gauge invariance. We can easily construct a simple example for one of the lements of the sum above by considering a graph Γ made up of three edges $\gamma_1,\gamma_2,\gamma_3$, to which we respectively associate the representations $1,\frac{1}{2},\frac{1}{2}$ of the group $SU(2)$, then we have:

$$\psi_{(\gamma_1\cup\gamma_2\cup\gamma_3)}[\mathscr{A}] = R^{(1)}\left(h_{\gamma_1}[\mathscr{A}]\right)^{ij} R^{(1/2)}\left(h_{\gamma_2}[\mathscr{A}]\right)_{AB} R^{(1/2)}\left(h_{\gamma_3}[\mathscr{A}]\right)_{CD} \sigma_i^{AC}\sigma_j^{BD}, \tag{6.31}$$

where $i,j=1,2,3$ are vector indexes, while $A,B,C,D=1,2$ are spinor indexes, and σ_i^{AC} are the 2×2 Pauli matrices. It is easy to check that the expression above is gauge invariant, in fact, the Pauli matrices are invariant tensors in the tensor product representation $1\otimes 1/2\otimes 1/2$ acting on the nodes of the graph $\Gamma=\gamma_1\cup\gamma_2\cup\gamma_3$. Generally, we can write a gauge invariant state function as

$$\psi_\Gamma[\mathscr{A}]=\left(\bigotimes_l R^{(j_l)}\left(h_{e_l}[\mathscr{A}]\right)\right)\cdot\left(\bigotimes_n i_n\right),$$

where the invariant tensors i_n assigned on the nodes of the graph are called intertwiners between the representations $j_1,\cdots,j_N$ associated to the edges joining in a node. The graph Γ, the labels j_k "coloring" the links and the intertwiners i_n "coloring" the nodes completely define a state, in particular a state defined by the triplet (Γ,j_k,i_n) is called *spin-network*.

Now, in order to physically characterize the states of the theory, represented by spin-networks, we construct some geometrical operators acting on them, defining their action on the single holonomy (the action on the complete state can be extracted by composition). The first operator we wish to define is the momentum operator, easily correlated to the triad which has a precise geometrical interpretation as stressed above. Specifically, the momentum P_i^α naturally acts on the holonomies as a functional derivative, i.e.

$$\widehat{P}[S,f]=\frac{\hbar}{i}\int_S ds^\beta ds^\gamma \varepsilon_{\alpha\beta\gamma}f^k\frac{\delta}{\delta\mathscr{A}_\alpha^k}, \tag{6.32}$$

where we introduced the Planck constant $\hbar$ (it is worth recalling that we set $8\pi G=1$ and $c=1$ from the very beginning). To compute the result of the action of the momentum operator on the holonomy, let us firstly note that

$$\frac{\delta}{\delta\mathscr{A}_\alpha^i(y)}h_\gamma[\mathscr{A}]=\int ds\,\delta\left(x(s),y\right)\frac{dx^\alpha}{ds}h_{\gamma_1}[\mathscr{A}]\,\tau_i h_{\gamma_2}[\mathscr{A}], \tag{6.33}$$

namely, the action of the functional derivative "cuts" the link γ at the point y where the derivative operator acts, inserting an $SU(2)$ generators in the middle of the holonomies valued on the two resulting pieces of the original holonomy. Given that, we easily get

$$\widehat{P}[S,f]\,h_\gamma[\mathscr{A}]=\frac{\hbar}{i}\int ds^1ds^2ds^3\varepsilon_{\alpha\beta\gamma}\frac{dx^\alpha}{ds^1}\frac{dx^\beta}{ds^2}\frac{dx^\gamma}{ds^3}\delta\left(\underline{x}(s^3),\underline{y}(s^1,s^2)\right)f^i h_{\gamma_1}[\mathscr{A}]\,\tau_i h_{\gamma_2}[\mathscr{A}]. \tag{6.34}$$

Notice that the integration is made in 3-dimensions so that the δ distribution can be safely integrated, moreover its presence ensures that the above integral vanishes if the link γ does not intersect the surface S. The points of intersection between links and surfaces are usually called *punctures*. By using the scalar product defined before, one can show that, in fact, the triad operator is self-adjoint.

After having established the action of the triad operator on the holonomy, we can now define an interesting quantum operator, namely the area operator $\widehat{\mathbf{A}}$. Let me initially refer to a single holonomy in the fundamental representation. Neglecting some (very important) subtleties, we have that:

$$\widehat{P}_i[S]\,\widehat{P}^i[S]\,h_\gamma[\mathscr{A}]=-\hbar^2 h_{\gamma_1}[\mathscr{A}]\,\tau^i\tau_i h_{\gamma_2}[\mathscr{A}]=\frac{3}{4}\hbar^2 h_\gamma[\mathscr{A}].$$

Considering a general irreducible representation of $SU(2)$, the action of the square of the momentum operator on the holonomy turns out to be

$$\widehat{P}_i[S]\,\widehat{P}^i[S]\,R^{(j)}\left(h_\gamma[\mathscr{A}]\right) = -\hbar^2\, j\,(j+1)\,R^{(j)}\left(h_\gamma[\mathscr{A}]\right)\,,$$

where we assumed that the surface S is punctured only once. For a generic surface S in space the situation is slightly more complicated and can be reduced to the previous simple case by the following procedure. Divide the surface S in N cells and consider the full area as a limit, $A_S = \lim_{N\to\infty} A_S^N$, where

$$A_S^N = \beta^2 \sum_{I=1}^{N} \sqrt{|P_i(S_I)P^i(S_I)|}\,, \tag{6.35}$$

$P_i(S_I)$ being the flux through the I-th cell. The factor β stems from the definition of the classical momentum, $P_i^\alpha = {}^{(\beta)}E_i^\alpha = E_i^\alpha/\beta$, remembering that the area operator has to be defined with the geometrical triad E_i^α. So that, the quantum area operator then simply becomes $\widehat{A}_S = \lim_{N\to\infty}\widehat{A}_S^N$.

Considering the result obtained above, the area operator turns out to be diagonal in the basis of spin-networks and reintroducing the physical constant, its spectrum is given by

$$\widehat{A}_S\,|\psi\rangle = \beta\,\ell_{\rm Pl}^2 \sum_p \sqrt{j_p\,(j_p+1)}\,|\psi\rangle\,, \tag{6.36}$$

where $\ell_{\rm Pl}$ is the Planck length. Notice that the cellular decomposition is chosen so that in the limit $N\to\infty$ each cell is punctured at most in a single point.

An analog procedure allows to define the kinematic volume operator $\widehat{V}$. The classical volume of a region, $\mathscr{R}$, of space can be written as

$$V(\mathscr{R}) = \int_{\mathscr{R}} d^3x \sqrt{\frac{1}{3!}\left|\varepsilon_{\alpha\beta\gamma}\varepsilon^{ijk}E_i^\alpha E_j^\beta E_k^\gamma\right|}\,, \tag{6.37}$$

which corresponds to a complicated quantum operator. In particular, to calculate the spectrum of the volume operator, a cellular decomposition analog to that performed in the case of the area operator results to be very useful. Specifically, the volume operator

$$\widehat{\mathbf{V}} = \lim_{N\to\infty} \beta^{3/2}\ell_{\rm Pl}^3 \int_{\mathscr{R}} d^3x \sqrt{\frac{1}{3!}\left|\varepsilon_{\alpha\beta\gamma}\varepsilon^{ijk}\widehat{P}_i^\alpha \widehat{P}_j^\beta \widehat{P}_k^\gamma\right|}\,, \tag{6.38}$$

can be evaluated and its spectrum results to be discrete depending on the quantum numbers coloring the nodes of the graph. It is worth noting that the Gauss constraint obliges the flux operator at a node to vanish, so that the volume of a three-valent node vanishes as well.

The fact that the quanta of area depend on the quantum numbers associated with the links or edges of the graph, while the quanta of volume depend on the quantum numbers of the nodes, a physical interpretation of a graph can be given. Specifically, any node of a graph represent a chunk of volume of the quantum space-time, while links describe the quantum properties of the surfaces between two volumes. This means that quantum space-time at a kinematical level is made up of quanta of volume separated by quanta of area.

It remains to describe how the canonical constraints can be implemented on the quantum states of the theory. This argument deserves to be carefully analyzed and is far from the scopes

of this paper. Nevertheless, it is important to say that the Gauss and vectorial constraints can be implemented and solved at the present stage of the development of the theory. They, in fact, have a pretty natural action on spin-networks, but serious difficulties appear as soon as the scalar constraint is regarded. Here I want to digress on the general procedure used to deal with such a problem.

As we said before, the space $\mathscr{F}$ can be equipped with the structure of an auxiliary Hilbert space by defining a normalized positive defined kinematical scalar product. Our final purpose is to define a physical scalar product, namely between states which satisfy the constraints. In order to give a general brief description of the problem, let us refer to a general constraint $\widehat{\mathscr{C}}$ and define the following projection operator

$$S_{\widehat{\mathscr{C}}} = \int \delta N e^{iN\widehat{\mathscr{C}}} . \tag{6.39}$$

This operator allows to formally define a physical inner product (for details and rigorous procedures see [1]). The idea is that $S_{\widehat{\mathscr{C}}}$ is formally equivalent to a delta function of the constraint operator $\widehat{\mathscr{C}}$, so it can select the states of the theory that satisfy the constraint, i.e.

$$\left(\left\langle \psi_{\Gamma,F} | \psi_{\Gamma',F'} \right\rangle\right)_{\text{Phys}} = \left\langle \psi_{\Gamma,F} | \delta\left(\widehat{\mathscr{C}}\right) | \psi_{\Gamma',F'} \right\rangle . \tag{6.40}$$

So, formally, the physical inner product corresponds to the following expression

$$\left\langle \psi_{\Gamma,F} | S_{\widehat{\mathscr{C}}} | \psi_{\Gamma',F'} \right\rangle = \left\langle \psi_{\Gamma,F} \left| \int \delta N e^{iN\widehat{\mathscr{C}}} \right| \psi_{\Gamma',F'} \right\rangle = \left(\left\langle \psi_{\Gamma,F} | \psi_{\Gamma',F'} \right\rangle\right)_{\text{Phys}} . \tag{6.41}$$

which can be made rigorous through the group averaging procedure.

Concluding, we stress that an important result obtained in the framework of LQG is the discreteness of the eigenvalues of geometric operators. But, simultaneously, this fact introduces an interesting issue. In fact, in the classical theory, to associate a precise physical meaning to geometrical quantities as the area and volume of a region of space-time, one has to define the surfaces and regions operationally, e.g., by using matter fields. Once this is done, one can simply calculate values of these observables using the geometrical formulas. An analogous situation characterizes the quantum theory. For instance, the area of the isolated horizon is a Dirac observable in the classical theory and the application of the quantum geometry area formula to this surface leads to physical results [1]. In this situation, the operators and their eigenvalues correspond to the proper lengths, areas an volumes of physical relevant objects.

Finally, answering to a question asked more than once during the School, it is important to emphasize that no tension exists between the discreteness of the eigenvalues of geometrical operators and Lorentz invariance. A simple example from quantum mechanics should clarify this point. Consider, e.g., the angular momentum operator in ordinary quantum mechanics, its eigenvalues are discrete and this is perfectly compatible with the rotational invariance of the theory.

Acknowledgments

The author wants to express his gratitude to the organizers of the 5th International School on Field Theory and Gravitation. Their exquisite hospitality, friendliness, and the stimulating atmosphere they have been able to create, have made this school extremely useful for both students and speakers. Particularly, for me and the other people who experienced the unexpected bath in the alligators' lake, this school has been and will literally be unforgettable. In this respect, I want to thank especially Carlos Pinheiro and Gentil Pires for their priceless help in solving all the bureaucratic troubles and for their friendly support.

I want also to thank Abhay Ashtekar for having suggested my name to the Organizers of the school to substitute him.

This research was supported in part by NSF grant PHY0854743, The George A. and Margaret M. Downsbrough Endowment and the Eberly research funds of Penn State.

A. Differential forms

In this appendix, we have collected the main definitions and formulas useful to deal with differential forms. This language has become pretty common in the literature and sometimes confuses people used to the index notation. We should say that in a theory like like GR, where the coordinates have no any physical meaning, differential forms represent the most natural formalism, even though in some problems the index notation is preferable. So, often one has the necessity in treating physical problems to switch notation, going from forms to indexes and vice versa. This has induced me to collect some main formulas in few pages, with the hope they will be useful to the readers as they has been so for me.

As a disclaimer, we have to stress that many different notations are used in the literature, anyone valid and motivated by precise choices. The definitions and formulas below are in accordance with a notation commonly used in physics and refers to an arbitrary number of dimensions (unless differently specified) and to any signature of the n-dimensional manifold.

Let M be an n-dimensional manifold of signature s.[33] Let me denote as $\Lambda^p(T^*M^n)$ the $\binom{n}{p}$-dimensional space of p-forms on the cotangent bundle. Let $e^a = e^a_\mu dx^\mu$ be the gravitational 1-form defined in § 5.1, transforming under the vectorial representation of the local symmetry group $SOn-s,s$. The canonical basis for $\Lambda^p(T^*M^n)$ is easily induced by the local basis by considering the wedge product of the 1-forms e^a. Specifically any p-form $\eta \in \Lambda^p_*(TM^n)$ can be expanded on the canonical basis according to the following definition

$$\eta = \frac{1}{p!}\,\eta_{[a_1\cdots a_p]}\, e^{a_1}\wedge\cdots\wedge e^{a_p}\,, \tag{A.1}$$

where the square brackets denote anti-symmetrization. In other words, the p-form η is a smooth map that associates to any $x \in M$ an antisymmetric tensor of type $(0,p)$, i.e.

$$\eta(e_{a_1},\cdots,e_{a_p}) = \eta_{[a_1\cdots a_p]}\,, \qquad \forall\,\eta \in \Lambda^p(T^*M^n)\,. \tag{A.2}$$

By definition we have that

$$e^{a_1}\wedge\cdots\wedge e^{a_p} = e^{\ a_1}_{[\mu_1}\cdots e^{\ a_p}_{\mu_p]} dx^{\mu_1}\wedge\cdots\wedge dx^{\mu_p} = (-1)^s \varepsilon^{a_1\cdots a_p} dV\,, \tag{A.3}$$

where dV denotes the volume element.

In order to give a precise operative meaning to the definition (A.2), we introduce the internal or scalar product between a differential p-form and a vector v defined on the tangent bundle TM. Let $\omega \in \Lambda^p(T^*M^n)$ and $v = v^a\tilde{e}_a$, where $\tilde{e}_a = e^\mu_a\partial_\mu$ is a vector field, then we can evaluate the differential form on the vector field, obtaining a $(p-1)$-form, according to the following prescription:

$$\begin{aligned}
\omega(v) &= \frac{1}{p!}\,\omega_{a_1\cdots a_p} v^b\,[e^{a_1}\wedge\cdots\wedge e^{a_p}]\lrcorner e_b \\
&= \frac{1}{p!}\sum_{i=1}^{p}(-1)^{p-i}\delta^{a_i}_b\,\omega_{a_1\cdots a_i\cdots a_p} v^b e^{a_1}\wedge\cdots\wedge e^{a_{i-1}}\wedge e^{a_{i+1}}\wedge\cdots\wedge e^{a_p} \\
&= \frac{1}{(p-1)!}\,\omega_{a_1\cdots a_{p-1}b} v^b e^{a_1}\wedge\cdots\wedge e^{a_{p-1}}\,,
\end{aligned} \tag{A.4}$$

[33] Namely, s corresponds to the number of minus signs in the metric.

where in the last line we moved the index saturated with the components of the vector v on the right by using the antisymmetry of the indexes of ω and renamed the others. By using the above formula for the internal product, the definition (A.2) can be easily derived as a trivial consequence.

Let us now introduce the exterior or wedge product "$\wedge$". It is defined as producing a map $\wedge : \Lambda^p(T^*M^n) \times \Lambda^q(T^*M^n) \to \Lambda^{p+q}(T^*M^n)$ $(p+q \leq n)$ according to the following prescription:

$$\begin{aligned}\omega \wedge \eta &= \frac{1}{p!q!}\omega_{[a_1\cdots a_p]}\eta_{[b_1\cdots b_q]}e^{a_1}\wedge e^{a_p}\wedge e^{b_1}\wedge e^{b_q} \\ &= \frac{1}{(p+q)!}\left(\frac{(p+q)!}{p!q!}\,\omega_{[a_1\cdots a_p}\eta_{a_{p+1}\cdots a_{p+q}]}\right)e^{a_1}\wedge e^{a_p}\wedge e^{a_{p+1}}\wedge e^{a_{p+q}}\,, \end{aligned} \tag{A.5}$$

so, the components of the resulting $(p+q)$-form are

$$\omega\wedge\eta(e_{a_1},\cdots,e_{a_{p+q}}) = \frac{(p+q)!}{p!q!}\,\omega_{[a_1\cdots a_p}\eta_{a_{p+1}\cdots a_{p+q}]}\,. \tag{A.6}$$

Another useful operator we want to introduce is the so-called *Hodge dual*, usually denoted by the symbol "$\star$". We prefer to define initially the action of the Hodge dual on the canonical basis p-form $e^{a_1}\wedge\cdots\wedge e^{a_p}$, i.e., $\star : \Lambda^p(T^*M^n) \to \Lambda^{n-p}(T^*M^n)$, according to

$$\begin{aligned}\star\left(e^{a_1}\wedge\cdots\wedge e^{a_p}\right) &= \frac{1}{(n-p)!}\varepsilon_{a_{p+1}\cdots a_n}{}^{a_1\cdots a_p}e^{a_{p+1}}\wedge\cdots\wedge e^{a_n} \\ &= \frac{1}{(n-p)!}(-1)^{p(n-p)}\varepsilon^{a_1\cdots a_p}{}_{a_{p+1}\cdots a_n}e^{a_{p+1}}\wedge\cdots\wedge e^{a_n}\,. \end{aligned} \tag{A.7}$$

The above definition of the Hodge dual is different from th standard one, but we find it more convenient for a reason that will be clear soon. An interesting consequence, in fact, is that the wedge product a p-form with its hodge dual generates the volume element according to the following formula:

$$\star\left(e^{a_1}\wedge\cdots\wedge e^{a_p}\right)\wedge\left(e_{b_1}\wedge\cdots\wedge e_{b_p}\right) = p!\delta^{a_1\cdots a_p}_{[b_1\cdots b_p]}dV \tag{A.8}$$

where dV is the natural volume element on the n-dimensional manifold, i.e. $dV = dx^1\wedge\cdots\wedge dx^n$. It is worth noting that no dependence on the signature or dimensions appear in the formula above, this is particularly convenient to rewrite actions in terms of differential forms. We note that applying the dual operator twice we obtain the initial form apart for a possible sign factor, i.e.

$$\star\star\left(e^{a_1}\wedge\cdots\wedge e^{a_p}\right) = (-1)^{s+p(n-p)}e^{a_1}\wedge\cdots\wedge e^{a_p}\,. \tag{A.9}$$

Now, using the definitions (A.1) and (A.7), we can easily extract the expression of the dual of a generic p-form. Specifically,

$$\star\omega = \frac{1}{p!}\,\omega_{a_1\cdots a_p}\star(e^{a_1}\wedge\cdots\wedge e^{a_p}) = \frac{1}{(n-p)!}\left(\frac{1}{p!}(-1)^{p(n-p)}\omega_{a_1\cdots a_p}\varepsilon^{a_1\cdots a_p}{}_{a_{p+1}\cdots a_n}\right)e^{a_{p+1}}\wedge\cdots\wedge e^{a_n}\,. \tag{A.10}$$

In other words, the dual of a generic p-form is the $(n-p)$-form of components

$$\star\omega(e_{a_1},\cdots e_{a_{n-p}}) = \frac{1}{p!}(-1)^{p(n-p)}\omega_{a_1\cdots a_p}\varepsilon^{a_1\cdots a_p}{}_{a_{p+1}\cdots a_n} \tag{A.11}$$

The formula above makes it possible to calculate the wedge product between a the p-form η and the $(n-p)$-form $\star\omega$. So let $\omega,\ \eta \in \Lambda^p_*(TM^n)$ be two p-form, we have:

$$\star\omega\wedge\eta = \frac{1}{p!}\omega_{a_1\cdots a_p}\eta^{a_1\cdots a_p}dV\,. \tag{A.12}$$

So, apart for the factor $1/p!$ the wedge product in (A.12) corresponds to the scalar product between the components of the two p-forms multiplied by the natural volume element. This can be rewritten as:

$$\star\omega\wedge\eta = \frac{1}{p!}\left(\omega(e_{a_1},\cdots,e_{a_p}),\eta(e_{a_1},\cdots,e_{a_p})\right)dV\,, \tag{A.13}$$

where the symbol $(\ldots,\ldots)$ denotes the internal product. We remark that wedging the p-form with the canonical basis the factorial of p disappears.

Another fundamental operator we want to define is the exterior derivative operator d, which is a map from $\Lambda^p(T^*M^n)$ to $\Lambda^{p+1}(T^*M^n)$ defined as

$$\Lambda^{p+1}(T^*M^n)\ni\eta = d\omega = \frac{1}{(p+1)!}\left((p+1)\partial_{[b}\omega_{a_1\cdots a_p]}\right)e^b\wedge e^{a_1}\wedge\cdots\wedge e^{a_p}\,,\quad \omega\in\Lambda^p(T^*M^n)\,, \tag{A.14}$$

where, as usual, we contained in parentheses the components of the resulting $(p+1)$-form. By the definition given above we can immediately extract an important property of the exterior derivatives, i.e. $d\circ d = 0$, namely the composition of two derivative operators is the vanishing operator. Moreover, assuming that $\omega\in\Lambda^p_*(TM^n)$ and $\eta\in\Lambda^q_*(TM^n)$, it is very easy to show the following formula

$$d(\omega\wedge\eta) = (d\omega)\wedge\eta + (-1)^p\omega\wedge d\eta\,. \tag{A.15}$$

The local $SO(n-s,s)$ symmetry requires the definition of a covariant derivative. We require that the covariant derivative operator acting on $SO(s,n-s)$ valued p-forms generates $SO(s,n-s)$ valued $(p+1)$-forms, namely the derivative operator has to transform in the adjoint representation of the local symmetry group. In this respect, let us introduce a $SO(s,n-s)$ valued connection 1-form ωab and define the new derivative operator $d^{(\omega)}$ as

$$d^{(\omega)}\cdots = d\cdots + \omega\wedge\cdots\,. \tag{A.16}$$

We claim that the above derivative operator has exactly the property required, as can be easily demonstrated. In order to operatively define the covariant operator, we firstly specify its action on the basis 1-form e^a, we have

$$d^{(\omega)}e^a = de^a + \omega^a_{\ b}\wedge e^b\,, \tag{A.17}$$

which as can be easily recognized is the definition of the torsion 2-form T^a. Specifically

$$T^a := d^{(\omega)}e^a = de^a + \omega^a_{\ b}\wedge e^b\,, \tag{A.18}$$

so the presence of torsion reflects on the failure of the covariant exterior derivative of annihilating the basis element. Sometimes the equation $d^{(\omega)}e^a = T^a$ is referred as second Cartan structure equation. It is worth noting that the composition of two covariant exterior derivative does not vanish trivially as the analog composition of two ordinary exterior derivatives, rather we have

$$d^{(\omega)}\circ d^{(\omega)}e^a = R^a_{\ b}\wedge e^b\,, \tag{A.19}$$

which allows to extract the following expression for the curvature 2-form

$$R^{ab} = d\omega^{ab} + \omega^a_{\ c} \wedge \omega^{cb}\,, \tag{A.20}$$

known as first Cartan structure equation. It is worth remarking that

$$R^{ab} = 0 \Longrightarrow \omega^a_{\ b} = \left(\Lambda^{-1}\right)^a_{\ c} d\Lambda^c_{\ b}\,, \tag{A.21}$$

namely the connection is a pure gauge, $\Lambda^{ab} = -\Lambda^{ba}$ being a representation of the local symmetry group. Then, one can demonstrate that by assuming

$$R^{ab} = 0 \Longrightarrow de^a = 0 \quad \text{iff } T^a = 0\,, \tag{A.22}$$

which implies that $e^a = dx^a$, where x^a are functions of the original set of coordinates. Moreover, the components $e^a_\mu = \partial_\mu x^a$, so they simply represent the soldering forms in flat space between two local arbitrary accelerated reference frames at the same point of the tangent bundle.

Two useful identities can be easily derived from the above definitions, i.e.

$$d^{(\omega)} R^a_{\ b} = 0\,, \tag{A.23a}$$

$$d^{(\omega)} T^a_{\ b} = R^a_{\ b} \wedge e^b\,, \tag{A.23b}$$

respectively known as first and second Bianchi identity.

We refer now to a specific case: we assume that $n = 4$ and $s = 1$, which means that we are referring to a 4-dimensional pseudo-Riemannian manifold M^4, which is locally isomorphic to Minkowski space-time with signature $(-,+,+,+)$, so the local symmetry group is $SO(3,1)$. In this framework the Hilbert-Palatini action for General Relativity can be rewritten as

$$S(e,\omega) = \frac{1}{2}\int \star(e_a \wedge e_b) \wedge R^{ab}\,, \tag{A.24}$$

indeed, remembering definition (A.1) and formula (A.8) we can easily write:

$$S(e,\omega) = \frac{1}{2}\int \star(e_a \wedge e_b) \wedge R^{ab} = \frac{1}{2}\int \frac{1}{2} R_{cd}^{\ \ ab} \star (e_a \wedge e_b) \wedge e^c \wedge e^d = \frac{1}{2}\int d^4x \det(e) R(\omega)\,, \tag{A.25}$$

where we used $dV = \det(e) d^4x$.

An analog procedure allows to rewrite also the Dirac action in the framework of differential forms, but before it is necessary to consistently define the action of the exterior covariant derivative on a spinor field, which is a 0-form transforming under the spinor representation of the $SO(3,1)$ local group. Here we do not enter in the details about the construction of the spinor bundle, we only say that the action of the exterior covariant derivative operator acts on the spinor fields ψ and $\overline{\psi}$ according to the following definitions

$$\mathscr{D}\psi = d\psi - \frac{i}{4}\,\omega^{ab}\Sigma_{ab}\psi\,, \tag{A.26a}$$

$$\overline{\mathscr{D}\psi} = d\overline{\psi} + \frac{i}{4}\,\overline{\psi}\Sigma_{ab}\omega^{ab}\,, \tag{A.26b}$$

where

$$\Sigma^{ab} = \frac{i}{2}\left[\gamma^a, \gamma^b\right] \tag{A.27}$$

are the generators of the Lorentz group. So, the Dirac action becomes

$$S(\psi,\overline{\psi},e,\omega)=\frac{i}{2}\int\star e_a\wedge\left[\overline{\psi}\gamma^a\mathscr{D}\psi-\overline{\mathscr{D}\psi}\gamma^a\psi+\frac{i}{2}me^a\overline{\psi}\psi\right], \tag{A.28}$$

which is equivalent to the usual Dirac action as can be easily demonstrated by using formula (A.8).

As a final remark, we recall that the Hilbert–Palatini action is not affected by a change in the signature as $(-,+,+,+)\to(+,-,-,-)$, while the Dirac action gets an overall minus sign. So, it is important to remember that they have a relative minus sign in signature $(+,-,-,-)$, particularly useful when dealing with spinors.

B. Large gauge transformations in Yang-Mills gauge theories

Let the $SU(N)$ valued connection $A_\alpha=\sum_I A^I_\alpha\lambda^I$ and its associated electric field $E^\gamma=\sum_K E^\gamma_K\lambda^K$ (where $I,J,K,\cdots$ are internal indexes running on $1,2,\cdots,N^2-1$) be a couple of conjugate variables in the framework of a canonical formulation of Yang-Mills gauge theories (see § 3.3). The evolution of the system is limited to a restricted region of the phase space by the first class Gauss constraint, expressed by the following weak equation:

$$G_I:=D_\alpha E^\alpha_I=\partial_\alpha E^\alpha_I+f_{IJ}{}^K A^J_\alpha E^\alpha_K\approx 0\,. \tag{B.1}$$

According to the Dirac quantization procedure [49, 108], the state functional describing the quantum physical system must satisfy the Gauss constraint (B.1), namely we have to require that

$$\widehat{G}_I\Phi(A)=-iD_\alpha\frac{\delta}{\delta A^I_\alpha}\Phi(A)=0\,, \tag{B.2}$$

where the usual quantum representation of the operators has been assumed.

The Gauss constraint in Eq.(B.1) formalizes the request of gauge invariance of the quantum state describing the physical system, namely it is equivalent to requiring that the state functional be invariant under the small component of the gauge group $G=SU(N)$, as can be easily realized. Since the global structure of the gauge group is non-trivial, in view of quantization, it is particularly interesting to study the behavior of the state functional under the large gauge transformations. It, in fact, can produce striking effects in the non-perturbative theory, as, e.g., P and CP violations, physically motivating this extension of the theory.

In this respect, let $\widehat{\mathscr{G}}$ be the generator of the large gauge transformations, acting on the state functional $\Phi(A)$. Considering that the Hamiltonian operator, $\widehat{\mathscr{H}}$, is invariant under the full gauge group (or, more formally, it commutes with the operator $\widehat{\mathscr{G}}$), we can construct a set of eigenstates for the quantum theory by diagonalizing simultaneously $\widehat{\mathscr{H}}$ and $\widehat{\mathscr{G}}$. In other words, the following equation

$$\widehat{\mathscr{G}}\Phi_w(A)=\Phi_w(A^g)=e^{i\theta w}\Phi_w(A)\,,\quad\text{where}\quad A^g=gAg^{-1}+gdg^{-1}\,, \tag{B.3}$$

is a super-selection rule for the states of the theory, which are now labeled by the *winding number* $w=w(g)$, according to their behavior under the action of the large gauge transformation operator. The constant θ introduced in Eq. (B.3) is an angular parameter, which indicates how much the state functional "rotates" under the action of the large gauge transformations operator. Specifically,

it represents a quantization ambiguity connected with the non-trivial global structure of the gauge group.

Eq.(B.3) implies that the wave functionals have to satisfy suitable θ-dependent boundary conditions passing from one "slab" to the next in the configuration space; or, a fully gauge invariant state functional can be constructed, transferring the θ dependence in the momentum operator. In this respect, we recall that the so-called *Chern-Simons functional*,

$$\mathscr{Y}(A) = \frac{1}{8\pi^2}\int \mathrm{tr}\left(F\wedge A - \frac{1}{3}A\wedge A\wedge A\right), \tag{B.4}$$

is characterized by the following remarkable property:

$$\mathscr{Y}\left(A^g\right) = \mathscr{Y}\left(A\right) + w(g)\,. \tag{B.5}$$

This directly implies that the new state functional,

$$\Phi'(A) = e^{-i\theta\mathscr{Y}(A)}\Phi_w(A)\,, \tag{B.6}$$

will be invariant under the full gauge group, as can be easily demonstrated. In other words we have

$$\widehat{\mathscr{G}}\Phi'(A) = \Phi'(A)\,. \tag{B.7}$$

So, by using the rescaling in Eq.(B.6), we have obtained a new fully gauge invariant quantum state functional, at the price of modifying the momentum operator, namely, the θ-dependence has been transferred from the boundary conditions to the momentum operator, which becomes:

$$E'_\alpha\Phi'(A) = e^{-i\theta\mathscr{Y}(A)}E_\alpha e^{i\theta\mathscr{Y}(A)}\Phi'(A) = -i\left[\frac{\delta}{\delta A^\alpha} - \frac{i\theta}{8\pi^2}\,\varepsilon_\alpha^{\ \beta\gamma}F_{\beta\gamma}\right]\Phi'(A)\,. \tag{B.8}$$

The above modification in the conjugate momentum reflects on the Hamiltonian operator, i.e.

$$H' = \int d^3x\,\mathrm{tr}\left[\frac{1}{2}\left(E_\alpha - \frac{\theta}{8\pi^2}\,\varepsilon_\alpha^{\ \beta\gamma}F_{\beta\gamma}\right)^2 + \frac{1}{4}F_{\alpha\beta}F^{\alpha\beta}\right], \tag{B.9}$$

generating a pseudo-vectorial term which prevents the new Hamiltonian H' from being invariant under the CP discrete symmetry.

The new Hamiltonian corresponds to a topological modification of the classical action, consisting in the presence of an additional term belonging to the Pontryagin class, i.e.

$$S_{\mathrm{new}}(A) = -\frac{1}{4}\int \mathrm{tr}\star F\wedge F + \frac{\theta}{8\pi^2}\int \mathrm{tr}F\wedge F\,. \tag{B.10}$$

The θ parameter appears as a multiplicative constant in front of the modification. It is worth mentioning that the new term does not affect the classical equations of motion, as we have already noticed in § 3.2, but modifies the vacuum to vacuum amplitude in the path-integral formulation of the quantum theory. In other words, it allows to take into account possible tunneling phenomena between distinct vacua characterized by different winding numbers, violating the CP discrete symmetry.

References

[1] A. Ashtekar and J. Lewandowski, *Class. Quant. Grav.* **21** (2004) R53.

[2] C. Rovelli, *Quantum Gravity*, Cambridge University Press, Cambridge 2004.

[3] T. Thiemann, *Modern Canonical Quantum General Relativity*, Cambridge University Press, Cambridge 2006.

[4] G. Amelino-Camelia, notes of the lectures of the course *Introduction to Quantum Gravity*, held at the University of Rome "La Sapienza" during the Academic Year 2005-06.

[5] G. Amelino-Camelia, J. Ellis, N.E. Mavromatos, D.V. Nanopoulos, and S. Sarkar, *Nature* **393** (1998) 763-765.

[6] T. Piran, lectures given at the 40th winter school of theoretical physics: *Quantum Gravity and Phenomenology*, Poland 2004, [astro-ph/0407462].

[7] S.W. Hawking, *The Path-Integral Approach to Quantum Gravity*, S.W. Hawking and W. Israel eds, in General Relativity an Einstein Centenary Survey, Cambridge University Press, Cambridge 1979.

[8] S.W. Hawking and R. Penrose, *Proc. Roy. Soc. Lond.* **A314** (1970) 529-548.

[9] R.M. Wald, *General Relativity*, The University of Chicago Press, Chicago 1984.

[10] S.W. Hawking and G.F.R. Ellis, *The Large Scale Structure of Space-Time*, Cambridge University Press, Cambridge 1973.

[11] M.H. Goroff and A. Sagnotti, *Nucl. Phys.* **B266** (1986) 709.

[12] P. Van Nieuwenhuizen, *Phys. Rep.* **68** (1981) 189.

[13] S. Deser, [hep-th/9905017].

[14] C. Rovelli, *Class. Quant. Grav.* **8** (1991) 1613.

[15] J. Barbour, [arXiv:0903.3489].

[16] C. Rovelli, [arXiv:0903.3832].

[17] F. Markopoulou, [arXiv:0909.1861].

[18] J. Barbour, *Relative or Absolute motion: the discovery of dynamics*, Cambridge University Press, Cambridge 1989.

[19] J. Barbour, *The End of Time*, Oxford University Press, Oxford 2000.

[20] J. Stachel, *Einstein's search for general covariance, 1912-15*, in Einstein and the History of General Relativity vol.1 of Einstein Studies, eds. D. Howard, J. Stachel, Birkhauser, Boston 1989.

[21] K. Kuchař, *Conceptual Problems of Quantum Gravity*, eds. A. Ashtekar, J. Stachel, Birkhäuser, Boston 1991, p 141.

[22] K. Kuchař, *Phys. Rev.* **D45** (1992) 4443.

[23] K. Kuchař, in proceedings of the 4th Canadian Conference on General Relativity and Relativistic Astrophysics, World Scientific, Singapore 1992.

[24] K. Kuchař and C. Torre, *Phys. Rev.* **D43** (1991) 419.

[25] C. Rovelli, *Class. Quant. Grav.* **8** (1991) 297.

[26] C. Rovelli, *Class. Quant. Grav.* **8** (1991) 317.

[27] S. Mercuri and G. Montani, *Int. Jour. Mod. Phys.* **D13** (2004) 165, [gr-qc/0310077].

[28] S. Mercuri and G. Montani, *Mod. Phys. Lett.* **A19** (2004) 1519, [gr-qc/0312077]; in proceedings of the X Marcel Grossmann meeting, Rio de Janeiro, Brasil, July 2003, eds. S. Perez-Bergliaffa e R. Ruffini, World Scientific, Singapore 2005, [gr-qc/0401127].

[29] L. Smolin, (2005), [hep-th/0507235].

[30] T. Thiemann, [astro-ph/0607380].

[31] J.A. Wheeler, in *Relativity, groups and topology*, B.S. DeWitt and C.M. DeWitt eds., Gordon and Breach, New York 1963.

[32] S.W. Hawking, *Nucl. Phys.* **B144** (1978) 349.

[33] J. Ellis, J.S. Hagelin, D.V. Nanopoulos, and M. Srednicki, *Nucl. Phys.* **B241** (1984) 381.

[34] J. Ellis, N. Mavromatos, and D.V. Nanopoulos, *Phys. Lett.* **B293** (1992) 37.

[35] J. Ellis, J. Lopez, N. Mavromatos, and D.V. Nanopoulos, *Phys. Rev.* **D53** (1996) 3846.

[36] P. Huet and M.E. Peskin, *Nucl. Phys.* **B434** (1995) 3.

[37] R. Adler *et al.*,*Phys. Lett.* B364 (1995) 239.

[38] I. Antoniadis, C. Bachas, J. Ellis, and D.V. Nanopoulos, *Phys. Lett.* **B257** (1991) 278.

[39] V.A. Kostelecky and S. Samuel, *Phys. Rev.* **D39** (1989) 683.

[40] R. Gambini and J. Pullin, *Phys. Rev.* **D59** (1999) 124021.

[41] J. Alfaro, H.A. Morales-Tcotl, and L. Urrutia, *Phys. Rev.* **D65** (2002) 103509.

[42] J. Lukierski, A. Nowicki, and H. Ruegg, *Ann. Phys.* **243** (1995) 90.

[43] G. Amelino-Camelia, *Phys. Lett.* **B392** (1997) 283.

[44] G. 't Hooft, *Class. Quant. Grav.* **13** (1996) 1023.

[45] U. Jacob and T. Piran, [hep-ph/0607145].

[46] M.R. Martinez and T. Piran, *JCAP*, **0604** (2006) 006; *JCAP*, **0605** (2006) 017.

[47] R.P. Geroch, *Jour. Math. Phys.* **11** (1970) 437.

[48] Y. Choquet-Bruhat, J. Isemberg, and D. Pollack, [gr-qc/0611009].

[49] P.A.M. Dirac, *Lectures on Quantum Mechanics*, Belfer Graduate School of Science, Yeshiva University, New York 1964.

[50] R. Arnowitt, S. Deser, and C. Misner, *Phys. Rev.* **116** (1959) 1322.

[51] R. Arnowitt, S. Deser, and C. Misner, *Phys. Rev.* **117** (1960) 1595.

[52] R. Arnowitt, S. Deser, and C. Misner, *J. Math. Phys.* **1** (1960) 434.

[53] R. Arnowitt, S. Deser, and C. Misner, in Gravitation: An Introduction to Current Research, L. Witten ed, Wiley, New York 1962, p 227.

[54] L.D. Landau and E.M. Lifshitz, *The Classical Theory of Fields*, Pergamon, Oxford 1962.

[55] K. Kuchař, *Canonical quantization of gravity*, Relativity, Astrophysics and Cosmology, Reidel, Dordrecht 1973, p 237.

[56] S. Mercuri and G. Montani, Nuovo Cim. **120B** 2005 1137, [gr-qc/0401102].

[57] J.D. Brown and K.V. Kuchař, *Phys. Rev.* **D51** 1995 5600.

[58] Y. Choquet-Bruhat, in *Gravitation: An Introduction to Current Research*, ed. L. Witten, Wiley, New York 1962.

[59] A. Ashtekar, *Phys. Rev. Lett.* **57** (1986) 2244; *Phys. Rev.* **D36** (1987) 1587.

[60] A. Ashtekar, J.D. Romano, and R.S. Tate, *Phys. Rev.* **D40** (1989) 2572.

[61] A. Ashtekar, *Lectures on non-perturbative canonical garvity*, World Scientific, Singapore 1991.

[62] A. Ashtekar, *Mathematics and General Relativity*, American Mathematical Society, Providence, Rhode Island 1987; *Contemporary Math.* **71** (1988) 39.

[63] F. Barbero, *Phys. Rev.* **D51** (1995) 5498; *Phys. Rev.* **D51** (1995) 5507.

[64] J. Samuel, *Class. Quant. Grav.* **17** (2000) 141.

[65] N. Barros e Sá, *Int. J. Mod. Phys.* **D10** (2001) 261.

[66] A. Ashtekar, C. Rovelli, and L. Smolin, *Phys. Rev. Lett.* **69** (1992) 237.

[67] B. Dittrich and T. Thiemann, [arXiv:0708.1721].

[68] C. Rovelli, [arXiv:0708.2481].

[69] L. Modesto, *Phys. Rev.* **D70** (2004) 124009.

[70] A. Ashtekar, T. Pawlosky, and P. Singh, *Phys. Rev.* **D74** (2006) 084003.

[71] A. Ashtekar and M. Bojowald, *Class. Quant. Grav.* **23** (2006) 391.

[72] C. Rovelli, *Phys. Rev. Lett.* **97** (2006) 151301.

[73] E. Bianchi, L. Modesto, C. Rovelli, and S. Speziale, *Class. Quant. Grav.* **23** (2006) 6989.

[74] J.D. Christensen, E.R. Livine, and S. Speziale, [arXiv:0710.0617].

[75] C. Rovelli and T. Thiemann, *Phys. Rev.* **D57** (1998) 1009.

[76] G. Immirzi, *Nucl. Phys. Proc. Suppl.* **57** (1997) 65; *Class. Quant. Grav.* **14** (1997) 177.

[77] S. Alexandrov, *Phys. Rev.* **D65** (2002) 024011.

[78] S. Alexandrov and D. Vassilevich, *Phys. Rev.* **D64** (2001) 044023.

[79] S. Alexandrov and E. Livine, *Phys. Rev.* **D67** (2003) 044009.

[80] F. Cianfrani and G. Montani, *Phys. Rev. Lett.* **102** (2009) 091301.

[81] S. Mercuri, *Phys. Rev.* **D77** (2008) 024036.

[82] S. Mercuri, [arXiv:0903.2270].

[83] R. Gambini, O. Obregon, and J. Pullin, *Phys. Rev.* **D59** (1999) 047505.

[84] S. Mercuri, *Phys. Rev.* **D73** (2006) 084016.

[85] S. Mercuri, in *Proceedings of the XI Marcel Grossmann meeting*, H. Kleinert, R.T. Jantzen, and R. Ruffini eds, World Scientific, Singapore 2008, [gr-qc/0610026].

[86] H.T. Nieh and M.L. Yan, *J. Math. Phys.* **23** (1982) 373.

[87] S. Holst, *Phys. Rev.* **D53** (1996) 5966.

[88] R.K. Kaul, *Phys. Rev.* **D77** (2008) 045030.

[89] O. Chandía and J. Zanelli, *Phys. Rev.* **D55** (1997) 7580.

[90] G. Date, R.K. Kaul, and S. Sengupta, *Phys. Rev.* **D79** (2009) 041901.

[91] S. Mercuri, *Phys. Rev. Lett.* **103** 2009 081302, [arXiv:0902.2764].

[92] G. Calcagni and S. Mercuri, *Phys. Rev.* **D79** (2009) 084004, [arXiv:0902.0957].

[93] S. Mercuri and V. Taveras, [arXiv:0903.4407].

[94] V. Taveras and N. Yunes, *Phys. Rev.* **D78** (2008) 064070.

[95] A. Torres-Gomez and K. Krasnov, *Phys. Rev.* **D79** (2009) 104014.

[96] S. Mercuri and C. Tomlin, in progress.

[97] M. Lattanzi and S. Mercuri, in progress.

[98] R. Percacci and D. Perini, *Phys. Rev.* **D68** (2003) 044018.

[99] A. Ashtekar, T. Pawlowski, P. Singh, and K. Vandersloot, *Phys.Rev.* **D75** (2007) 024035.

[100] M. Bojowald, *Gen. Rel. Grav.* **40** (2008) 2659.

[101] A. Ashtekar, A. Corichi, and P. Singh,*Phys. Rev.* **D77** ,2008 024046.

[102] A. Ashtekar, [arXiv:0812.4703].

[103] A. Ashtekar, *Gen. Rel. Grav.* **41** (2009) 707.

[104] M. Bojowald, *Class. Quant. Grav.* **26** (2009) 075020.

[105] A. Ashtekar and E. Wilson-Ewing, *Phys. Rev.* **D79** (2009) 083535.

[106] L. Modesto, *Class. Quant. Grav.* **23** (2006) 5587.

[107] A. Ashtekar, V. Taveras, and M. Varadarajan, *Phys. Rev. Lett.* **100** (2008) 211302.

[108] M. Henneaux, C. Teitelboim, *Quantization of gauge systems*, Princeton University Press, Princeton 1992.

[109] J.Lewandowski, A. Okolow, H. Sahlmann, and T. Thiemann, [gr-qc/0504147].

[110] C. Fleischhack, [math-ph/0407006].

[111] T. Thiemann, [gr-qc/0305080].