The Pennsylvania State University

The Graduate School

INTERPLAY BETWEEN TOPOLOGY, GAUGE FIELDS AND GRAVITY

A Thesis in
Physics
by
Alejandro Corichi Rodriguez Gil

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 1997

Abstract

In this thesis we consider several physical systems that illustrate an interesting interplay between quantum theory, connections and knot theory. It can be divided into two parts.

In the first one, we consider the quantization of the free Maxwell field. We show that there is an important role played by knot theory, and in particular the Gauss linking number, in the quantum theory. This manifestation is twofold. The first occurs at the level of the algebra of observables given by fluxes of electric and magnetic field across surfaces. The commutator of the operators, and thus the basic uncertainty relations, are given in terms of the linking number of the loops that bound the surfaces. Next, we consider the quantization of the Maxwell field based on self-dual connections in the loop representation. We show that the measure which determines the quantum inner product can be expressed in terms of the self linking number of thickened loops. Therefore, the linking number manifests itself at two key points of the theory: the Heisenberg uncertainty principle and the inner product.

In the second part, we bring gravity into play. First we consider quantum test particles on certain stationary space-times. We demonstrate that a geometric phase exists for those space-times and focus on the example of a rotating cosmic string. The geometric phase can be explicitly computed, providing a fully relativistic gravitational Aharonov-Bohm effect. Finally, we consider 3-dimensional gravity with non-vanishing cosmological constant in the connection dynamics formulation. We restrict our attention to Lorentzian gravity with *positive* cosmological constant and Euclidean signature with *negative* cosmological constant. A complex transformation is performed in phase space that makes the constraints simple. The reduced phase space is characterized as the moduli space of flat complex connections. We construct the quantization of the theory when the initial hyper-surface is a torus. Two important issues relevant two full $3+1$ gravity are clarified, namely, the incorporation of the "reality conditions" in the quantum theory and the role played by the signature of the classical metric in the quantum theory.

Contents

1

Introduction

The development of physical theories is often driven by the quest for an unified description of the world. An example of this can be seen in the work of Newton. The universal law of gravitation and the theory of dynamics of classical bodies, as we know it nowadays, grew from the realization that the same principles should be obeyed by objects on earth as well as 'celestial' bodies. Another celebrated unification was a product of the work of Faraday and Maxwell who combined electricity and magnetism into a unique electro-magnetic field. Similar considerations now motivate the search of a unified description of all interactions, or as it is sometimes called, the 'theory of everything'.

It is not uncommon that such 'revolutions' are accompanied by the creation of new branches of mathematics, or by the application of existing techniques that were previously thought not to be relevant to the 'real world'. A classical example of the first scenario is the creation of calculus by Newton and Leibniz that came with the Newtonian dynamics. Examples of the second type are the application of the "absolute differential calculus" of Levi and Civita to the mathematical framework of the general theory of relativity and, more recently, the application of the theory of connections on fiber bundles to the description of gauge theories.

As these two examples show, *geometry* has had great impact in the formulation of physical theories. Differential Geometry, as created by Gauss and Riemann, has proved to be a cornerstone in our understanding of space-time, the backdrop on which physics 'happens'. Einstein taught us that geometry is not a mere spectator but is an active part of the play; it is *dynamical*. The idea that a manifold with a Lorentzian metric describes not only space-time but also the gravitational interaction is a very attractive one. It is natural then that great effort was put into the generalization of this idea to include electro-magnetism, the other long range interaction known at that time, into the same framework.

With the invention of quantum mechanics it became clear that general relativity, with all its beauty, could not be the 'last word'. Not only did Nature decide to behave in a non-classical fashion when probed in the microscopic domain, but She also decided that there should be additional interactions, namely, the weak and strong forces. Therefore, it is natural to seek an unified description of all interaction in terms of a metric theory, in possibly, higher dimensions. This attempts, that go under the name of Kaluza-Klein theories, have failed to provide a consistent unification.

Quantum mechanics also brought with itself an imbalance in the mathematical description of Nature: Its initial formulation was algebraic rather than geometrical. At the forefront are notions like Hilbert spaces and algebras of operators, very far removed from the language of differential manifolds, curvature, geodesic distance and the like that appear in geometry. It was only recently that quantum mechanics was re-cast as a geometrical theory [1]. The ultimate implications of this reformulation are still far from being clear. Similarly, the Holy grail of having a description for all interactions in terms of a common mathematical language was elusive for a long time. The strong and electro-weak interactions are formulated, in the standard model, as theories of connections; thus, one would like to have gravity described also as a theory of connections. It was only about ten years ago that a new set of variables for the description of gravity enabled one to put general relativity on the same footing as other interactions [2, 3]. This observation has also lead to a considerable advance in the attempt to reconcile general relativity and quantum mechanics [4].

It turns out that geometry is not the only branch of mathematics to be closely related to modern physics. In fact, a branch of topology now known as *knot theory* had its origins in physics. During the second half of last century, Lord Kelvin was trying to describe atoms as knotted vortex flux of ether [5]. He realized that knots are very stable constructs that could not be continuously deformed into each other. Therefore it was natural to regard knots as models of atoms since they were thought to be indestructible. However, there are *many more* knots than elements, so the idea did not work as an atomic theory. Knot theory departed from physics and developed on its own until very recently (for a very readable account of the history of knot theory see [6]). Knot theory came to be much more relevant for theoretical physics when Atiyah and Witten realized that certain topological quantum field theories (TQFT), and in particular a Chern-Simons field theory in 3 dimensions, would yield certain knot invariants as vacuum expectation values of "Wilson Loops" [7, 8, 9]. This construct in fact was one of the first attempts to reconcile quantum mechanics with topology, or more precisely, quantum field theory with diffeomorphism invariance. Furthermore, it involved a theory of connections. Topology and connections had already been intertwined in what is known as Donaldson theory, relating 4-dimensional topology with Yang Mills theory of self-dual connections [10]. Witten also connected the Donaldson theory with a TQFT in 4-dimensions [8]. By now, knots and physics are such a well established couple that there are even some monographs dedicated to them [11, 12]. Note that while Chern-Simons and self-dual Yang-Mills theories do play a role in physics, a more dramatic example of the application of knot theory to fundamental physics is provided by quantum gravity. Attempts to quantize general relativity, when expressed in terms of connections, have also related it to knot theory. For, natural gauge invariant functions of connections, the so-called Wilson loops, are traces of holonomies along loops. One can therefore regard quantum states as certain functionals of loops. Diffeomorphism invariance then leads us to regard certain knots invariants as physical states [13, 14, 15, 16].

Thus, twentieth-century physics has seen a growing interconnection between geometry, gauge theories and topology. The aim of this thesis is to present new results which bring out certain facets of this interplay. The thesis can be divided into two parts. The first one deals with the quantization of the free Maxwell field. As the simplest theory of connections, it has been used as testing ground for many new ideas and techniques that one hopes, will be extended to non-Abelian theories. The second part brings gravity into the picture. First, we consider test particles, both classical and quantum, on certain space-times. The metric is not dynamical but serves as a mere background. Finally, we consider the gravitational field as a dynamical entity. We focus our attention in $2+1$ gravity when regarded as a theory of connections. For a special choice of the topology of space-time we proceed to quantize the gravitational degrees of freedom.

As mentioned above, the relation between theories of connections and knot theory emerged via the Chern-Simons theory. More precisely, the expectation value of the Wilson loop of a U(1) connection along a loop γ with a measure given by $\exp(\frac{i}{\hbar}S_{\rm CS})$, with $S_{\rm CS} = \frac{k}{4\pi}\int A \wedge {\rm d}A$, the action of the Chern-Simons theory, is given by the *Gauss self-linking* number of γ [17]. In the process of the calculation one needs to introduce a *framing* in order to remove certain ambiguities in the integral. A framing is a continuous association of a transverse vector field to each point of the loop γ. Intuitively one can think of it as the limit of an infinitesimally displaced loop γ' as it approaches γ. Using Gauss' results, one can compute the linking number $\mathcal{GL}(\gamma,\gamma')$ of the two loops, and *define* the self-linking of the framed loop γ to be equal to $\mathcal{GL}(\gamma,\gamma')$. Thus, the calculation of the U(1) Chern-Simons "path integral" yields the self-linking number of the framed loop γ [9]. When non-Abelian gauge groups are considered, one gets more complicated link invariants, such as the Jones polynomial. However, until now knot theory played no role in the quantum Maxwell theory. This could in principle be expected since Chern-Simons theory is diffeomorphism invariant whereas the Maxwell theory requires a background metric. On the other hand, there is a particular aspect of the Maxwell theory that has not been widely recognized: At the kinematical level, that is, in the definition of the symplectic structure (and Poisson brackets) of the theory, one does not need a background metric. This is particularly transparent in the canonical approach where the basic variables are a connection 1-form A_a and a vector density of weight one $\tilde{E}^a$, so one can naturally integrate expression like $\int {\rm d}^3x\, \tilde{E}^a A_a$ that appear in the expression of the symplectic structure. (In terms of the covariant formulation one can also recognize this fact without difficulty (see Appendix C).) The background metric structure only appears in the definition of the Hamiltonian and therefore, in the dynamics of the theory. Since the symplectic structure of the classical theory is intimately related to the canonical commutation relations (CCR) of the quantum theory, one can anticipate a diffeomorphism invariant formulation of the commutation relations. As we shall see, this expectation is proven to be realized in a very concrete fashion.

Knot theory and in particular the Gauss linking number has a prominent role in the quantization of the electro-magnetic field. This manifestation is at two levels. The first occurs at the level of the algebra of observables, that is, in the Heisenberg uncertainty relations between canonically conjugate variables. In Chapter 2 we consider the observables of the free Maxwell field given by (suitable regularized) fluxes of electric and magnetic field across surfaces bounded by loops. This observables are intimately related with the holonomy of the magnetic potential A_a around loop γ since the holonomy is given by $\exp(\frac{i}{\rm e}\oint_\gamma A\cdot {\rm d}l) = \exp(\frac{i}{\rm e}B[\gamma])$, where $B[\gamma]$ is the magnetic flux across S_γ. Given two loops α and β, the commutator of the operators $\hat{B}[\alpha]$ and $\hat{E}[\beta]$, where $E[\beta]$ is the flux of electric field across S_β, is given by $[\hat{B}[\alpha], \hat{E}[\beta]] = i\hbar\, \mathcal{GL}(\alpha, \beta)$, with $\mathcal{GL}(\alpha, \beta)$ the Gauss linking number between α and β. For physical considerations, that is, in order to represent the basic observables as well defined operators on Fock space, we need to introduce a regularization of the loops. A natural choice is to pick a framing of the loops α and β. In the case of non intersecting loops, the result is simply given in terms of the linking number. For the intersecting case one needs to be careful in the procedure; in the extreme case, where α and β coincide, the commutator of the operators is given by the self-linking number of the framed loop α.

The second aspect of the Maxwell field, in which knot invariants play an important role, is in the expression of the inner product. This is brought out in the quantization based on certain canonical variables of the *self-dual* connection and the electric field. More precisely, one considers the holonomies of the self-dual connection (of a "thickened loop") and the real electric field as fundamental variables in the quantization. The variables one is using are 'mixed' in the sense that the connection is complex valued and the electric field is real, in complete analogy with the variables introduced by Ashtekar for gravity [3]. Contrary to the positive-frequency decomposition used in ordinary field theory, the self-dual decomposition does not require any extra background structure other that the metric. If one tries to follow the same path for quantization as one does in the positive frequency case, one gets a diffeomorphism invariant expression for the inner-product. However, it is *not* positive definite and therefore, physical incorrect (see Appendix C for details). Can we find a correct quantization that uses self-dual connections and that is invariant at the same time? In Chapter 3 we show that the answer is in the affirmative. We consider the loop representation, that is, states are functions of embedded loops. It turns out that, in order to have a well defined quantum theory, one needs to consider quantum states of regularized, thickened loops. One step in the program is to impose the *reality conditions* in the quantum theory, that is, one requires that real observables of the classical theory get promoted to self-adjoint operators in the quantum theory. The imposition of these reality-Hermiticity conditions implies that the measure that dictates the inner-product in the quantum theory is expressible entirely in terms of the self-linking number of the loops. Thus, there is a sense in which knot theory manifests itself in the inner-product of the theory, revealing a very interesting interplay between connections and loops. As these two examples demonstrate, the Gauss linking number lies at the heart of the quantum Maxwell theory, i.e., both in the uncertainty principle and the inner product of the theory.

As previously mentioned, quantum mechanics can be cast in a geometrical language. The idea is to consider the space of states $\mathcal{P}$, the *rays* in the Hilbert space $\mathcal{H}$, as fundamental. Now, observables are no longer self-adjoint operators on $\mathcal{H}$ but certain functions on $\mathcal{P}$. One still can see the Hilbert space $\mathcal{H}$ as a fiber bundle over $\mathcal{P}$ where the fiber is the complex numbers $\mathbb{C}$ minus the origin. There is a canonically defined connection on this bundle, now seen as a Hermitian U(1) bundle. The holonomy of this connection along loops γ in $\mathcal{P}$ are 'phases' (living in U(1)) that multiply the original state. Part of this phase factor can be accounted for by the Hamiltonian evolution of the system, but there might be some remnant left. This is known as the Aharonov-Anandan *geometric phase*

[18]. It is a generalization of *Berry's phase* where the (adiabatic) evolution was only considered to be in some parameter space [19].

When one considers quantum test particles on a fixed gravitational background, there are situations in which such geometric phases can be explicitly computed. In particular, in Chapter 4 we restrict our attention to certain stationary space-times that can be seen as a 'perturbation' of a fiducial static metric. An example of such space-times is given by a rotating cosmic string in 4-dimensions (or alternatively a spinning particle in 3-dimensions). In this case the space-time is locally flat outside the string, but it is not globally Minkowskian; it is a topological 'defect'. This situation has resemblance with the electro-magnetic Aharonov-Bohm in many ways, where loosely speaking, in making the analogy, one is identifying the curvature in one case with the magnetic field in the other. This can explain why this situation has often been referred as the *gravitational Aharonov-Bohm* effect. In the gravitational case one can show that there is a geometric phase when parallel transporting wave functions around the string and that this phase is present only because of the non-trivial topological space-time in which it is considered. Therefore, there is an interesting interplay between topology of the space-time and the holonomy of the universal connection over $\mathcal{P}$.

The quantization of the gravitational field, or more precisely, the reconciliation of general relativity and quantum mechanics, remains one of the grand challenges of theoretical physics. A somewhat conservative approach intends to apply the principles of quantum mechanics to the equations of general relativity. This program is sometimes called *quantum gravity*, or, more precisely, non-perturbative general relativity. In the past years there has been a considerable advance in this program so that one can say there is now a mathematically well defined framework in which fundamental questions can be addressed. There are however, certain conceptual difficulties that still need to be resolved. To address these, it is often useful to consider 'toy models' since one can then isolate the particular difficulties one is interested in, without having to worry about all the other problems that the complete theory has. One of such reduced models is given by $2+1$ dimensional gravity. In terms of the metric formulation of general relativity, the model is on one hand as difficult as the $3+1$ case and on the other, it is trivial. More precisely, when one looks at the equations of motion, or alternatively, at the constraints of the Hamiltonian theory, one finds that the theory is as difficult as its 4-dimensional counterpart in the sense that the equations are again non-linear and non-polynomial in the basic canonical variables. On the other hand, the theory is technically trivial since we *know* that in 3-dimensions the Riemann tensor is completely determined by the Ricci tensor and therefore, in the absence of matter, the space-time is flat. Therefore, we can solve the theory by producing flat space-times with the required boundary conditions.

General Relativity in 3 space-time dimensions can also be cast as a theory of connections. It was first noticed by Achúcarro and Townsend [20] and later used by Witten in [21]. They showed that the Einstein-Hilbert-Palatini action can be rewritten in terms of a non-Abelian Chern-Simons theory. The particular gauge group depends on the sign of the cosmological constant. Therefore, $2+1$ gravity is another example of a TQFT. One does not need to consider the Chern-Simons theory but one can remain in the Palatini formulations in term of connections [22]. In Chapter 5 we shall focus on this approach. In this case there is a fundamental difference in the theory depending on whether we have a non-vanishing cosmological constant or not. The case of zero cosmological constant is well understood and we shall not consider it here. We restrict our attention to particular cases of signature of the space-time and sign of the cosmological constant $((++-)$ with $\Lambda > 0$ and $((+++)$ with $\Lambda < 0)$, such that a complex change of coordinates on phase space simplifies the constraints considerably. Thus, we are in a situation similar to the $3+1$ case in which the constraints simplify when considering complex connections [3]. The (reduced) phase space of the theory for both cases mentioned above is the moduli space of flat connections. There are now some natural questions arising: Can we reconcile the fact that we have the same phase space description for space-times with different signature? Can we quantize the theory and solve the "reality conditions"? As we shall see, this questions can be answered in the affirmative, where, for simplicity, we restrict our attention to the case of spatial hyper-surface given by a closed two-torus T^2.

There are four appendices in which we give a somewhat elementary introduction to several topics of relevance to this work. Appendix A gives basic definitions and results about fiber bundles, connections and gauge fields. Appendix B demonstrates the equivalence between the intersection number of a loop and a bounded surface and the Gauss linking number of the corresponding loops. The classical and quantum theory of the Maxwell field is the subject of Appendix C. Finally, in Appendix D we construct in detail the Segal-Barmann-Hall transform for the group U(1).

2

Uncertainty Principle for the Maxwell Field

2.1 Introduction

In 1833, Gauss noticed a striking fact about electro-magnetism [23]. He considered a loop L_1 carrying a constant current I and computed the work W done in moving a magnetic monopole of strength m along a closed path L_2 in the magnetic field produced by the current:

$$W = \frac{mI}{4\pi} \oint_{L_1} \mathrm{d}s \oint_{L_2} \mathrm{d}t \; \epsilon_{abc} \, \dot{L_1}^a(s) \, \dot{L_2}^b(t) \frac{L_1^c(s) - L_2^c(t)}{|L_1(s) - L_2(t)|^3} \, . \tag{2.1}$$

He then made a deep observation which can be stated in the modern mathematical terminology as follows: although the double integral

$$\mathcal{GL}(L_1, L_2) := \frac{1}{mI} W \tag{2.2}$$

makes use of Euclidean geometry in several ways, its value is in fact a topological invariant, a measure of the linking between the loops L_1 and L_2. In particular, even if one deforms the loops, the value of the double integral does not change so long as the loops do not touch or cross each other. This is a remarkable property and Gauss expressed the belief that the quantity $\mathcal{GL}(L_1, L_2)$ may have a fundamental significance. The view was shared by others. In particular, in his celebrated treatise on electricity and magnetism, Maxwell returns to this property and further elaborates on it [24, 25].

It turns out that the double integral $\mathcal{GL}(L_1, L_2)$ does have a fundamental significance in electro-magnetism, which however, could not have been guessed before the advent of quantum field theory. To see this, consider source-free Maxwell theory (in Minkowski space-time). In the Hamiltonian treatment, the vector potential $A_a(\bar{x})$ and the electric field $E^a(\bar{x})$ (on a constant time hyper-plane) serve as the basic canonically conjugate fields with the Poisson bracket relations:

$$\{A_a(\bar{x}), \, E^b(\bar{y})\} = \delta_a^b \, \delta^3(\bar{x}, \bar{y}). \tag{2.3}$$

(Throughout this thesis, curly brackets will denote Poisson brackets.) The vector potential itself is not an observable since it fails to be gauge invariant. However, we can integrate it over a closed

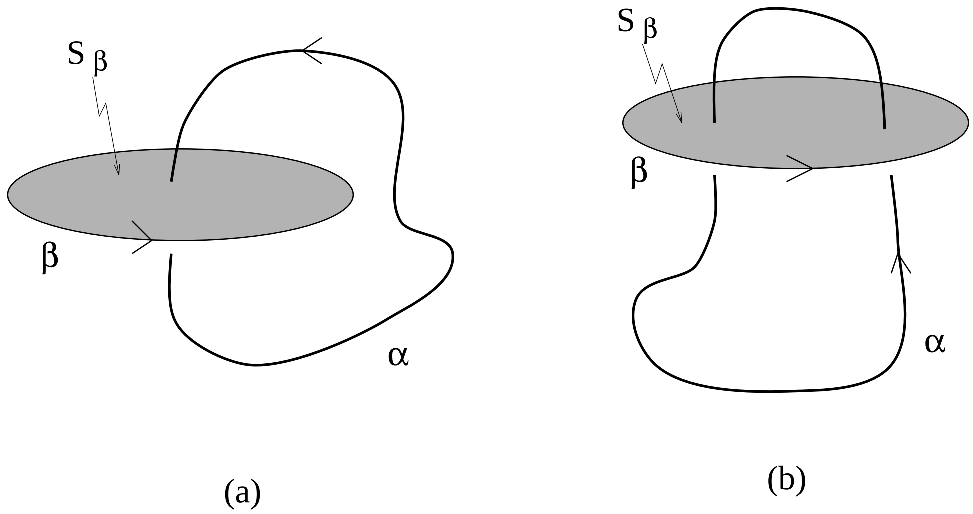

Figure 2.1: The Gauss linking number $\mathcal{GL}(\alpha, \beta)$ equals the oriented intersection number $I(\alpha, S_\beta)$. In the case (a) this number is 1 and in (b) it is 0.

loop α to obtain a gauge invariant functional:

$$B[\alpha] := \oint_\alpha A_a \mathrm{d}l^a \equiv \int_{S_\alpha} B^a \mathrm{d}^2 S_a \,, \tag{2.4}$$

where S_α is *any* 2-surface bounded by the loop α. Similarly, given any 2-surface S_β bounded by a closed loop β we can define the flux of the electric field:

$$E[\beta] := \int_{S_\beta} E^a \mathrm{d}^2 S_a \,, \tag{2.5}$$

which depends only on the loop β (and not on the specific surface S_β with boundary β) because E^a is divergence-free. The observables $B[\alpha]$ and $E[\beta]$ are (over)complete in the sense that their values at a point (B^a, E^a) of the physical phase space suffice to determine that point uniquely.

It is straightforward to compute the Poisson brackets between these observables if α and β have no point in common. The result is:

$$\begin{aligned} \{B[\alpha], E[\beta]\} &= \oint_\alpha \mathrm{d}l^a(\bar{x}) \int_{S_\beta} \mathrm{d}^2 S_a(\bar{y})\, \delta^3(\bar{x}, \bar{y}) \\ &= I(\alpha,\, S_\beta)\,, \end{aligned} \tag{2.6}$$

where $I(\alpha, S_\beta)$ denotes the oriented intersection number between the loop α and the surface S_β. But, as a geometrical picture makes it clear (see Figure 2.1), this intersection number is precisely the linking number between loops α and β. (An analytic calculation showing the equality of $I(\alpha, S_\beta)$ with $\mathcal{GL}(\alpha, \beta)$ is given in Appendix B.) Thus, we have:

$$\{B[\alpha], E[\beta]\} = \mathcal{GL}(\alpha, \beta)\,. \tag{2.7}$$

The fact that the Poisson bracket is metric independent may seem surprising at first. But note that, since the vector potential A_a is a 1-form and the electric field E^a (being canonically conjugate

Figure 2.2: The (absolute value of the) Gauss linking number between loops α and β is 1. In this case, the Heisenberg uncertainty between the magnetic flux through the surface S_α and the electric flux through S_β is $\frac{\hbar}{2}$.

to A_a) is naturally a vector density of weight one, neither the symplectic structure (2.3) nor the definitions of the observables $E[\alpha]$, $B[\beta]$ require a metric (or any other background field) for their definitions. Hence, if well-defined, the right side of (2.7) *has to be* a topological invariant of loops labeling the observables.

Let us pass to the quantum theory heuristically. One would expect that (if α and β have no point in common), the commutator between the magnetic and electric flux operators would be given by:

$$\left[\hat{B}[\alpha],\, \hat{E}[\beta]\right] = i\hbar\, \mathcal{GL}(\alpha, \beta)\,, \tag{2.8}$$

and hence the Heisenberg uncertainties should satisfy:

$$(\Delta\hat{B}[\alpha])(\Delta\hat{E}[\beta]) \geq \frac{\hbar}{2}\, |\mathcal{GL}(\alpha, \beta)|\ . \tag{2.9}$$

This implies that there is an intrinsic uncertainty in the simultaneous measurements of fluxes of electric and magnetic fields across finite surfaces if the Gauss linking numbers of the loops bounding the two surfaces is not zero (see Figure 2.2). In this sense, as suspected by Gauss and others, the linking number does have a fundamental significance in electro-magnetism.

A number of questions arise immediately. Can one make these quantum considerations precise? If the quantization is *based* on the algebra of operators generated by $\hat{B}[\alpha]$ and $\hat{E}[\beta]$, the answer is clearly in the affirmative. However, in such representations, the Maxwell Hamiltonian operator fails to be well-defined. In the standard Fock representation where the Hamiltonian *is* well-defined, operators $\hat{E}[\alpha]$ and $\hat{B}[\beta]$ fail to be well-defined! (See, e.g., [26].) Can one nonetheless give meaning to the topological uncertainty relations (2.9) in the Fock representation *in a suitable limiting sense*? Secondly, the uncertainty relation given above fails to be meaningful if the loops α and β coincide. However, it is known that the Gauss self-linking number of a loop *is* well-defined if the loop is framed [25]. Is there perhaps a framing that is naturally introduced in the process of regularization of the operators in question? If so, can one give meaning to the commutator of $\hat{B}[\alpha]$ and $\hat{E}[\alpha]$? In this Chapter, we analyze these issues. We will find that the specific questions raised here can be answered affirmatively.

This Chapter is organized as follows. In Sec. 2.2, we will consider the Fock space of photons and show that suitably 'thickened', regulated versions of the above 'flux operators' are indeed well-defined on the Fock space. This will enable us to regard $\hat{B}[\alpha]$, and $\hat{E}[\beta]$ as certain limits of well-defined operators. We will see that the commutation relations (2.8) also holds in a limiting sense. In this limit, the thickening of surfaces goes to zero. However, the loop that bounds the limiting surface carries the 'memory' of the thickening in the form of a framing. We will see in Sec. 2.3 that the limits of the commutators of the regulated flux operators are functionals of these framed loops. In particular, the framing enables one to evaluate the (limits of) commutators without ambiguities even when the loops intersect and overlap. We end the chapter with a discussion in Sec. 2.4

We will conclude this section with a few remarks.

i) One often defines *dimension-less* observables: $B'[\alpha] = \frac{c}{e} B[\alpha]$ and $E'[\beta] = \frac{1}{e} E[\beta]$, which can be exponentiated to obtain Weyl commutation relations. (The exponential of $iB'[\alpha] \equiv i\frac{c}{e} \oint_\alpha \mathrm{d}l^a A_a$, for example, is the $U(1)$ holonomy.) In terms of these primed observables, the uncertainty principle reads: $(\Delta B'[\alpha])(\Delta E'[\beta]) \geq (1/2\alpha_{\rm fine})|\mathcal{GL}(\alpha,\beta)|$, where $\alpha_{\rm fine}$ is the fine structure constant.

ii) In the non-Abelian case, one can replace $\exp iB'[\alpha]$ by the trace of the holonomy of the connection along the loop α. The analog of $E'[\beta]$ is trickier. To ensure gauge invariance, one now has to thicken the loop to a ribbon. (See, e.g., [27, 28].) The commutator is then again 'topological'. However, the physical meaning of these observables is now less transparent.

iii) The equation (2.6) was first written down by Widom and Clark in [29], but they did not analyze the existence and regularization of the operators. Finally, the results of this chapter have been reported in a paper [30].

2.2 Regulated Flux Operators

2.2.1 Preliminaries

Let us begin by recalling a few facts about the Fock representation of photons. Details can be found in Appendix C. Since we are interested in (fluxes of) electric and magnetic operators, it will be convenient to adapt the discussion to a canonical framework. A vector V in the 1-photon Hilbert space $\mathcal{H}$ is then represented by a pair (A_a, E^a) of divergence-free vector fields on a constant time hyper-plane and the Hilbert space norm is given by (see, e.g., [26], and Appendix C):

$$\langle V|V\rangle = \frac{1}{2\hbar}\int_\Sigma \mathrm{d}^3x\, \left[A_a(\Delta^{1/2}A^a) + E^a(\Delta^{-1/2}E_a)\right] . \tag{2.10}$$

Denote by $\mathcal{F}_{\rm s}(\mathcal{H})$ the symmetric Fock space based on $\mathcal{H}$. Electric and magnetic fields are represented by operator valued distributions involving the standard linear combinations of creation and annihilation operators on $\mathcal{F}_{\rm s}(\mathcal{H})$. Consider, for example, the smeared object,

$$\hat{E}[f] := \int \mathrm{d}^3x\, \hat{E}^a(x) f_a(x)\, , \tag{2.11}$$

where f_a is a test-field, i.e., vector field of compact support. Since E^a is divergence-free, we have: $E[f] = E[f + \partial g]$, for any test function g. This is a well-defined operator on $\mathcal{F}_{\rm s}(\mathcal{H})$ provided the vector $V = ({}^{\mathrm{T}}f_a, 0)$ lies in the Hilbert space $\mathcal{H}$, i.e., provided the norm

$$\langle V|V\rangle = \frac{1}{2\hbar}\int \mathrm{d}^3x\, \left[{}^{\mathrm{T}}\!f_a\,(\Delta^{1/2}\ {}^{\mathrm{T}}\!f^a)\right] \equiv \frac{1}{2\hbar}\int \mathrm{d}^3k|k|\ |{}^{\mathrm{T}}\!\tilde{f}_a|^2 \tag{2.12}$$

is finite, where ${}^{\mathrm{T}}\!f_a$ is the transverse part of f, and ${}^{\mathrm{T}}\!\tilde{f}$, its Fourier transform. (The transverse projection removes the "gauge freedom" of adding a gradient to f_a.) In that case, $\hat{E}[f]$ is expressible as a sum of the creation and annihilation operators associated with the state V.

The situation with the magnetic field operator is completely analogous. The commutator between smeared electric and magnetic fields is given by:

$$\begin{aligned}\left[\hat{B}[f]\,,\hat{E}[g]\right] &= i\hbar\int \mathrm{d}^3x\,\epsilon^{abc}(\partial_a f_b(\bar{x}))g_c(\bar{x})\\ &= \hbar\int \mathrm{d}^3k\,\epsilon^{abc}k_a(\tilde{f}_b(\bar{k}))^\star\,\tilde{g}_c(\bar{k})\,,\end{aligned} \tag{2.13}$$

where $\tilde{f}_b$ denotes the Fourier transform of f_b and $\star$ denotes complex conjugation.

2.2.2 Regularization

Let us now consider the formal expression of the electric flux operator:

$$\hat{E}[\beta] = \int_{S_\beta} \hat{E}^a \mathrm{d}^2S_a\,. \tag{2.14}$$

It can be expressed as a smeared electric field, $E[\beta] = \int \mathrm{d}^3x\, E^a(\bar{x})f_a^{(\beta)}(\bar{x})$, where, however, the test field $f_a^{(\beta)}(\bar{x})$ is a *distribution* with support on S_β:

$$f_a^{(\beta)}(\bar{x}) = \int_{S_\beta} \mathrm{d}^2S_a\ \delta^3(\bar{x},\bar{s}_\beta)\,, \tag{2.15}$$

where $\bar{s}_\beta$ denotes a point on the surface S_β. Hence, the corresponding operator $\hat{E}[\beta]$ is only formal; it fails to be well-defined on the Fock space. We must regulate it.

We will proceed in two steps (the first of which is the crucial one). Geometrically, the problem arises because $f_a^{(\beta)}$ is a distribution with two dimensional support. We can remedy this situation by an appropriate "thickening" of the surface S_β. Let us therefore replace the loop β by a strip (or ribbon) Σ_β of height ϵ (see Figure 2.3). More precisely, let us proceed as follows. Let us first equip β with a framing (i.e., let us introduce, at each point of β, a vector in a direction transverse to $\dot{\beta}^a$, the tangent to β.) Then, for each $\tau \in [0,\epsilon]$, let us denote by β_τ the loop obtained by displacing β a distance τ along the framing. (Thus $\beta_0 \equiv \beta$). This construction uses the flat Euclidean metric on the spatial hyper-plane. But the key final results will not depend on this flat metric.) Let S_τ denote a surface bounded by the loop β_τ (such that the assignment $\tau \to S_\tau$ is smooth.) The family of loops β_τ constitute the strip Σ_β and the three-dimensional region swept out by the family of surfaces S_τ constitutes a 'pill-box' P_β with boundary Σ_β.

We can now consider the flux of the electric field through the *three dimensional* pill-box region

$$E[P_\beta] = \int_0^\epsilon \mathrm{d}\tau \int_{S_\tau} \mathrm{d}^2S_a\, E^a(\bar{s}_\tau). \tag{2.16}$$

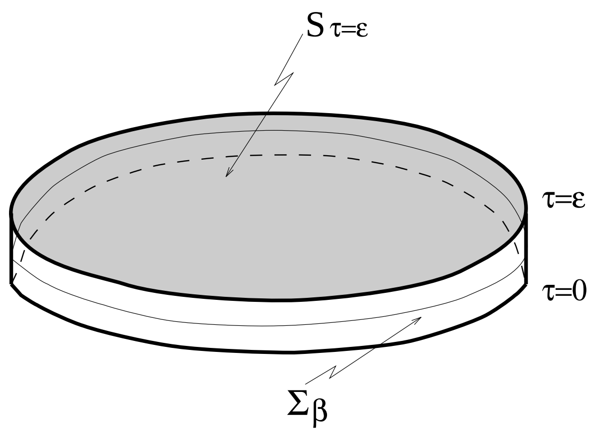

Figure 2.3: The surface S_β with boundary β is thickened to a three dimensional pill-box P_β bounded by the strip Σ_β. As ϵ tends to zero, P_β shrinks to S_β and Σ_β tends to the loop β. The 'memory' of the strip is retained by the initial framing attached to β.

(From now on, loops will be assumed to be framed. However, for simplicity of notation, we will continue to denote them just by greek letters $\alpha, \beta, ...$) This yields a smearing of the electric field with a test field $f_a^{(P_\beta)}$ with support in *three*-dimensions,

$$E[P_\beta] = \int \mathrm{d}^3x\, E^a(x) f_a^{(P_\beta)}(x), \quad \text{with} \quad f_a^{(P_\beta)} = \int_0^\epsilon \mathrm{d}\tau \int_{S_\tau} (\mathrm{d}S_\tau)_a\, \delta^3(\bar{x}, \bar{s}_\tau)\,, \tag{2.17}$$

and one can hope that the corresponding operator would be well-defined in quantum theory. This completes our first step in regularization.

The key question now is the following: Is $V = ({}^{\mathrm{T}}f_a^{(P_\beta)}, 0)$ normalizable with respect to the inner product (2.12)? It turns out that, although the 3-dimensional smearing softens the singularity of $f_a^{(\beta)}$ considerably, the norm $\langle V|V\rangle$ still has a logarithmic ultra-violent divergence (see below). This arises because the three dimensional pill-box P_β, on which $f_a^{(P_\beta)}$ is supported, has sharp boundaries. This is where the second step in the regularization procedure comes in. The problem can be handled in a number of ways. We will use the simplest one and just introduce an ultra-violet cut-off at $|\bar{k}| = \Lambda$. That is, we begin with $f_a^{(P_\beta)}$ as in (2.17), take the Fourier transform of its transverse part, multiply it by the step function which is unity if $|\bar{k}| \leq \Lambda$ and zero otherwise and consider the inverse Fourier transform $f_a^{(P_\beta,\Lambda)}$ of the resulting function. Then, $\hat{E}[f^{(P_\beta,\Lambda)}]$ is a well-defined operator on the Fock space. This operator can be regarded as the regulated version of the heuristic expression $\hat{E}[\beta]$ since, in the classical theory, we have:

$$\lim_{\epsilon\to 0} \frac{1}{\epsilon}\Big(\lim_{\Lambda\to\infty} \int \mathrm{d}^3x E^a(x) f_a^{(P_\beta,\Lambda)}(x)\Big) = E[\beta]\,. \tag{2.18}$$

The situation with the magnetic-flux operator, of course, is identical. Thus, given a framed loop α, we can introduce a strip Σ_α of height ϵ and consider a pill-box P_α it bounds. Then, $\hat{B}[f^{(P_\alpha,\Lambda)}]$ is a well-defined operator on the Fock space.

2.2.3 Quantum Flux Operators

Recall from the discussion above that the smeared operators $\hat{E}[f^{(P_\beta)}]$ are well-defined on the Fock space $\mathcal{F}_s(\mathcal{H})$ if and only if the $\mathcal{H}$-norm of $V = ({}^{\mathrm{T}}\!f_a^{(P_\beta)}, 0)$ is finite. In this sub-section we will compute this norm and show that it has a logarithmic divergence thereby establishing the necessity of an ultra-violet cut-off.

Let us consider the simplest case. In cylindrical coordinates (ρ, ϕ, z), let the strip Σ_β defining the pill-box region P_β be circles of radius $\rho = R$ and $z = \tau$; thus $\vec{\alpha}_\tau(s) = (R, 2\pi\, s, -\epsilon/2 + \tau)$. We can take the S_τ surfaces to be parallel to the $z = $ const plane. In this geometry, the smearing co-vector field $f_a(x)$ is the 'step function': $f_a(x) = \nabla_a z$ if $\rho < R$ and $z \in [-\epsilon/2, \epsilon/2]$; $f_a(x) = 0$ otherwise. The Fourier transform will have non-vanishing component only in the k_z direction:

$$\tilde{f}_{k_z}(k) = \frac{1}{(2\pi)^{3/2}} \int_0^R \int_0^{2\pi} \int_{\epsilon/2}^{\epsilon/2} \mathrm{d}z\, \mathrm{d}\phi\, \rho\, \mathrm{d}\rho\, e^{ik_\rho \rho \cos\phi}\, e^{ik_z z}\,. \tag{2.19}$$

Using the identity,

$$\mathrm{J}_0(z) = \frac{1}{\pi} \int_0^\pi e^{iz\cos\theta}\, \mathrm{d}\theta\,, \tag{2.20}$$

and the recurrence formulae of Bessel functions we arrive at

$$\tilde{f}_{k_z}(k) = \sqrt{\frac{2}{\pi}} R\, \frac{\mathrm{J}_1(k_\rho R)}{k_\rho}\, \frac{\sin(k_z \epsilon)}{k_z}\,. \tag{2.21}$$

The transverse part any field $\tilde{f}_a^T(\bar{k})$ is the projection of the that field orthogonal to the radial vector k^a. Therefore,

$$|\tilde{f}^T|^2 = |\tilde{f}_{k_z}|^2 \frac{k_\rho^2}{(k_\rho^2 + k_z^2)}\,. \tag{2.22}$$

The expression $\int_\Sigma \mathrm{d}^3k\, |k|\, |\tilde{f}_a^T(k)|^2$ now takes the form

$$\int_\Sigma \mathrm{d}^3k\, |k|\, |\tilde{f}_a(k)|^2 = 4R^2 \int_{-\infty}^\infty \int_0^\infty \mathrm{d}k_z\, \mathrm{d}k_\rho \frac{k_\rho^2}{\left[k_z^2 + k_\rho^2\right]^{1/2}} \frac{\mathrm{J}_1^2(k_\rho R)}{k_\rho}\, \frac{\sin^2(k_z\epsilon)}{k_z^2}\,, \tag{2.23}$$

where $|k| = [k_z^2 + k_\rho^2]^{1/2}$. It is now obvious that the integral diverges logarithmically since $\mathrm{J}_1(x) \sim x^{-1/2}$ when $x \to \infty$.

2.3 Removal of the regulators

Let us compute the commutator between the regularized flux operators using (2.13) and then, in the result, remove the regulators by taking appropriate limits. Removing the ultra-violet cut-off yields:

$$\begin{aligned}
\lim_{\Lambda\to\infty}\left[\hat{B}[f^{(P_\alpha,\Lambda)}],\ \hat{E}[f^{(P_\beta,\Lambda)}]\right] &= \hbar \lim_{\Lambda\to\infty}\int_\Lambda \mathrm{d}^3k\, \epsilon^{abc}k_a\left(\tilde{f}^{(P_\alpha)}(\bar{k})\right)^\star \tilde{f}^{(P_\beta)}(\bar{k}) \\
&= i\hbar\int \mathrm{d}^3x\, \epsilon^{abc}\left(\partial_a f_b^{(P_\alpha)}(\bar{x})\right) f_c^{(P_\beta)}(\bar{x}) \\
&= i\hbar\int_0^\epsilon \mathrm{d}\sigma \oint_{\alpha_\sigma} \mathrm{d}l^a \int_0^\epsilon \mathrm{d}\tau\, f_a^{(\beta_\tau)} \\
&= i\hbar\int_0^\epsilon \mathrm{d}\sigma\int_0^\epsilon \mathrm{d}\tau\{B[\alpha_\sigma],\ E[\beta_\tau]\} \qquad (2.24)
\end{aligned}$$

where, in the first step, the subscript Λ denotes that the integration is carried out over the ball $|K| < \Lambda$. The final result is not surprising: the right side is just $i\hbar$ times the well-defined Poisson bracket between the classical 'thickened flux' observables:

$$\lim_{\Lambda\to\infty}\left[\hat{B}[f^{(P_\alpha,\Lambda)}],\ \hat{E}[f^{(P_\beta,\Lambda)}\right] = i\hbar\,\{B[f^{(P_\alpha)}],\ E[f^{(P_\beta)}]\}. \qquad (2.25)$$

Thus, as far as the commutator is concerned, the ultra-violet cut-off plays no essential role.

To remove the regulator ϵ, we have to compute:

$$\begin{aligned}
\lim_{\epsilon\to 0}\frac{1}{\epsilon^2}\left(\lim_{\Lambda\to\infty}\left[\hat{B}[f^{(P_\alpha,\Lambda)}],\ \hat{E}[f^{(P_\beta,\Lambda)}]\right]\right) &= i\hbar\lim_{\epsilon\to 0}\frac{1}{\epsilon^2}\left(\int_0^\epsilon \mathrm{d}\sigma\int_0^\epsilon \mathrm{d}\tau\,\{B[\alpha_\sigma], E[\beta_\tau]\}\right) \\
&= i\hbar\lim_{\epsilon\to 0}\frac{1}{\epsilon^2}\left(\int_0^\epsilon \mathrm{d}\sigma\int_0^\epsilon \mathrm{d}\tau\,\mathcal{P}(\alpha_\sigma,\,\beta_\tau)\right), \text{ say.} \qquad (2.26)
\end{aligned}$$

This calculation is more subtle. We will divide the discussion in four cases which bring out the role of framing in handling pathologies that arise when the loops intersect and overlap.

i) The simplest case arises when the loops have no point in common. Then, for a sufficiently small ϵ, there is no intersection between any of the loops α_σ and β_τ. Hence, the Poisson bracket $\mathcal{P}(\alpha_\sigma,\,\beta_\tau)$ can be calculated exactly as in Sec. 2.1. It is independent of σ and τ and equals $\mathcal{GL}(\alpha,\beta)$. Hence,

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon^2}\int_0^\epsilon \mathrm{d}\sigma\int_0^\epsilon \mathrm{d}\tau\,\mathcal{P}(\alpha_\sigma,\beta_\tau) = \mathcal{GL}(\alpha,\beta). \qquad (2.27)$$

Thus, in this case, the limiting procedure gives a precise meaning to the calculation of Sec. 2.1: The uncertainty relation (2.9) holds in the sense that:

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon^2}\left(\lim_{\Lambda\to\infty}\left[\hat{B}[f^{(P_\alpha,\Lambda)}],\ \hat{E}[f^{(P_\beta,\Lambda)}]\right]\right) = i\hbar\,\mathcal{GL}(\alpha,\beta)\,. \qquad (2.28)$$

ii) Let us now consider the case when α and β intersect at a single point, say p. (Intersections at a finite number of points requires only a trivial extension of this case.) Now, the result depends on the thickening, or more precisely, on the framing at p initially chosen to carry out the thickening. Let $\dot{\alpha}^a$ and $\dot{\beta}^a$ denote the tangent vectors to the two loops at p. Consider the two dimensional plane they span in the tangent space of p. Suppose that the frame vectors of the loops α and β lie on opposite sides of the plane. Then (for sufficiently small ϵ) among loops α_σ and β_τ, the only ones which intersect are α_0 and β_0, the original loops. Hence, $\mathcal{P}(\alpha_\sigma, \beta_\tau)$ is well-defined in the σ, τ space except at the single point, $(\sigma = 0, \tau = 0)$, which is of measure zero. Furthermore, at all other points, $\mathcal{P}(\alpha_\sigma, \beta_\tau)$ is independent of σ and τ. Its value is precisely the Gauss linking number $\mathcal{GL}(\alpha, \beta') = \mathcal{GL}(\alpha', \beta)$, where the primed loops are obtained by moving the unprimed ones slightly along the framing vectors. Thus, we have:

$$\lim_{\epsilon\to 0} \frac{1}{\epsilon^2} \Big(\lim_{\Lambda\to\infty} \Big[\hat{B}[f^{(P_\alpha,\Lambda)}]\,, \hat{E}[f^{(P_\beta,\Lambda)}]\Big]\Big) = i\hbar\, \mathcal{GL}(\alpha, \beta') = i\hbar\, \mathcal{GL}(\alpha', \beta)\,, \tag{2.29}$$

which, in this case, is ($i\hbar$ times) the natural Gauss linking number associated with the *framed* loops.

iii) Let us now consider the case where the two loops intersect at a single point p as before but where frame vectors at p (lie on the same side of the plane spanned by the two tangents and) are parallel. Then the two strips Σ_α and Σ_β intersect in a line rather than a single point. In this case the limit is more delicate. A loop α_σ on Σ_α intersects a loop β_τ on Σ_β if and only if $\sigma = \tau$. Thus, the calculation of Sec. 2.1 for computing $\mathcal{P}(\alpha_\sigma, \beta_\tau)$ goes through for the entire region of the parameter space $(\sigma, \tau) \in [0, \epsilon] \times [0, \epsilon]$ except for the diagonal. Again, since the diagonal is a set of measure zero, we can ignore it. However, now the integrand $\mathcal{P}(\alpha_\sigma, \beta_\tau)$ is no longer a constant on the entire parameter space. As a simple geometric picture reveals, it takes one value, $\mathcal{GL}(\alpha, \beta')$, on one side of the diagonal and another value, $\mathcal{GL}(\alpha', \beta)$, on the other, where as before the primed loops are obtained by displacing the unprimed loops slightly in the direction of framing. Hence, we now have:

$$\lim_{\epsilon\to 0} \frac{1}{\epsilon^2} \Big(\lim_{\Lambda\to\infty} \Big[\hat{B}[f^{(P_\alpha,\Lambda)}]\,, \hat{E}[f^{(P_\beta,\Lambda)}]\Big]\Big) = \frac{i\hbar}{2}\Big(\mathcal{GL}(\alpha, \beta') + \mathcal{GL}(\alpha', \beta)\Big)\,. \tag{2.30}$$

The right side is ($i\hbar$ times) the average of the two possible linking numbers one can obtain by displacing the loops α and β infinitesimally using the assigned framing, i.e., the 'natural' (extension of the) Gauss linking number associated with the two given *framed* loops.

iv) Finally, let us consider the commutator between fluxes of electric and magnetic associated with the *same* framed loop α. In this case the strips Σ_α and Σ_β as well as the 'pill-boxes' P_α and P_β coincide. Now, if $\sigma \neq \tau$ the loops α_σ and β_τ have no points in common. Hence, the integrand $\mathcal{P}(\alpha_\sigma, \beta_\tau)$ is well-defined everywhere except along the diagonal. However, in this case, (outside the diagonal which we can ignore) the value of the integrand is in fact constant, namely, the Gauss linking number $\mathcal{GL}(\alpha, \alpha')$, where α' is again obtained by displacing α slightly along the framing. Thus, in this case, we have:

$$\lim_{\epsilon\to 0} \frac{1}{\epsilon^2} \Big(\lim_{\Lambda\to\infty} \Big[\hat{B}[f^{(P_\alpha,\Lambda)}]\,, \hat{E}[f^{(P_\alpha,\Lambda)}]\Big]\Big) = i\hbar\, \mathcal{GL}(\alpha, \alpha')\,. \tag{2.31}$$

The right side is precisely the *self-linking number* of the framed loop α [25]. Note in particular that if the framing is trivial (e.g., if all the frame vectors are parallel in the three dimensional Euclidean space), the right side vanishes (even before taking the limit $\epsilon \to 0$) . Thus, in this case, one can simultaneously measure the fluxes of electric and magnetic fields with *arbitrary* accuracy.

There are of course other cases one can analyze. For example, one can consider loops with isolated intersections, where, however, framing at the intersection points is not of the type considered in cases ii) and iii) above. Given the two framings, the calculation of the limit of the commutator is generally straightforward. Since the regularization is geometric and the final limiting procedure refers only to properties of ribbons obtained from framing, it is natural to interpret the result as the Gauss linking number of framed loops in those cases as well. However, we shall not go into the details of those cases.

2.4 Discussion

In this chapter we have pointed out that there is a remarkable relation between the Gauss linking number, the simplest link invariant, and the Heisenberg uncertainty between fluxes of electric and magnetic fields, the basic observables of the quantum Maxwell theory. This uncertainty is intrinsic in that it arises because of the fundamental quantum fluctuations and persists even in the vacuum state.

The precise sense in which this relation holds is rather subtle especially if the loops in question intersect or overlap. In the classical theory, given any closed loop α, we can compute the fluxes $B[\alpha]$ and $E[\alpha]$ of magnetic and electric fields through any surface S_α bounded by the loop α. To obtain the corresponding quantum observables, however, we have to 'thicken' the surfaces in question. A natural strategy is to frame the initial loop since a framing provides a canonical thickening. When this is done (and an ultra-violet cut-off is introduced) one obtains regulated flux operators which are well-defined on the Fock space. We can compute their commutators and *then* remove the regulators. The limit is just ($i\hbar$ times) the Gauss link invariant of the framed loops. Even when the loops intersect or coincide (as in cases ii), iii) and iv) considered in Sec. 3), the limit of the commutator equals the Gauss linking number of the *framed* loops.

For simplicity, we worked in Minkowski space-time. However, the entire discussion can be carried over without any difficulty to general stationary space-times (where the norm of the Killing field is bounded away from zero). In this case, one can use geodesics tangential to the framing to thicken the loops and work in the canonical Fock representation selected by the Killing field [31, 32]. In the non-stationary context, there is no canonical representation of the CCR. However, one can again construct the algebra of smeared flux operators and the basic results will hold on any Hilbert space on which this algebra can be represented. This is to be expected because the final results are topological and do not refer to the Minkowskian geometry used in the intermediate stages.

The Gauss linking number also plays a key role in the expression of the measure which dictates the inner product on the photon Hilbert space in the so called self-dual representation (where states are appropriate functionals of self-dual connections). That is the subject of the next chapter.

3

Loops and Photon Inner-product

3.1 Introduction

Source-free electrodynamics is the simplest physical theory based on connections. From a geometric point of view, a natural set of observables of this theory is given by holonomies of the connection around closed loops. It is then natural to ask if topological invariants associated with loops play a physically significant role in this description. For the observable algebra, as was shown in Chapter 2, the answer is in the affirmative: the fundamental Heisenberg uncertainty relations can be formulated in terms of the Gauss-linking number. One can ask if such a topological invariant of loops also plays a role in the description of the Hilbert space of quantum states.

In this chapter we show that the answer is again in the affirmative. Furthermore, the specific calculation we wish to present is based on an interesting interplay between self-duality, loop representation and knot theory and may well be a reflection of a deeper structure that underlies these three notions.

The basic idea is the following. The standard Fock description of photons can be reformulated in terms of loops, so that the states can be regarded as functionals of loops (rather than connections) (see, e.g., [34] and chapter 14 in [35]). There are, however, several such loop representations. In the one most directly related to the Fock-Bargmann representation [34], it is the negative frequency electric field that is diagonal and the conjugate operator represents the holonomy of positive frequency connections. Alternatively, one can work with real electric fields and connections. But then to obtain the Fock representation, the loops have to be thickened [33, 25]. In this chapter, we will work with yet another choice: our loop representation will be based on *real* electric fields but *self-dual* connections (without any reference to positive and negative frequencies). Again, to get the Fock representation, we will have to thicken our loops. Thus, quantum states are expressed as functionals of thickened loops and the basic operators are the holonomies of self-dual connections and real electric fields. The measure that dictates the inner-product in this representation has a Gaussian form where the exponent is given by the self-linking number of thickened loops.

More precisely, the situation is the following. Given a loop α and a weighting function f on $\mathbb{R}^3$, we can define a "canonical thickening" α_f of the loop. The self-linking number of a thickened loop can be computed by adding the Gauss linking numbers of the loops involved in the thickening with weights given by f. The measure that dictates the inner product is just the exponential of the self-linking number of α_f. Now, as is well-known, the Fock inner product depends on the Minkowski metric. It is quite interesting that one can put all the information about the space-time metric in the construction that associates quantum states with functionals of loops and then express the inner-product itself in terms of the Gauss linking number which is a topological invariant. A further striking fact is that this "coding" of the inner-product information in a topological invariant works only in the loop representation based on *self-dual* (or anti-self-dual) connections.

The plan of this chapter is as follows. In Section 3.2, we will collect some mathematical preliminaries. These are used in Section 3.3 to construct the loop representation based on self-dual connections. The main result then follows in Section 3.4. Since our primary goal is to bring out the interplay between self-duality of connections, loop representations and the linking number, we will keep the functional analytic details to a minimum. However, it should be rather straightforward to see how one can complete our discussion to obtain a rigorous treatment. Throughout this chapter, we use units where $c = 1$, but write $\hbar$ and e explicitly. Finally, let us mention that the results of this chapter have been reported in [36].

3.2 Mathematical Preliminaries

This section is divided in to two parts. In the first, we recall the phase space formulation of the Maxwell field using self-dual variables and in the second we introduce the notion of "form factors" associated with loops and thickened loops.

3.2.1 Self-Dual Variables for the Maxwell Field

Let us begin with a brief summary of the standard phase space formulation of Maxwell fields (see Appendix C for details). Denote by Σ a space-like three-plane in Minkowski space, and by q_{ab}, the induced positive definite (flat) metric thereon. The configuration variable for the Maxwell field is generally taken to be the connection one-form $A_a(\bar{x})$ (the vector potential for the magnetic field) on Σ. Its canonically conjugate momentum is the electric field $E^a(\bar{x})$ on Σ. ($E^a(\bar{x})$ naturally arises as a vector density. However, since we have an underlying metric, q_{ab}, which can be used to add or remove density weights, we will ignore density weights in this chapter). The fundamental Poisson bracket is:

$$\{A_a(\bar{x}), E^b(\bar{y})\} = \delta_a^b \delta^3(\bar{x}, \bar{y})\,. \tag{3.1}$$

The system has one first class constraint, $\partial_a E^a(\bar{x}) = 0$. One can therefore pass to the reduced phase space by fixing transverse gauge. The true degrees of freedom are then contained in the pair $(A_a^{\rm T}(\bar{x}), E_{\rm T}^a(\bar{x}))$ of transverse (i.e. divergence-free) vector fields on Σ. Denote by Γ the phase space spanned by these fields. On Γ, the only non-vanishing fundamental Poisson bracket is:

$$\{A_a^{\rm T}(\bar{x}), E_{\rm T}^b(\bar{y})\} = \delta_a^b \delta^3(\bar{x}, \bar{y}) - \Delta^{-1}\partial^b\partial_a \delta^3(\bar{x}, \bar{y})\,, \tag{3.2}$$

where Δ is the Laplacian operator compatible with the flat metric q_{ab}. It is convenient to write $A^{\rm T}_a(\bar{x})$ and $E^a_{\rm T}(\bar{x})$ in terms of their Fourier decomposition. Then, the true degrees of freedom are contained in the new dynamical variables $q_j(\bar{k})$, $p_j(\bar{k})$ with $j = 1, 2$:

$$A^{\rm T}_a(\bar{x}) = \frac{1}{(2\pi)^{3/2}} \int {\rm d}^3k \, {\rm e}^{ik\cdot x} (q_1(\bar{k}) m_a(\bar{k}) + q_2(\bar{k}) \bar{m}_a(\bar{k})) \tag{3.3}$$

$$E^a_{\rm T}(\bar{x}) = \frac{1}{(2\pi)^{3/2}} \int {\rm d}^3k \, {\rm e}^{ik\cdot x} (p_1(\bar{k}) m^a(\bar{k}) + p_2(\bar{k}) \bar{m}^a(\bar{k})) \,, \tag{3.4}$$

where m_a and $\bar{m}_a$ are transverse (complex) vectors satisfying: $m_a k^a = 0$, and $m_a \bar{m}^a = 1$. The Poisson brackets (3.2) for the transverse components are,

$$\{q_i(-\bar{k}), p_j(\bar{k}')\} = -\delta_{ij}\delta^3(\bar{k}, \bar{k}') \,, \tag{3.5}$$

while the fact that $A^{\rm T}_A(\bar{x})$ and $E^a_{\rm T}(\bar{x})$ are real translates to the "reality conditions":

$$\bar{q}_i(\bar{k}) = q_i(-\bar{k}) \quad \text{and} \quad \bar{p}_i(\bar{k}) = p_i(-\bar{k}) \,. \tag{3.6}$$

In order to construct the self dual connection, we will use $d^{\rm T}_a(\bar{x})$, the transverse vector potential of the electric field ($E^a_{\rm T}(\bar{x}) = \epsilon^{abc}\partial_b d^{\rm T}_c(\bar{x})$)

$$d^{\rm T}_a(\bar{x}) = \frac{1}{(2\pi)^{3/2}} \int \frac{{\rm d}^3k}{|k|} \, {\rm e}^{ik\cdot x} (p_1(\bar{k}) m_a(\bar{k}) - p_2(\bar{k}) \bar{m}_a(\bar{k})) \,. \tag{3.7}$$

Let us define the *self dual connection* as

$$^{\uparrow}\!A^{\rm T}_a(\bar{x}) := -iA^{\rm T}_a(\bar{x}) + d^{\rm T}_a(\bar{x}) \,. \tag{3.8}$$

We want to use the pair ($^{\uparrow}\!A^{\rm T}_a(\bar{x})$, $E^a_{\rm T}(\bar{x})$) as the basic variables. In terms of the (q_j, p_j) coordinates, the self dual connection takes the form,

$$\begin{aligned} ^{\uparrow}\!A^{\rm T}_a(\bar{x}) &= -\frac{1}{(2\pi)^{3/2}} \int \frac{{\rm d}^3k}{|k|} \, {\rm e}^{ik\cdot x} [(-p_1(\bar{k}) + i|k| q_1(\bar{k})) m_a(\bar{k}) + (p_2(\bar{k}) + i|k| q_2(\bar{k})) \bar{m}_a(\bar{k})] \\ &= -\frac{1}{(2\pi)^{3/2}} \int \frac{{\rm d}^3k}{|k|} \, {\rm e}^{ik\cdot x} [z_1(\bar{k}) m_a(\bar{k}) + z_2(\bar{k}) \bar{m}_a(\bar{k})] \,, \end{aligned} \tag{3.9}$$

where

$$z_1(\bar{k}) := -p_1(\bar{k}) + i|k| q_1(\bar{k}) \quad \text{and} \quad z_2(\bar{k}) := p_2(\bar{k}) + i|k| q_2(\bar{k}) \,. \tag{3.10}$$

The basic Poisson brackets for the pairs $(z_i(\bar{k}), p_j(\bar{k}))$, – the Fourier components of the self-dual connection and the real electric field– are:

$$\{p_i(\bar{k}), z_j(-\bar{k}')\} = i|k|\delta_{ij}\delta^3(\bar{k}, \bar{k}') \,, \tag{3.11}$$

and the "reality conditions" (3.6) now become:

$$z_1(-\bar{k}) + \bar{z}_1(\bar{k}) = -2p_1(-\bar{k}) \quad \text{and} \quad z_2(-\bar{k}) + \bar{z}_2(\bar{k}) = 2p_2(-\bar{k}) \,. \tag{3.12}$$

Finally, for later convenience, let us examine the self-dual magnetic field $B^a := \epsilon^{abc}\partial_b{}^{\uparrow}A_c$. Its Fourier components of the magnetic field are given by

$$B_1(\bar{k}) = -z_1(\bar{k})\,, \quad \text{and} \quad B_2(\bar{k}) = z_2(\bar{k})\,. \tag{3.13}$$

The reality conditions for the magnetic field read

$$B_j(\bar{k}) + \overline{B_j(-\bar{k})} = 2p_j(\bar{k})\,. \tag{3.14}$$

These conditions will play an important role in the selection of the inner-product in the loop representation.

3.2.2 Loops

Let us begin with some definitions. By a *loop* we shall mean a continuous and piecewise smooth mapping γ from S^1 to Σ, where $s \in [0, 2\pi]$. Two loops γ and β will be said to be *holonomically equivalent* if, for every smooth connection A_a, we have $\oint_\gamma A_a \mathrm{d}s^a = \oint_\beta A_a \mathrm{d}s^a$. It turns out that two holonomically equivalent loops, γ and β, can differ from each other only through: i) reparametrization, $\gamma(s) = \beta(s')$ for some (orientation-preserving) reparametrization $s \to s'$ of the curve $\beta(s)$; ii) retracing identity, $\gamma = l \cdot \beta \cdot l^{-1}$, where l is a line segment and $\cdot$ indicates composition of segments [37]. Each equivalence class will be referred to as a *holonomic loop*. Since loops will primarily enter our discussions through holonomies, it is these equivalence classes –rather than individual loops– that will be directly relevant to our discussion. To keep the notation simple, we will use the same symbols –say γ– to denote both an individual loop and the holonomic loop it defines; the context should suffice to resolve the resulting ambiguity.

An analytic characterization of holonomic loops can be given through certain distributional vector densities, called *form factors*. Given a loop α, its *form factor*, $F^a[\alpha, \bar{x})$, is defined via:

$$\int \mathrm{d}^3x F^a[\alpha, \bar{x}) w_a(\bar{x}) = \oint_\alpha w_a \mathrm{d}s^a\,. \tag{3.15}$$

Thus, $F^a[\alpha, \bar{x})$ may be more directly expressed as

$$F^a[\alpha, \bar{x}) = \oint_\alpha \mathrm{d}s\, \dot{\alpha}^a(s)\,, \delta^3(\bar{x}, \alpha(s)) \tag{3.16}$$

where $\alpha(s)$ is a point on the loop α at parameter value s and $\dot{\alpha}^a(s)$ the tangent vector to α at $\alpha(s)$. Note that the form factor $F^a[\alpha, \bar{x})$ is automatically divergence free,

$$\partial_a F^a[\alpha, \bar{x}) = 0\,, \tag{3.17}$$

because $\oint_\alpha \partial_a f\, \mathrm{d}s^a = 0$. It is often convenient to perform a Fourier transform to obtain the momentum space representation of $F^a[\alpha, \bar{x})$. We have:

$$\begin{aligned} F^a[\alpha, \bar{k}) &:= \frac{1}{(2\pi)^{3/2}} \int \mathrm{d}^3x\, \mathrm{e}^{-i\bar{k}\cdot\bar{x}} F^a[\alpha, \bar{x}) \\ &= \frac{1}{(2\pi)^{3/2}} \oint_\alpha \mathrm{d}s\, \dot{\alpha}^a(s) \mathrm{e}^{-i\bar{k}\cdot\alpha(s)}\,. \end{aligned} \tag{3.18}$$

Let us note a few properties of these form factors. First, two loops α and β will have the same form factors if and only if they are holonomic. Thus, $F^a[\alpha, \bar{x})$ can be used to characterize holonomic loop α. Next, since $F^a[\alpha, \bar{x})$ is divergence-free its Fourier transform is transverse ($k_a F^a[\alpha, \bar{k}) = 0$). We can write the two independent components as:

$$\begin{aligned} F_1[\alpha, \bar{k}) \equiv F^+[\alpha, \bar{k}) &= \frac{1}{(2\pi)^{3/2}} \oint_\alpha \mathrm{d}s\, \dot{\alpha}^a(s) \bar{m}_a(\bar{k}) \mathrm{e}^{-i\bar{k}\cdot\alpha(s)} \\ F_2[\alpha, \bar{k}) \equiv F^-[\alpha, \bar{k}) &= \frac{1}{(2\pi)^{3/2}} \oint_\alpha \mathrm{d}s\, \dot{\alpha}^a(s) m_a(\bar{k}) \mathrm{e}^{-i\bar{k}\cdot\alpha(s)}\,, \end{aligned} \tag{3.19}$$

so that

$$F^a[\alpha, \bar{k}) = F^+[\alpha, \bar{k}) m^a(\bar{k}) + F^-[\alpha, \bar{k}) \bar{m}^a(\bar{k})\,. \tag{3.20}$$

(We have introduced the $\pm$ notation because in the quantum theory, F_1 will capture positive helicity and F_2 the negative.) This transversality of form factors will play an important role in the loop-representation because it captures in a natural way the gauge invariance of the theory, i.e. the transversality of the photon. The next property follows from the fact that $F^a[\alpha, \bar{x})$ is real. Consequently, its Fourier transform $F^a[\alpha, \bar{k})$ satisfies the "reality condition"

$$\overline{F}_j[\alpha, \bar{k}) = -F_j[\alpha, -\bar{k})\,. \tag{3.21}$$

Finally, given two holonomic loops α and β, we define a new holonomic loop, $\alpha\#\beta$ as follows: $\alpha\#\beta = l \cdot \alpha \cdot l^{-1} \cdot \beta$ where l is any line segment joining a point on α to a point on β. (Because of its geometric picture $\alpha\#\beta$ is sometimes called the "eye-glass loop".) Using the definition of the form factors, we now have

$$F_j[(\alpha\#\beta), \bar{k}) = F_j[\alpha, \bar{k}) + F_j[\beta, \bar{k}) \tag{3.22}$$

In order to construct the quantum theory in the loop representation, we will need to thicken the loops appropriately. We will conclude this section by indicating how this can be done. Fix an averaging function $f_r(\bar{x})$ such that $\int_{R^3} d^3x f_r(\bar{x}) = 1$, and goes to a delta function when $r \to 0$. A convenient choice is [38]:

$$f_r(\bar{y}) = \frac{1}{(2\pi r)^{3/2}} \exp\left(-\frac{y^2}{2r}\right) \tag{3.23}$$

To make the discussion concrete, we will make this choice and thus characterize the thickening completely by a real parameter r. (However, the generalization to arbitrary smearing functions is obvious.) Now, given a loop α we take the loop $\alpha + y$ obtained by rigidly shifting the loop by the vector y^a,

$$(\alpha + y)^a(s) = \alpha^a(s) + y^a \tag{3.24}$$

Next, we can average over y using the weight $f_r(\bar{y})$ and define a "smeared form factor" via:

$$\begin{aligned} F_r^a[\alpha, \bar{x}) &:= \int \mathrm{d}^3y f_r(\bar{y}) F^a[\alpha + \bar{y}, \bar{x}) \\ &= \int \mathrm{d}^3y f_r(\bar{y}) \oint_\alpha ds\, \dot{\alpha}^a(s) \delta^3(\bar{x} - \alpha(s) - \bar{y}) \\ &= \oint_\alpha \mathrm{d}s\, \dot{\alpha}^a(s) f_r(\bar{x} - \alpha(s)) \,. \end{aligned} \tag{3.25}$$

Its Fourier transform $F_r^a[\alpha, \bar{k})$ satisfies

$$F_r^a[\alpha, \bar{k}) = \exp\left(-\frac{r^2 k^2}{2}\right) F^a[\alpha, \bar{k}) \,. \tag{3.26}$$

We will see that these $F_r^a[\alpha, \bar{k})$ can be used as "generalized coordinates" for loops. More precisely, once the weight functions $f_r(\bar{y})$ are chosen, we can associate with any loop a transverse, *smooth* vector field,

$$\alpha \longrightarrow F^a(\bar{k}) := F_r^a[\alpha, \bar{k}) \,. \tag{3.27}$$

As is well-known, photon states can be expressed as suitable functionals, $\Phi[F]$, of smooth vector fields $F^a(\bar{k})$ (in the representation in which the electric field is diagonal). These can be pulled-back to the loop space to yield functionals $\Psi_r(\alpha) = \Phi[F]|_{F=F_r[\alpha,\bar{k})}$. Thus, the entire Fock space of photon states can be expressed in terms of suitable functionals of loops. This fact will be exploited in the next section.

3.3 Quantum Theory

This section is divided in to three parts. In the first, we recall a general quantization program (for details, see [39, 28]), in the second we construct a $\star$-algebra of operators based on loop variables and in the third we construct the loop representation.

3.3.1 Quantization Program

Consider a classical system with phase space Γ. To construct the quantum theory, we can proceed in the following steps.

i) Choose a subspace $\mathcal{S}$ of the space of complex valued functions on Γ which is closed under the Poisson bracket operation and large enough so that any well behaved function on Γ can be expressed as (possible the limit of) a sum of products of elements of $\mathcal{S}$. Elements of $\mathcal{S}$ are called *elementary classical variables* and are to have unambiguous quantum analogs.

ii) Associate with each f in $\mathcal{S}$ an *elementary quantum operator* $\hat{f}$ and consider the free associative algebra generated by these abstract operators. Impose on this algebra the (generalized) canonical commutation relations

$$[\hat{f}, \hat{g}] = i\hbar \widehat{\{f, g\}} \,, \tag{3.28}$$

for all f and g in $\mathcal{S}$. In addition, if the set $\mathcal{S}$ is over-complete, impose on the algebra also 'anti-commutation relations', namely the relations that capture the algebraic relations that exist between elements of $\mathcal{S}$. For instance if f, g and $h = fg$ are all in $\mathcal{S}$, then $\hat{f} \cdot \hat{g} + \hat{g} \cdot \hat{f} = 2\hat{h}$. Denote the resulting associative algebra by $\mathcal{A}$.

iii) Introduce an involution, $*$, on $\mathcal{A}$ by setting

$$(\hat{f})^* = \hat{\bar{f}} \tag{3.29}$$

for all elementary variables f (the bar denotes complex conjugation as before) and requiring that $*$ satisfies the defining properties of an involution: $(\hat{A} + \lambda\hat{B})^* = \hat{A}^* + \bar{\lambda}\hat{B}^*$; $(\hat{A}\hat{B})^* = \hat{B}^*\hat{A}^*$ and $(\hat{A}^*)^* = \hat{A}$, for all $\hat{A}$, $\hat{B}$ in $\mathcal{A}$ and complex numbers λ. Denote the resulting $*$-algebra by $\mathcal{A}^*$.

iv) Choose a linear representation of $\mathcal{A}$ on a complex vector space V. (The $*$-relations are ignored at this step).

v) Introduce on V an inner product $\langle\, ,\, \rangle$ so that the "quantum reality conditions" are satisfied

$$\langle \Psi, \hat{A}\Phi \rangle = \langle \hat{A}^*\Psi, \Phi \rangle \tag{3.30}$$

for all Φ, Ψ in V and $\hat{A}$ in $\mathcal{A}^*$. Thus, it is the $*$-relations that are to select the inner product.

The program requires two external inputs: the choice of $\mathcal{S}$ in step (i) and the choice of the carrier space V of the representation in step (iv). If the choices are viable, i.e. if the program can be completed at all, the resulting inner product is unique on each irreducible sector of the representation of $\mathcal{A}$ on V [40]. In the framework of this program, the textbook treatments of field theories correspond to choosing for elements of $\mathcal{A}$ the smeared field operators, and, for V, the Fock space or, alternatively, suitable functionals of fields (see Appendix C). In the loop quantization, on the other hand, one changes this strategy. both $\mathcal{S}$ and V are now constructed from holonomic loops.

3.3.2 Algebra Based on Loop Variables

Let us now implement this program for the Maxwell field using loop variables. Let us define the *smeared holonomy* of self-dual connections as:

$$\begin{aligned} h_r[\alpha] \;&:=\; \exp\left(\tfrac{1}{e}\int \mathrm{d}^3x \oint_\alpha \mathrm{d}s\, \dot{\alpha}^a(s)\, {}^{+}\!A_a(\bar{x}) f_r(\bar{x}-\alpha)\right) \\ &=\; \exp\left(\tfrac{1}{e}\int \mathrm{d}^3x\, F_r^a[\alpha,\bar{x}]\, {}^{+}\!A_a(\bar{x})\right)\,, \end{aligned} \tag{3.31}$$

or equivalently,

$$h_r[\alpha] = \exp\left[-\tfrac{1}{e}\int \frac{\mathrm{d}^3k}{|k|}(z_1(\bar{k})\bar{F}_1(\bar{k}) + z_2(\bar{k})\bar{F}_2(\bar{k}))\exp\left(-\frac{r^2k^2}{2}\right)\right]\,. \tag{3.32}$$

Being a function of the self dual connection it can be regarded as a "configuration variable". As a momentum variable we will take the (real) electric field $E^a(\bar{x})$, or its Fourier transform $E^a(\bar{k})$. (Strictly speaking we should take the smeared observable $E[f] = \int_\Sigma E^a f_a\, d^3x$, but this smearing will not be relevant for our results.) Hence, $h_r[\alpha]$ and $E^a(\bar{k})$ provide us with a (over-) complete coordinatization of the phase space. The space $\mathcal{S}$ of elementary classical variables required in the first step of the quantization program shall be the vector space generated by the $h_r[\alpha]$ and $E^a(\bar{k})$. It is closed under Poisson-bracket operation because

$$\{h_r[\alpha], E^a(\bar{k})\} = \tfrac{i}{e} F_r^a[\alpha, \bar{k}) h_r[\alpha] \,. \tag{3.33}$$

The next step in the quantization program is the construction of the algebra $\mathcal{A}$ of quantum operators. Let us associate with each $h_r[\alpha]$ in $\mathcal{S}$ an operator $\hat{h}_r[\alpha]$ and with each $E^a(\bar{k})$ an operator $\hat{E}^a(\bar{k})$ and consider the associative algebra generated by finite sums of products of these elementary quantum operators. On this algebra impose the commutation relations:

$$\left[\hat{h}_r[\alpha], \hat{h}_r[\beta]\right] = 0 \qquad ; \qquad \left[\hat{E}^a(\bar{k}), \hat{E}^b(k')\right] = 0$$
$$\left[\hat{h}_r[\alpha], \hat{E}^a(\bar{k})\right] = -\tfrac{\hbar}{e} F_r^a[\alpha, \bar{k}) \hat{h}_r[\alpha] \tag{3.34}$$

Furthermore, we must incorporate in this quantum algebra the fact that $h_r[\alpha]$ is over-complete. i.e. there are algebraic relations among them; $h_r[\alpha]h_r[\beta] = h_r[\alpha\#\beta]$. This is achieved by imposing on the algebra the relations $\hat{h}_r[\alpha]\hat{h}_r[\beta] = \hat{h}_r[\alpha\#\beta]$ for all holonomic loops α and β. The result is the algebra $\mathcal{A}$ of quantum operators.

3.3.3 Loop Representation

The next step in the program is to choose a vector space V and a representation of the quantum operators. The procedure involved is generally exploratory. Thus, one does not specify all the required regularity conditions right in the beginning; the precise definition of spaces considered becomes clear only at the end of the construction. This will also be the case in our construction.

We wish to choose for V a vector space of suitable functionals of loops. As noted at the end of Section 3.2.2, in the standard electric field representation, one can choose states as suitably regular functionals $\Phi[F]$ of smooth, vector fields $F^a(\bar{k})$ which are transverse, i.e., satisfy $F^a(\bar{k})k_a = 0$. Now, in section 3.2.2 (for each choice of a smearing function f_r), we set up a mapping $\alpha \mapsto F_r^a[\alpha, \bar{k})$ from loops to smooth transverse vector fields in the momentum space. We can just pull back the functionals $\Phi(F)$ via this map to obtain certain functionals $\Psi(\alpha)$ on the loop space:

$$\Psi(\alpha) = \Phi[F]|_{F=F_r[\alpha,\bar{k})} \,. \tag{3.35}$$

(Using the regularity conditions on Φ that come from the standard electric-field representation, it is not difficult to check that the map has no kernel, i.e., $\Psi(\alpha) = 0$ for all α if and only if $\Phi[F] = 0$.) Since the transverse vector fields $F^a(\bar{k})$ have only two components $F^{\pm}(\bar{k})$, from now on we will regard Φ as functionals of the two fields $F^{\pm}$.

Thus, for the representation space V, we will use the functionals Ψ on the loop space of the form (3.35). Using the procedure that was successful in the loop representation adapted to the positive-frequency connections [34], the action of the basic operators $\hat{h}_r[\alpha]$ and $\hat{E}^a(\bar{k})$ will be taken to be:

$$\begin{aligned} \hat{h}_r[\alpha]\Psi(\gamma) &= \Psi(\gamma\cdot\alpha) \\ \hat{E}^a(\bar{k})\Psi(\gamma) &= \tfrac{\hbar}{e}F_r^a[\gamma,\bar{k})\Psi(\gamma) \end{aligned} \tag{3.36}$$

As is usual in the loop representation, the electric field is diagonal in the representation. The only non-vanishing commutator between the basic operators is

$$\left[\hat{h}_r[\alpha], \hat{E}^a(\bar{k})\right] = -\tfrac{\hbar}{e}F_r^a[\alpha,\bar{k})\hat{h}_r[\alpha] \tag{3.37}$$

Finally, for later convenience, we note the action of the magnetic field operators $\hat{B}^{\pm}$ on these states.

$$\hat{B}^{\pm}(\bar{k})\Psi(\alpha) = \pm e|k|\left[\frac{\delta}{\delta F^{\pm}(-\bar{k})}\Phi[F^{\pm}(\bar{k})]\right]\Bigg|_{F^{\pm}(\bar{k})=F_r^{\pm}[\alpha,\bar{k})}, \tag{3.38}$$

which is nothing but the "loop derivative" evaluated at $F_r^{\pm}$ (see, e.g. [25]).

Our next task is to find an inner-product on V so that the "quantum reality conditions" (3.30) are satisfied. Let us begin with an inner product of the form

$$\langle\Psi|\Psi'\rangle := \int\prod_{k,\pm}\mathrm{d}F^{\pm}(\bar{k})\,\mathrm{e}^{-T[F^{\pm}(\bar{k})]}\,\overline{\Phi[F^{\pm}]}\,\Phi'[F^{\pm}] \tag{3.39}$$

and determine the measure by imposing the reality conditions. The property (3.21) of form factors implies that $T[F]$ should be real. It also implies that the reality condition on the electric field is automatically satisfied. The other condition one should impose, namely the quantum version of (3.14) is

$$\begin{aligned} \langle\Psi|(\hat{B}^{\pm}(\bar{k}))^{\dagger}\chi\rangle &= \overline{\langle\chi|\hat{B}^{\pm}(\bar{k})\Psi\rangle} \\ &= \langle\Psi| - \hat{B}^{\pm}(-\bar{k}) + 2\hat{p}^{\pm}(-\bar{k})|\chi\rangle\,. \end{aligned} \tag{3.40}$$

Using the form of the operators (3.36) and (3.38) for $\hat{p}^{\pm}(\bar{k})$ and $\hat{B}^{\pm}(\bar{k})$, we conclude that the reality condition (3.14) is satisfied if and only if

$$\frac{\delta T}{\delta F^{\pm}(\bar{k})} = \mp\frac{2\hbar}{e^2|k|}\overline{F^{\pm}(\bar{k})}\,. \tag{3.41}$$

The solution to this equation is:

$$T[F] = -\frac{2\hbar}{e^2}\int\frac{\mathrm{d}^3k}{|k|}\left[|F_r^{+}(\bar{k})|^2 - |F_r^{-}(\bar{k})|^2\right] \tag{3.42}$$

Hence, the explicit form of the inner product (3.39) is given by:

$$\langle\Psi|\Psi'\rangle = \int\prod_{k,\pm}\mathrm{d}F^{\pm}(\bar{k})\,\mathrm{e}^{\left[\frac{2\hbar}{e^2}\int\frac{\mathrm{d}^3k}{|k|}\left(|F^{+}(\bar{k})|^2-|F^{-}(\bar{k})|^2\right)\right]}\,\overline{\Phi[F^{\pm}]}\,\Phi'[F^{\pm}]\,. \tag{3.43}$$

Notice that the basic form of (3.43) is the same as that of the inner-product for a free-field in the configuration (i.e., Schrödinger) representation[1]. There are, however, two important differences.

[1]Although $F^{\pm}(\bar{k})$ are complex-valued, they arise as Fourier components of a *real* field $F^a(\bar{x})$ and hence satisfy the reality conditions $\overline{F^{\pm}(\bar{k})} = -F(-\bar{k})$. The configuration space underlying our loop representation is thus real and states $\Phi(F^{\pm})$ are *arbitrary* complex-valued functions of $F^{\pm}$ (i.e., not subject o any "holomorphicity" condition.)

First, our states are functionals of loops rather than of a configuration field variable (such as the connection or the electric field). Second, for the positive component, the Gaussian is exponentially growing rather than damping. Hence, while we can take the states to be polynomials in F^- as in the Schrödinger representation, we have to assume that they are exponentially damped in their dependence on F^+. Thus, for example, we can take elements of V to be the functionals $\psi(\alpha)$ on the loop space of the form:

$$\Psi(\alpha) = P[F_r^{\pm}[\alpha, \bar{k})] \exp\left[-\frac{2\hbar}{e^2}\int \frac{\mathrm{d}^3k}{|k|}\left(|F_r^+(\bar{k})|^2\right)\right], \tag{3.44}$$

where $P[F_r^{\pm}[\alpha, \bar{k})]$ is a polynomial in $F^{\pm}$. As usual, the Cauchy completion will enlarge this space; the Hilbert space of all states will contain more general functionals. In this description, F^+ captures positive helicity while F^- captures the negative helicity of the photon. Thus, as one might have expected from our use of only the self-dual part of the connection, the description is asymmetric in the two helicities.

To summarize, the elementary operators are $\hat{h}_r(\alpha)$ and $\hat{E}^a$. The space of quantum states is given by functionals $\Psi(\alpha)$ of holonomic loops which are normalizable with respect to the inner-product (3.43) and the action of the elementary operators is given by (3.36). For every $r > 0$, this loop representation is naturally isomorphic to the Fock representation[2] (where the isomorphism, however, depends on the value of r.) The fact that we are using a loop representation adapted to self-dual connections is reflected in the measure that dictates the inner product (3.43). In the loop representation adapted to positive-frequency fields [34], for example, the measure has the same form but the squares of both $|F^{\pm}|$ appear with negative signs in the exponent.

3.4 Measure and the Gauss Linking Number

Recall that our quantum states are functionals of thickened loops α_r, or equivalently, of their form factors $F_r^{\pm}(\alpha, k)$; it is for technical convenience that in the intermediate stages of calculations that we extended them to functionals on the vector space of all fields $F^{\pm}(\bar{k})$. Therefore, it is instructive to examine the measure that dictates the inner-product also directly in terms of the thickened loops. This is easy to achieve: we can just pull-back the "Gaussian" $\exp(-T)$ that dictates the inner-product to the space of thickened loops. The result is trivially given by:

$$\exp\left(-T[F_r[\alpha, \bar{k})]\right) = \exp\left[\frac{2\hbar}{e^2}\int \frac{\mathrm{d}^3k}{|k|}\left(|F_r^+[\alpha, \bar{k})|^2 - |F_r^-[\alpha, \bar{k})|^2\right)\right]. \tag{3.45}$$

We will now show that this loop functional can be expressed in terms of the Gauss linking number.

Let us begin by recalling the definition of the linking number. Given non-intersecting loops α and β the Gauss linking number $\mathcal{GL}(\alpha, \beta)$ between them can be expressed in terms of their form factors as (see Appendix B):

$$\mathcal{GL}(\alpha, \beta) = \int \mathrm{d}^3x\, F^a[\alpha, \bar{x}) w_a[\beta, \bar{x}) \tag{3.46}$$

[2] If we let r go to zero, the smearing function $f_r(\bar{x})$ tends to the δ-distribution and the thickened loop α_r reduces to the loop α. However, now the exponent T in the measure diverges and the loop representation ceases to exist.

where $F^a[\alpha,\bar{x})$ is the form factor for α and $w_a[\beta,\bar{x})$ is a potential for the form factor of β: $\epsilon^{abc}\partial_b w_c[\beta,\bar{x}) = F^a[\beta,\bar{x})$. The integral is independent of the specific choice of the potential $\omega_a[\beta,\bar{x})$ because $F^a[\alpha,\bar{x})$ is divergence free. Note that neither the definition of the form factor $F^a(\alpha,x)$ nor that of the potential $\omega_a(\beta,x)$ requires any background fields on the underlying oriented 3-manifold $\mathbb{R}^3$; in particular, there is no reference to the 3-metric. (Since F^a is a vector density, the ϵ^{abc} in the definition of $\omega_a[\beta,\bar{x})$ is the Levi-Civita density which is naturally available on any oriented 3-manifold.) This is to be expected since the Gauss linking number is a topological invariant.

Nonetheless, as shown in Appendix B, one can use the flat metric q_{ab} on R^3 to express the linking number in more familiar terms. First, we have the well-known form used by Gauss himself [23]:

$$\mathcal{GL}(\alpha,\beta) := \frac{1}{4\pi}\int \mathrm{d}s \int \mathrm{d}t\, \epsilon_{abc}\dot{\alpha}^a(s)\,\dot{\beta}^b(t)\frac{\alpha^c(s)-\beta^c(t)}{|\alpha(s)-\beta(t)|^3}\,. \tag{3.47}$$

For our purposes, a more convenient form is the one involving the Fourier transforms of the form factors. The Fourier transform of the potential has the form:

$$\begin{aligned} F^a[\beta,\bar{k}) &= iw_c[\beta,\bar{k})k_b\epsilon^{abc} \\ &= w_c[\beta,\bar{k})|k|(m^a\bar{m}^c - \bar{m}^a m^c) \\ &= |k|(m^a w^+[\beta,\bar{k}) - \bar{m}^a w^-[\beta,\bar{k})) \end{aligned} \tag{3.48}$$

whence,

$$F^+[\beta,\bar{k}) = |k|w^+[\beta,\bar{k}), \quad \text{and} \quad F^-[\beta,\bar{k}) = -|k|w^-[\beta,\bar{k}). \tag{3.49}$$

Therefore, the Gauss linking number takes the form

$$\begin{aligned} \mathcal{GL}(\alpha,\beta) &= \int \mathrm{d}^3k \overline{F^a[\alpha,\bar{k})}w_a[\beta,\bar{k}) \\ &= \int \mathrm{d}^3k(\overline{F^+[\alpha,\bar{k})m^a} + \overline{F^-[\alpha,\bar{k})\bar{m}^a})w_a[\beta,\bar{k}) \\ &= \int \frac{\mathrm{d}^3k}{|k|}(\overline{F^+[\alpha,\bar{k})}F^+[\beta,\bar{k}) - \overline{F^-[\alpha,\bar{k})}F^-[\beta,\bar{k}))\,. \end{aligned} \tag{3.50}$$

Finally, we will need the notion of Gauss number of the thickened loops α_r and β_r. This is just the total linking number of loops in α_r with those in β_r:

$$\begin{aligned} \mathcal{GL}(\alpha_r,\beta_r) &:= \int \mathrm{d}^3y \int \mathrm{d}^3z\, f_r(\bar{y})f_r(\bar{z})\,\mathcal{GL}(\alpha^a+y^a,\beta^a+z^a) \\ &= \int \frac{\mathrm{d}^3k}{|k|}(\overline{F_r^+[\alpha,\bar{k})}F_r^+[\beta,\bar{k}) - \overline{F_r^-[\alpha,\bar{k})}F_r^-[\beta,\bar{k})) \end{aligned} \tag{3.51}$$

Hence the self-linking number of a thickened loop α_r is given by:

$$\mathcal{GL}(\alpha_r,\alpha_r) = \int \frac{\mathrm{d}^3k}{|k|}\left[\overline{F_r^+[\alpha,\bar{k})}F_r^+[\alpha,\bar{k}) - \overline{F_r^-[\alpha,\bar{k})}F_r^-[\alpha,\bar{k})\right] \tag{3.52}$$

whence the "Gaussian" (3.45) on the space of thickened loops which dictates the inner-product can be expressed as:

$$\exp\left(-T[F_r[\alpha,\bar{k})]\right) = \exp\left[\frac{2\hbar}{e^2}\mathcal{GL}(\alpha_r,\alpha_r)\right] . \tag{3.53}$$

This is the result we were seeking. (Note, incidentally, that the coefficient of the linking number is 2 over the fine structure constant.)

We conclude with a remark. Had we used positive frequency connections [34], for example, the loop functional (3.45) would have been replaced by

$$\int \frac{\mathrm{d}^3k}{|k|}(\overline{F_r^+[\alpha,\bar{k})}F_r^+[\alpha,\bar{k}) + \overline{F_r^-[\alpha,\bar{k})}F_r^-[\alpha,\bar{k}))$$

which has no obvious interpretation in terms of the topology of loops. Similarly, if we had worked in the self-dual connection representation, the measure would have been dictated by a "Gaussian" on the space of *connections* (see chapter 11.5, especially Eq 42b in [35], and [41]) and would therefore also have had no relation to topological invariants of loops. We need both self-duality of the connection *and* the loop representation to relate the photon inner product with the Gauss linking number.

4

Geometric Phases in Gravitational Systems

4.1 Introduction

In this second part of the thesis, we shall consider gravitational systems and show interesting relations between quantum theory, connections and topology. In this chapter we study test particles on a fixed background, and in particular, a generalization to gravity of the familiar Aharonov-Bohm (AB) effect. Recall that the AB effect is a result of the topological non-triviality of the configuration space accessible to the electron. It is also of great importance since is was the first evidence of the physical relevance of the magnetic vector potential $\vec{A}$, the electro-magnetic connection, in quantum mechanics. Previously, the only quantity of physical relevance was taken to be the magnetic field $\vec{B}$, the curvature of the connection.

The term "gravitational Aharonov-Bohm effect" (gAB) has been used to describe a variety of systems with diverse properties [42, 43, 44]. Usually it implies a behavior different from the one in Minkowski space-time even when the motion is restricted to regions in which the space-time is (locally) flat. At the classical level, the parallel transport of vectors and spinors round a cosmic string, for example, is nontrivial when the mass M of the string is different from zero. When the angular momentum J is non vanishing, the time required for null (light) rays to travel around the cosmic string depends on the direction of motion [45]. For a quantum test (Klein Gordon) particle the effect is manifested in different ways depending on the situation. When bound states are studied [43] it is seen that the energy spectrum is modified depending on the mass and angular momentum of the string. In the scattering of particles, the scattering amplitude is non trivial and the phase shifts are modified also depending on M and J [46].

It has also been noted that in linearized gravity, i.e. with $g_{ab} = \eta_{ab} + h_{ab}$ where h_{ab} is small, the deparametrized square-root Hamiltonian of a slowly moving particle takes the form: $H = (p_i - A_i)^2 + \frac{1}{2}mh_{00}$ where $A_i = \frac{1}{2}h_{0i}$. A_i transforms like a U(1) connection under infinitesimal diffeomorphism $x^i \to x^i + \zeta^i$ and if $\partial_{[i}A_{j]} = 0$, the Schrödinger equation looks like the one in the AB case, h_{0i} playing the role of the electro-magnetic potential [47, 48]. Using this analogy, Anandan obtained a geometric phase for a particle satisfying the Schrödinger equation in the weak field region around a cosmic string by interfering beams passing on opposite sides of the string. The wave function thus constructed is, however, not single-valued [49]

We will show that this last construction can be extended to general metrics (without the weak field assumption) and for the fully relativistic Klein-Gordon equation (so that the velocity is not necessarily small) using only single-valued solutions. The g_{0i} part of the metric still behaves like a potential. Both the Hamiltonian constraint and the Klein Gordon equation have a very appealing form in terms of a fiducial static metric on which the potential A_i "lives". This will allow us to identify g_{0i} as a vector potential and, in the case of locally flat space-times (which include cosmic strings), discuss a fully relativistic gAB effect.

The similarities between charged quantum particles in the presence of a long thin solenoid and scalar quantum particles in the space-time generated by a cosmic string, are at two levels: 1) The Schrödinger equation for the charged particle and the Klein-Gordon equation have a similar form; 2) There exists a "vector potential" in the linearized case that behaves like a electro-magnetic U(1) connection. Nevertheless, it is not a priori clear to what extent g_{0i} can be taken as a genuine connection and what its role is in the gAB effect, as compared to the role of A_i in the electro-magnetic AB effect. The aim of this chapter is to clarify this point.

The structure of this chapter is as follows. In Sec. 4.2, we consider the Klein-Gordon equation for a specific class of stationary space-times that can be decomposed into a fiducial static background metric plus a "potential A_i term". We show that the solutions can be constructed from solutions on the "background" metric using a prescription introduced by Dirac. For the specific case of the space-time of a spinning cosmic string further properties hint at a gAB effect. In Sec. 4.3 we review the electro-magnetic Aharonov-Bohm effect. Following Berry's suggestion, it can be interpreted as an Aharonov-Anandan geometric phase i.e. the holonomy of a natural connection in the Hilbert space of states. Section 4.4.1 recalls that, for a stationary space-time, the space of solutions of the Klein-Gordon equation can be given the structure of a complex Hermitian Hilbert space. In Sec. 4.4.2, we combine the results of Secs. 4.2, 4.3 and 4.4.1 to conclude that a geometric phase does exist for the spinning cosmic string. The phase is explicitly constructed. It provides a fully relativistic gravitational Aharonov-Bohm effect.

Throughout this chapter units where $G = \hbar = c = 1$ are assumed. Many of the results of this chapter are contained in the paper [50]. Geometric phases in more general space-times have been recently considered in [51].

4.2 Test Particle on Stationary Space-times

Stationary solutions of vacuum Einstein equations can describe the space-time geometry outside rotating bodies. As is well known such metrics are characterized by having a time-like Killing vector field t^a. Let us consider the following stationary line-element [52]:

$$\mathrm{d}s^2 = -V^2(\mathrm{d}t - A_i \mathrm{d}x^i)^2 + h_{ij}\mathrm{d}x^i \mathrm{d}x^j \,, \tag{4.1}$$

where V, A_i, h_{ij} are functions on a Cauchy surface Σ coordinatized by x^i, $i = 1, 2, 3$. A_i has vanishing curl $\partial_{[i}A_{j]} = 0$. The time coordinate t is the affine parameter along the time-like Killing vector field t^a. The freedom left in the choice of t is $t \longrightarrow t + \Lambda(x^i)$ whereas x^i can be transformed among themselves completely arbitrarily without changing the general form of the metric. Under this coordinate transformation, the object A_i transforms as $A_i \longrightarrow A_i + \nabla_i \Lambda$. Under diffeomorphisms of the 3 manifold Σ, A_i transforms as a covector. We can therefore interpret A_i as a U(1) connection under gauge transformations, living on the fiducial static space-time with metric $\overset{\circ}{g}_{ab} = -V^2 \nabla_a t \nabla_b t + h_{ij} \nabla_a x^i \nabla_b x^j$.

Let us consider a classical test particle of mass m. The Hamiltonian constraint for the particle is given by,

$$\begin{aligned} H &= g^{ab} p_a p_b + m^2 \\ &= \left(h^{ij} A_i A_j - \frac{1}{V^2} \right) p_0^2 + 2 h^{ij} A_i p_j p_0 + h^{ij} p_i p_j + m^2 = 0 \end{aligned} \tag{4.2}$$

We can solve for p_0 in order to get a true "Hamiltonian",

$$p_0 = \frac{\vec{A} \cdot \vec{p} + \left[h^{ijkl} A_i A_j p_k p_l - (\frac{\vec{p}^2}{V^2} - m^2 \vec{A}) + \frac{m^2}{V^2} \right]^{1/2}}{\frac{1}{V^2} - \vec{A}^2}, \tag{4.3}$$

where $h^{ijkl} = h^{ik} h^{jl} - h^{ij} h^{kl}$. It is straightforward to show that in the weak-field and low-velocity limit, this Hamiltonian reduces to the one considered by DeWitt in [47].

Assume now that we have massive scalar quantum test particles whose behavior is determined by the Klein-Gordon equation,

$$(\Box - m^2) \Psi = 0 . \tag{4.4}$$

For the line element (4.1) it takes the form,

$$\left[\overset{\circ}{\Box} + h^{ij} A_i A_j {\partial_0}^2 + 2\, h^{ij} A_j \partial_0 \partial_i + \frac{1}{V\sqrt{h}} \partial_i \left(h^{ij} A_j V \sqrt{h} \partial_0 \right) - m^2 \right] \Psi = 0 , \tag{4.5}$$

where $\overset{\circ}{\Box}$ is the operator corresponding to $\overset{\circ}{g}_{ab}$, h^{ij} is the inverse of h_{ij} and $h = \det(h_{ij})$. Since the background is time independent, we can look for solutions of the form

$$\Psi(t, \bar{x}) = e^{-iEt} \Phi(\bar{x}) \tag{4.6}$$

Substituting (4.6) into (4.5) we obtain the equation for $\Phi(\bar{x})$,

$$\left[\left(\frac{E^2}{V^2} - h^{ij} A_i A_j \right) - 2iE h^{ij} A_i \partial_j - \frac{iE}{V\sqrt{h}} \partial_i (A^i V \sqrt{h}) - iE A^i \partial_i - m^2 \right] \Phi(\bar{x}) = 0 \tag{4.7}$$

It is a curious fact that its solution can be written in terms of the solution $\Phi_0(x)$ to the Klein-Gordon equation $(\overset{\circ}{\Box} - m^2)\Psi_0 = 0$ on the background metric $\overset{\circ}{g}_{ab}$, where we have assumed $\Psi_0 = e^{-iEt}\Phi_0$,

$$\Phi(x) = e^{iE \int_{\mathbf{x}_0}^{\mathbf{x}} A_i dx^i} \Phi_0(x) \tag{4.8}$$

This is the Dirac phase factor method. Although $\Psi(t,x)$ is a solution, it is multi-valued and therefore should be treated with care as we will see in Section 4.4.2. For a non-contractible closed path C the multi-valued feature of solution (4.6) is transparent.

One remark is in order. The requirement that $\partial_{[i}A_{j]} = 0$ means that, locally, the connection A_i is pure gauge, i.e., that locally the metric (4.1) can be put in a static form. More precisely, it implies that the twist of the Killing field t^a is zero and therefore the nonintegrabillity of the hyper-surfaces has a topological origin[1], namely, whenever the first cohomology group of the hyper-surface is non-vanishing: $H^1(\Sigma) \neq 0$.

4.2.1 The Cosmic String

A particular case of (4.1) is the space-time generated by a spinning cosmic string, on which we will focus on this part. In cylindrical coordinates (t,ρ,θ,z) the line element is given by (4.1) with

$$\begin{aligned} V^2 &= 1 \\ A_i &= -4J\nabla_i\theta \\ h_{ij} &= \nabla_i\rho\nabla_j\rho + (\alpha\rho)^2\nabla_i\theta\nabla_j\theta + \nabla_i z\nabla_j z\,, \end{aligned} \tag{4.9}$$

where J is the angular momentum of the string and $\alpha = 1 - 4\mu$ with μ the linear mass density. This space-time is flat (outside the string) and can be locally written as the Minkowski line element. Consider the local change of coordinates given by: $\tilde{t} := t + 4J\theta$ and $\tilde{\theta} := \alpha\theta$, the line element takes the form,

$$\mathrm{d}s^2 = -\mathrm{d}\tilde{t}^2 + \mathrm{d}\rho^2 + \rho^2\mathrm{d}\tilde{\theta}^2 + \mathrm{d}z^2 \tag{4.10}$$

Now, the range of the 'angle' $\tilde{\theta}$ is $[0, 2\pi\alpha)$. Therefore as a consequence of non vanishing μ the spatial slices (t=constant) are planes with a wedge of angle $\beta = 8\pi\mu$ removed and edges identified, i.e., a cone. We see that the new 'time' coordinate $\tilde{t}$ can not be globally defined. In geometrical terms, for a non vanishing J, the hyper-surfaces which are locally orthogonal to the time-like Killing vector field, can not be globally integrated.

If one takes the Killing vector field $(\frac{\partial}{\partial z})^a$ and reduces along its orbits, one gets a $2+1$ space-time with a massive spinning particle at the origin. The mass of the particle is μ and its angular momentum is J. For a detailed description of such space-time see [53]. Therefore there exist two equivalent descriptions, one four dimensional and the other three dimensional.

There are some surprising results related to this space-time already in classical physics that are worth pointing out. Let us examine the effects of parallel transport of spinors and vectors along non-contractible loops. Although this space-time is locally flat everywhere (except at the $\rho = 0$ line), holonomies are non trivial. The components of vectors undergo a change which is sensitive to the angular defect β as well as the angular momentum J. On the other hand the components

[1] If we regard $F_{ab} = \nabla_a t_b$ as an "Electro-magnetic two form", its magnetic field (proportional to the twist) is zero; formally the situation is the same as in the electro-magnetic AB case.

of spinors detect just the angular defect β; therefore, spinors do *not* see the rotation of the cosmic string.

Finally, let us consider quantum particles. Using (4.6) and (4.8), it follows that the solution to the Klein-Gordon equation in this background is given by

$$\Psi(t,\rho,\theta,z) = e^{-iEt}\,\Phi(\rho,\theta,z)\,, \tag{4.11}$$

where

$$\Phi(\rho,\theta,z) = e^{-i4JE\int_{\theta_0}^{\theta} d\theta}\,\Phi_0(\rho,\theta,z)\,, \tag{4.12}$$

$\Phi_0(x)$ being a solution on the space-time of a non rotating string. Concretely, the equation (4.7) takes the form,

$$\left[\frac{\partial^2}{\partial\rho^2} + \frac{1}{\rho}\frac{\partial}{\partial\rho} + \frac{1}{(\alpha\rho)^2}\frac{\partial^2}{\partial\theta^2} + \frac{\partial^2}{\partial z^2} - m^2 + E_n\right]\Phi_{0\,n} = 0\,, \tag{4.13}$$

where $\Phi_n = \exp(-i4JE_n\theta)\Phi_{0\,n}$. The solutions are given by,

$$\Phi_{0\,n} = N_n\,e^{ipz}\,e^{il\theta}\,\mathrm{J}_\nu(k\rho) \tag{4.14}$$

with $k = \sqrt{E_n^2 - (p^2+m^2)}$, $\nu = \frac{l}{\alpha}$ and N_n is a normalization constant.

In the next section we will see how the phase-factor gives rise to some interesting results.

4.3 Geometric Phase and the Aharonov-Bohm Effect

The property that the solution of a "complicated" equation can be found from solutions to a "simpler" equation using the Dirac phase factor is a well known feature of the electro-magnetic Aharonov-Bohm (AB) wave function. In that case the solution for the Schrödinger equation for a charged particle in the exterior of a solenoid containing a magnetic flux ϕ can be constructed from the solution with no flux. One common exposition of the AB effect is to consider two beams of particles passing on opposite sides of the solenoid. The wave function for each beam is constructed from the solution without magnetic flux using the Dirac phase factor along each path. The wave functions of the two beams will be related by $\exp(i\oint A_i \mathrm{d}x^i)$, which is precisely the holonomy of the electro-magnetic U(1) connection around the flux source. One then argues that the AB effect is due to the non triviality of the holonomy even though the curvature of the connection $F_{\mu\nu}$ vanishes. There is, however, a problem with this reasoning: the wave function is not single valued. This implies that, in spite of manifest axisymmetry, the angular momentum of the particle is not conserved as one switches on the flux on the solenoid adiabatically and is therefore not acceptable [54]. Indeed, in the analysis of a scattering situation, studied by Aharonov and Bohm in detail, it is important that the wave function is required to be single valued and the effect is observed as a nontrivial scattering amplitude [55]. Thus, strictly speaking one cannot use the naive argument.

Nevertheless, one can still use the Dirac phase factor and relate the AB effect to the holonomy of a connection following Berry's suggestion that the AB effect is a particular case of Berry's phase [19]. However, the experimental setup has changed in this setting: one considers a particle contained in a box (not containing the solenoid), and the interference occurs between two particles one of which was transported around the flux line. Berry's phase is a particular case of the Aharonov-Anandan geometric phase which we will now introduce [18].

Consider a Hilbert space $\mathcal{H}$ and the set $\mathcal{H}_o$ of normalized vectors, i.e.,

$$\mathcal{H}_o = \{|\Psi\rangle \in \mathcal{H} \,/ \langle\Psi|\Psi\rangle = 1\} \,. \tag{4.15}$$

Define an equivalence relation $\sim$ in $\mathcal{H}_o$ by: $|\Psi\rangle \sim |\Phi\rangle$ iff $|\Psi\rangle = e^{if}|\Phi\rangle$ where f is real. The ray space $\mathcal{P}$ is defined as the quotient space,

$$\mathcal{P} = \mathcal{H}_o / \sim \tag{4.16}$$

and represents the space of all physically distinct states. $\mathcal{H}_o$ has the structure of a U(1) principal bundle over $\mathcal{P}$. Suppose that the system undergoes a cyclic evolution in $\mathcal{P}$ generated by a Hamiltonian H, $i\frac{d}{dt}|\Psi\rangle = H(t)|\Psi\rangle$. Denote by C the corresponding curve, $C:\ [0,\tau] \longrightarrow \mathcal{P}$. The initial and final states of a curve $\hat{C}$ in $\mathcal{H}_o$ that is projected to C in the ray space $\mathcal{P}$ will differ by a phase,

$$|\Psi(\tau)\rangle = e^{i\beta}|\Psi(0)\rangle. \tag{4.17}$$

This phase β can be decomposed into two parts, $\beta = \gamma + \delta$, where δ is called the *dynamical phase* which depends on the Hamiltonian,

$$\delta = -\int_0^\tau \langle\Psi|H(t)\Psi\rangle dt \,, \tag{4.18}$$

while the remainder, γ, is the *geometrical phase* which depends only on the curve C in $\mathcal{P}$. There is a natural, "universal", U(1) connection defined in $\mathcal{H}_o$ such that the holonomy around the path yields the geometric phase γ. Given C in $\mathcal{P}$, let us define the horizontal lift C' in $\mathcal{H}_o$, by requiring that $|\Phi(t)\rangle$ satisfies,

$$\langle\Phi|\frac{\mathrm{d}}{\mathrm{d}t}\Phi\rangle = 0 \,. \tag{4.19}$$

Then the state $|\Phi(t)\rangle$ acquires a phase when parallel-transported:

$$|\Phi(\tau)\rangle = e^{i\gamma}\,|\Phi(0)\rangle \,. \tag{4.20}$$

We can define a *connection* $A_s := \langle\Psi(s)|\frac{\mathrm{d}}{\mathrm{d}s}\Psi(s)\rangle$. Under "gauge transformation" $|\Psi\rangle \longrightarrow e^{i\lambda(s)}|\Psi\rangle$, A_s transforms as $A_s \longrightarrow A_s + \dot{\lambda}$. Taking $|\Psi'\rangle$ such that $|\Psi'(\tau)\rangle = |\Psi'(0)\rangle$ the geometric phase is

$$\gamma = i\oint\langle\Psi'|\dot{\Psi}'\rangle \,. \tag{4.21}$$

Berry's phase is recovered when the evolution in $\mathcal{H}_o$ is adiabatic and the time dependence of the Hamiltonian is completely contained on external parameters $R_i(t)$, such that $R_i(\tau) = R_i(0)$. The system evolves according to the Schrödinger equation

$$i\,\frac{\mathrm{d}}{\mathrm{d}t}|\Psi(t)\rangle = H(R_i(t))\,|\Psi(t)\rangle\,. \tag{4.22}$$

The adiabaticity is imposed by assuming that an eigenstate $|E(R_i(0))\rangle$ will evolve with H and will be in the state $|E(R_i(t))\rangle$ at time t.

The geometric phase is then

$$\gamma(C) = i\,\oint_C \langle E(R_i)|\nabla_R E(R_i)\rangle\,\mathrm{d}R\,. \tag{4.23}$$

In this framework, the Aharonov-Bohm effect can be described as follows [19]. Consider a particle of charge q inside a 'perfectly reflecting' box such that the wave function $\Psi_n(r)$ is non zero only in the interior of the box. Call R_i the vector from the origin of the coordinate system (where the solenoid is located) to the center of the box. When there is no magnetic potential the wave functions have the form $\Psi_n(r-R)$ with energies E_n independent of R_i. With non-zero flux the wave functions $\langle\, r|\, n(R_i)\rangle$ are obtained by the Dirac phase factor inside the box,

$$\langle\, r|\, n(R_i)\rangle = \exp\left[iq\int_R^r \mathrm{d}r'\cdot A(r')\right]\,\Psi_n(r-R)\,. \tag{4.24}$$

Let the box be transported around a circuit C threaded by the flux line. The geometric phase (4.23) can be then computed using the fact that

$$\begin{aligned}\langle n(R_i)|\,\nabla_R n(R_i)\rangle &= \int \mathrm{d}^3r\,\Psi_n^*(r-R)\,[-iqA(R)\Psi_n(r-R) + \nabla_R\Psi_n(r-R)] \\ &= -i\,q\,A(R)\,,\end{aligned} \tag{4.25}$$

which implies

$$\gamma(C) = q\,\oint A(R)\cdot\mathrm{d}R = q\,\phi\,. \tag{4.26}$$

Then, the effect manifests itself as an interference between the particle in the transported box and the one in a box which was not transported round the circuit.

We conclude with two remarks. First note that the solution is single valued and therefore well defined everywhere. Secondly, the factor $e^{i\gamma}$ coincides with the holonomy of the electro-magnetic connection, but as we have seen the connection in question is the U(1) connection on the unit sphere $\mathcal{H}_o$. This fact will allow us to relate the electro-magnetic and gravitational AB effects.

4.4 Geometric Phase in Gravity

4.4.1 Space of Solutions

For a geometric phase as described in the last section to exist, the only conditions that must be satisfied are the existence of a Hilbert Space with a Hermitian inner product and a unitary operator generating time evolution[2]. The existence of such structures in the space-times considered in Sec. 4.2 is the subject of this section.

In ordinary non-relativistic quantum mechanics, the time evolution of the system is governed by a differential equation which is first order in time, namely the Schrödinger equation. The Hamiltonian operator is the infinitesimal generator of time translations and it must be self-adjoint with respect to the Hermitian inner product in order to generate unitary evolution. The Klein-Gordon equation is on the other hand a *second order in time* equation and it is not clear a-priori that there exist for general space-times a well defined "square root" that can be identified with the Hamiltonian.

The Hilbert space structure is however well defined in the case when the underlying space-time is stationary [31] and a Hamiltonian operator can be constructed as the generator of time translations. For completeness, let us recall this construction where the one-particle Hilbert space $\mathcal{H}$ is obtained from the vector space V of real solutions to the Klein-Gordon equation. $\mathcal{H}$ must have the structure of a complex Hilbert space in order to represent quantum states of a particle. For that we need to introduce on V a complex structure J and define, on the complex vector space (V, J), a Hermitian inner product $\langle\cdot|\cdot\rangle$. Recall that a complex structure J is a linear map $J : V \to V$ such that $J^2 = -1$.

The one particle Hilbert space $\mathcal{H}$ would be the Cauchy completion of the complex inner product space $(V, J, \langle\cdot|\cdot\rangle)$. The complex structure must be compatible with the natural symplectic structure Ω on V;

$$\Omega(\Psi_1, \Psi_2) := \int_\Sigma \mathrm{d}S^a(\Psi_2 \nabla_a \Psi_1 - \Psi_1 \nabla_a \Psi_2)\,, \tag{4.27}$$

where Σ is any (3-dimensional) Cauchy surface on (M, g^{ab}). J is said to be compatible with Ω iff $\Omega(J\Phi_1, \Phi_2)$ is a symmetric, positive definite metric $g(\cdot,\cdot)$ on V.

The Hermitian inner product is thus defined,

$$\langle \Psi_1 |\, \Psi_2\rangle := \tfrac{1}{2} g(\Psi_1, \Psi_2) + \tfrac{1}{2} i\Omega(\Psi_1, \Psi_2)\,. \tag{4.28}$$

In the case of stationary space-times the complex structure can be defined from the generator of *time evolution* $\pounds_t$ on V ($\pounds_t$ is a well defined operator in V because it commutes with the Klein-Gordon operator). The complex structure

$$J := -(-\pounds_t \pounds_t)^{-\frac{1}{2}} \cdot \pounds_t \tag{4.29}$$

satisfies all the required properties and therefore gives to $(V, J, \langle\cdot|\cdot\rangle)$ the structure of a complex inner product space. Given a complex structure J, one can recover the more familiar language of positive and negative frequency decomposition. In this case the Hilbert Space $\mathcal{H}$ consists of positive frequency solutions Ψ^+ of the Klein-Gordon equation,

$$\Psi^+ := \tfrac{1}{2}(\Psi - i\, J\, \Psi)\,, \tag{4.30}$$

[2] More general conditions (i.e. non unitary evolution) have been considered in the literature [56], but we will not consider them here.

and the inner product takes the form,

$$\langle\Psi_1|\Psi_2\rangle = -i\,\Omega(\overline{\Psi_1^+}, \Psi_2^+)\,. \tag{4.31}$$

On this space the Schrödinger equation is given by

$$H\cdot\Psi^+ = i\,\hbar\,\pounds_t\Psi^+ \tag{4.32}$$

This covariant approach (in which one takes real solutions of the KG equation in the full space-time to be the phase space) is completely equivalent to the 'canonical' approach in which one considers the phase space to be pairs (f,p) where f is the initial wave function on Σ: $f = \Psi|_\Sigma$, and p is the initial normal derivative: $p = n^a\nabla_a\Psi|_\Sigma$ (See Appendix C for details).

4.4.2 Gravitational Aharonov-Bohm Effect

We can now investigate the gravitational Aharonov-Bohm effect by combining the results that were established in the previous sections. In Sec. 4.2, we found a solution to the Klein-Gordon equation outside a spinning cosmic string. Let us now confine our quantum system to a box situated at $R_i = (R_0, \theta_0, z_0)$ for $R_0 > 4J/(1-4\mu)$. This procedure allows us to get rid of two problems: the multivaluedness of the wave function and having to deal with closed time-like curves [3].

The boundary condition for the wave function inside the box is $\Psi|_{\partial\text{box}} = 0$. The Hilbert space structure for such solutions can be defined following the procedure outlined in Sec. 4.4.1.

Following the electro-magnetic case, we proceed by transporting the box round a closed circuit C. Since the space-time is axisymmetric we can transport the box along orbits of the Killing vector field R^a. The action of $\pounds_R$ commutes with the Klein-Gordon operator $(\Box - m^2)$. Consequently the action of transporting the box maps the Hilbert space into itself. Therefore we can restrict our analysis to the Hilbert space $\mathcal{H}_R$ for the box in the position R. Since we are considering the covariant approach, the unitary motion generated by $\pounds_R$ in $\mathcal{H}_R$ takes the state Ψ back to the same point in the Hilbert space.

All the framework required to discuss Berry's phase is established. We can now calculate the gravitational geometric phase using (4.21). We have

$$\gamma = i\oint_C \langle\Psi(R_i)|\pounds_R\Psi(R_i)\rangle\,, \tag{4.33}$$

where $\Psi(R_i)$ is given by (4.12).

The integrand in (4.33) is evaluated using the Klein-Gordon inner product (4.31) as follows

$$\begin{aligned}\langle\Psi(R_i)|\pounds_R\Psi(R_i)\rangle &= i\int_\Sigma \mathrm{d}S^a(\bar\Psi\nabla_a(\pounds_R\Psi) - (\pounds_R\Psi)\nabla_a\bar\Psi)\\ &= -i4EJ \end{aligned}\tag{4.34}$$

[3]For $R_0 < \frac{4J}{(1-4\mu)}$ the rotational Killing vector field $R^a = \left(\frac{\partial}{\partial\phi}\right)^a$, which has closed orbits, becomes time-like.

The geometric phase (4.33) is then

$$\gamma(C) = 8\,\pi\,J\,E \tag{4.35}$$

The effect can be observed by interfering the wave-function in the transported box and another that followed the orbits of the time-like Killing vector field t^a. Note that the geometric phase depends on the energy E of the particle, as opposed to the electro-magnetic AB case in which the phase is independent of the energy. Therefore, in order to have a cyclic evolution (required by the formalism) we may not superimpose eigenfunctions with different energy eigenvalues. However, there are bounded eigestates of the energy satisfying our boundary conditions. Such states can be constructed by taking superpositions of wave-functions of the form (4.14) as follows,

$$\Phi_n = e^{-4iJE_n\theta} \sum_l \int \mathrm{d}k\, \tilde{f}_l(k)\, e^{-il\theta}\, e^{i\sqrt{E_n^2-m^2-k^2}\,z}\, \mathrm{J}_\nu(k\rho)\,, \tag{4.36}$$

where $f(\rho,\theta) = \sum_l \int \mathrm{d}k\, \tilde{f}_l(k)\, e^{il\theta}\, \mathrm{J}_\nu(k\rho)$ is a localized function, that is, of compact support.

Let us conclude this section with a remark. We have avoided the region of closed time-like curves by restricting the particle to the box and therefore we do not have problems with the non-Hermiticity of the Hamiltonian [46].

4.5 Discussion

Let us summarize the main results of this chapter.

We have found that for those stationary space-times whose Killing vector field has a vanishing twist, solutions to the Klein-Gordon equation can be found using the Dirac phase factor. A geometric phase can thus be found for those space-times given that the existence of a Hermitian inner product in the Hilbert space of solutions is assured. The construction of such phase and its comparison with the electro-magnetic AB effect has been possible because we have interpreted both cases as the holonomy of a U(1) connection *on the space of states.* The phase depends on the geometry of the closed path in the ray space, but will be nonzero when the topology of the underlying space-time is nontrivial. We can say, therefore, that the phase is of a topological origin. For a spinning cosmic string, we have constructed the phase for a particle going around the string.

Two remarks are in order. In $2+1$ gravity, the form of the metric (4.9) for the spinning particle is the general asymptotic form at infinity for 'asymptotically flat' space-times. Therefore, the effect is always present near infinity for nontrivial topologies ($\Pi_1(\Sigma) \neq 0$), when the total angular momentum is non vanishing ($J \neq 0$). Secondly, our result is a generalization of the geometric phase for a Klein Gordon particle in Minkowski space-time studied in [57].

We have seen in this chapter that there is an interesting interplay between the topology of the background space-time and the holonomy of the natural connection on the space of rays $\mathcal{P}$. So far, the gravitational field has been considered to be fixed and non-dynamical. In the next Chapter we consider a dynamical space-time in 3-dimensions.

5

Gravity in 3-dimensions with a Cosmological Constant

5.1 Introduction

One of the main conceptual difficulties for the construction of a quantum theory of gravity comes from the absence of a background metric. The metric is now the dynamical variable, whose degrees of freedom are to be quantized. We can no longer apply ordinary methods of quantization that benefit from such background structure. Thus, it is always useful to consider simpler "toy models" that capture the same qualitative features of the full $3+1$ dimensional theory, but that are at the same time free of many technical difficulties. This is particularly true of $2+1$ gravity. Before 1984, roughly speaking, the theory was considered to be "too trivial" to deserve any attention. It was with the work of Deser, Jackiw and t'Hooft [53], who considered point particle solutions, and later with the paper by Witten [21], that an avalanche of papers on $2+1$ gravity followed. A very wide spectrum of issues have been addressed in this years, from the "problem of time" to the "loop representation", and more recently, black hole solutions. For recent reviews see [58, 59, 60] and particularly [61]. For a review of the $2+1$ black hole see [62].

There are still several conceptual difficulties of the full $3+1$ case that have not been addressed in the $2+1$ playground. The first one has to do with the so called "reality conditions". The new variables that Ashtekar introduced for $3+1$ gravity involve a complex connection [3]. At the classical level, the reality conditions that one should impose in order to recover the real theory are well understood. At the quantum level, however, one does not have yet a satisfactory treatment that yields the quantum theory corresponding to real Lorentzian general relativity (however, see [63, 64] and [65] for several proposals). As mentioned in the Introduction, $2+1$ gravity can be formulated as a theory of connections. The key difference with the $3+1$ case is that all this formulations are in terms of real connections, so there are apparently no reality conditions to worry about. There are two ways of writing the 3-dimensional action that involve connections. The Einstein-Hilbert-Palatini action has as independent fields a "spin" connection A and a (non-degenerate) co-triad field e. The "pure-connection" formulation considers a Chern-Simons action where A and e are united into a single connection. Since we are at the end interested in learning something about the full 4 dimensional theory, it is convenient to deal with the connection formulation that is closer to the

$3+1$ case. This is the Einstein-Hilbert-Palatini theory. In the case of a vanishing cosmological constant, the theory has been extensively studied and understood. The same is not true, however, for the case of a non vanishing constant. For, in this case, the geometric meaning of the equations is not so transparent as in the flat case.

In this chapter we consider the Palatini formulation of 3-dimensional gravity with a cosmological constant. We analyze the $2+1$ decomposition of the theory in order to get a Hamiltonian description. From the four possible combinations of signature of the underlying space-time and sign of the cosmological constant, we restrict out attention to two of them. These are Lorentzian signature with *positive* cosmological constant and Euclidean space-time with a *negative* constant. We are interested in this cases since a complex change of variables on phase space simplifies the constraints considerably (the complementary cases can be simplified by a similar transformation but in terms of real connections). The new constraints tell us that we have a theory of flat connections, and the reduced space is simply the moduli space under gauge transformations. However, the theory is now formulated in terms of complex connections, in analogy with the $3+1$ case. Therefore, the stage is set for studying the issue of the reality conditions. Can we quantize the theory and recover the real theory we started with? As we shall see, this question can be answered in the affirmative.

The second issue we shall consider in this chapter has to do with the recovery of the space-time from the canonical quantum theory. When one works in the so called "frozen formalism", i.e., in the reduced phase space of the theory, one is dealing with the "true degrees of freedom" of the theory. That is, roughly speaking, an equivalence class of classical solutions where one does not distinguish between configurations that are related by "gauge". In the quantum theory, this true degrees of freedom are "fluctuating" excitations at the microscopic level. By general arguments, one should expect that the signature of space-time, the classical limit of the theory, be somehow "encoded" in the microscopic theory. One anticipates then, that the details of the quantum theory dictate unambiguously the signature of the resulting space-time. As $2+1$ gravity shows us this expectation is, however, not satisfied. For, in the cases of interest ($(-++)$ signature with $\Lambda > 0$ and $(+++)$ with $\Lambda < 0$), the reduced phase spaces *coincide*. Thus, we have a mathematical identical description of the two systems. This is quite surprising and disturbing at first. But, as it is the case in many physical situations, the "physics" is in the interpretation of the formalism. As we shall see, there is a precise way to reconstruct two space-times, each one with a different signature, for each point of the (reduced) phase space. This construct also provides an explicit isomorphism between this space-times, or using a "Wheelerism", a "Wick rotation without Wick rotation". Thus, extra input is needed in order to recover the precise theory we want to consider. The implications of this fact for the $3+1$ theory are not clear yet.

This chapter is organized as follows. In Sec. 5.2 we consider the Palatini formulation of 3-dimensional gravity. We perform a $2+1$ decomposition in order to go to the Hamiltonian formulation. A complex transformation is performed in phase space that simplifies the constraints and allows for a simple characterization of the reduced phase space of the theory. In Sec. 5.3 we restrict our attention to the simplest case of spatial topology: a two torus. We give explicit parametrization of the reduced phase space. The quantization of this reduced phase space is considered in Sec. 5.4. In Sec. 5.5 we relate the phase space with the space-time picture, making the isomorphism between Euclidean and Lorentzian space-times manifest. We conclude with a discussion is Sec. 5.6.

Throughout this chapter we use units in which $c = 1$, but write G and $\hbar$ explicitly.

5.2 Palatini Formulation of $2+1$ Gravity

In this section we will recall the Palatini formulation of 3-dimensional gravity. We will start from the standard Einstein-Hilbert action written in terms of triads and a connection taking values in the Lie algebra of the relevant gauge group. We will then perform the $2+1$ decomposition of the action, arriving to a first order expression from which we can "read off" the Hamiltonian formulation. We will then define new complex coordinates on the phase space in which the constraints simplify in a significant way. We can give a clear interpretation to the geometrical content of the constraints and characterize completely the reduced phase space.

5.2.1 The Action

In this chapter we consider 3-dimensional gravity with a cosmological constant. The signature of the metric could be Euclidean or Lorentzian. Since we are going to consider gravity to be a theory of connections, the only thing that will change will be the gauge group of the connection. In the case of Euclidean signature, the group is SO(3) (or its double coverture SU(2)). The gauge group for Lorentzian space-times is SO(2,1) (SU(1,1)). The Einstein-Hilbert action for 3-dimensional gravity [22, 66],

$$S_{\mathrm{EH}}[g^{ab}] := \frac{1}{G}\int_M \mathrm{d}^3x\,\sqrt{|g|}(R-2\Lambda)\,, \tag{5.1}$$

can be rewritten in terms of a connection and a co-triad field as follows,

$$S_{\mathrm{P}}[A_a, e_b] := \frac{1}{4G}\int_M \mathrm{d}^3x\,\tilde{\eta}^{abc}e_{aI}\epsilon^{IJK}\left(F_{bcJK} - \frac{2\Lambda}{3}e_{bJ}e_{cK}\right)\,, \tag{5.2}$$

where the capital indices denote internal vectors, that is, vectors in a fiducial 3 dimensional vector space W (see Appendix A). The field e_a^I is the soldering form that maps internal vectors to tangent vectors to the 3 manifold M: $v^I := e_a^I v^a$. In the internal vector space W there is a fixed fiducial metric g_{IJ} whose signature coincides with the one of the space-time metric. Therefore, in the case of Euclidean gravity $g_{IJ} = \delta_{IJ}$ and for Lorentzian signature it is the 3 dimensional Minkowski metric with signature $(-,+,+)$. There is a preferred volume element ϵ_{IJK} in W defined from g_{IJ}. The ϵ with "upstairs" indices is defined from ϵ_{IJK} by "raising the indices" with g^{IJ}. We can recover the space-time metric g_{ab} from the soldering form e_a^I,

$$g_{ab} := e_a^I e_b^J g_{IJ}\,. \tag{5.3}$$

The inverse of e_a^I is denoted by e_I^a and satisfies

$$g_{ab}e_I^a e_J^b = g_{IJ}\,. \tag{5.4}$$

The action functional depends also on a connection ${A_{aI}}^J$ on M that "acts" on internal indices. In other words, there is a covariant derivative that acts on objects with space-time and internal indices in the following way,

$$\mathcal{D}_a k_{bI} := \partial_a k_{bI} + {A_{ab}}^c k_{cI} + {A_{aI}}^J k_{bJ}. \tag{5.5}$$

Given such a "generalized derivative operator" $\mathcal{D}_a$ we can construct generalized curvature tensors,

$$2\mathcal{D}_{[a}\mathcal{D}_{b]} k_I \quad =: \quad {F_{abI}}^J k_J \tag{5.6}$$

$$2\mathcal{D}_{[a}\mathcal{D}_{b]} k_c \quad =: \quad {F_{abc}}^d k_d\,. \tag{5.7}$$

We are considering the Palatini theory which means that the fields e and A are to be varied independently in the action principle. Finally, $\tilde{\eta}^{abc}$ is the naturally defined Levi-Civita density-one tensor field on M. In the next subsection we will consider the equations of motion coming from the action for both signatures.

Equations of Motion

We can find the equations of motion coming from the action (5.2) by varying with respect to the independent variables (A, e). The variation with respect to ${A_{aI}}^J$ yields,

$$\mathcal{D}_a(\tilde{\eta}^{abc}\epsilon_{IJK}e_c^K) = 0\,, \tag{5.8}$$

and the equation from e is,

$$\tilde{\eta}^{abc}\epsilon_{IJK}(F_{bc}^{JK} - \Lambda e_b^J e_c^K) = 0. \tag{5.9}$$

The first equation implies that the covariant derivative defined by ${A_{aI}}^J$ coincides with the one compatible with the co-triad, that is, ${A_{aI}}^J = {\Gamma_{aI}}^J$. We can therefore replace F by R in (5.9) and we arrive at Einstein equations:

$$G^{ab} + \Lambda g^{ab} = 0\,. \tag{5.10}$$

Recall that this equation implies that the space-time has constant scalar curvature proportional to Λ:

$$g_{ab}(R^{ab} - \frac{1}{2}Rg^{ab} + \Lambda g^{ab}) = -\frac{1}{2}R + 3\Lambda = 0\,. \tag{5.11}$$

Therefore, $R = 6\Lambda$. Note that this result is independent of the signature of the space-time and therefore, of the gauge group we are considering.

At this point it is convenient to use a fact from 3 dimensions: the dimensions of the "Lorentz group" and the manifold coincide. That is, the Lie algebra $so(2,1)(so(3))$ and the internal space W can be identified. Furthermore, the Killing-Cartan metric on the Lie algebra k_{IJ} is proportional to the internal metric g_{IJ}. The idea now is to change the connection from a two-index object to a

"Lie-Algebra valued" connection. We can achieve this by using the Levi-Civita symbol ϵ_{IJK} in W. Let us define,

$$A_{aI}{}^{J} =: A_a^K \, \epsilon^J{}_{IK} \tag{5.12}$$

In terms of representation of Lie algebras we are passing from the defining to the adjoint representation. It is in this step that the Euclidean and Lorentzian theories have different expressions. This can be understood from the fact that we are using the internal metric g to raise the indices of the ϵ_{IJK} in equation (5.12): $\epsilon^J{}_{IK} := g^{JM}\epsilon_{MIK}$. We can now rewrite the generalized covariant derivative in terms of A_a^I as,

$$\mathcal{D}_a v^I = \partial_a v^I + [A_a, v]^I, \tag{5.13}$$

where $[A_a, v]^I = \epsilon^I{}_{JK}\, A_a^J\, v^K$. From (5.12) it also follows that the generalized curvature tensor F_{ab}^I is,

$$F_{ab}^I = 2\,\partial_{[a} A_{b]}^I + [A_a, A_b]^I. \tag{5.14}$$

The Palatini actions can be rewritten as,

$$\begin{aligned} S_L[e,A] &= \frac{1}{2G}\int_M \mathrm{d}^3x\, \tilde{\eta}^{abc} e_{aI}\left(F_{bc}^I - \frac{\Lambda}{3}\epsilon^{IJK} e_{bJ} e_{cK}\right) && \text{Lorentzian,} \\ S_E[e,A] &= -\frac{1}{2G}\int_M \mathrm{d}^3x\, \tilde{\eta}^{abc} e_{aI}\left(F_{bc}^I + \frac{\Lambda}{3}\epsilon^{IJK} e_{bJ} e_{cK}\right) && \text{Euclidean.} \end{aligned} \tag{5.15}$$

The equations of motions from this actions are just equivalent to (5.8) and (5.9).

5.2.2 Legendre Transform

In this subsection, we will perform a $2+1$ decomposition of the manifold M and we will accordingly rewrite the actions (5.15) in a $2+1$ fashion.

In order to make the decomposition, we have to assume that the manifold M is of the form $\Sigma \times \mathbb{R}$ and that there exists a globally defined function t (with nowhere vanishing gradient $(\mathrm{d}t)_a$) such that the hyper-surfaces of constant value Σ_t are diffeomorphic to Σ. For simplicity we will restrict to Σ compact and orientable. We have to assume that there exist a globally defined vector field t^a defining the "time flow" satisfying $t^a(\mathrm{d}t)_a = 1$.

The next step is to decompose $\tilde{\eta}^{abc}$ in terms of t^a and $\tilde{\eta}^{ab}$ (The Levi-Civita tensor density of weight one on Σ). Using $\tilde{\eta}^{abc} = 3t^{[a}\tilde{\eta}^{bc]}\mathrm{d}t$, we get (modulo a surface integral)

$$\begin{aligned} S[e,A] &= \pm\frac{1}{2G}\int_M \mathrm{d}^3x\, \tilde{\eta}^{abc} e_{aI}\left(F_{bc}^I \mp \frac{\Lambda}{3}\epsilon^{IJK} e_{bJ} e_{cK}\right) \\ &= \pm\frac{1}{G}\int dt \int_\Sigma \mathrm{d}^2x\, \tilde{E}_I^c \pounds_t A_c^I + (\mathcal{D}_a \tilde{E}_I^a)(A\cdot t)^I + \frac{1}{2}(e\cdot t)_I(\tilde{\eta}^{ab} F_{ab}^I \mp \Lambda\epsilon^{IJK} \underset{\sim}{\eta}_{ab} \tilde{E}_J^a \tilde{E}_K^b)\,. \end{aligned} \tag{5.16}$$

The variables are $(e\cdot t)_I := t^a\, e_{aI}$, $\tilde{E}_I^a := \tilde{\eta}^{ab} e_{bI}$ and $(A\cdot t)^I := t^a\, A_a^I$. A_a^I is now the pull-back to Σ of the original connection. We have used $(+,-)$ to denote Lorentzian and Euclidean signatures respectively.

From this 2+1 form of the action we can read off what the canonical variables are. The expression in already in first order form and has the structure $\int p\dot{q} - H$. The Lie derivative with respect to the vector field t^a is to be interpreted as "time derivative", so the conjugate momenta to A is the field $\tilde{E}^a_I$[1]. The fields $(A\cdot t)_I$ and $(e\cdot t)_I$ play the role of Lagrange multipliers. We can see then that the phase space $(\Gamma_G, \Omega_{\mathbf{G}})$ is coordinatized by the pair $(A^I_a, \tilde{E}^a_I)$ and has symplectic structure,

$$\Omega_G = \frac{1}{G}\int_\Sigma \mathbf{d}\tilde{E}^a_I \wedge \mathbf{d}A^I_a\,, \tag{5.17}$$

where $G = \mathrm{SO}(2,1)$ for Lorentzian gravity and $G = \mathrm{SO}(3)$ for Euclidean. With this symplectic structure (5.17) the Poisson Bracket of the basic variables is,

$$\{A^I_a(y), \tilde{E}^b_J(x)\} = G\delta^b_a\delta^I_J\delta^3(x,y)\,. \tag{5.18}$$

The Hamiltonian is given by

$$H_G(A,\tilde{E}) = \mp\frac{1}{G}\int_\Sigma \frac{1}{2}(e\cdot t)_I(\tilde{\eta}^{ab}F^I_{ab} \mp \Lambda\epsilon^{IJK}\underset{\sim}{\eta}_{ab}\tilde{E}^a_J\tilde{E}^b_K) - (A\cdot t)^I(\mathcal{D}_a\tilde{E}^a_I)\,\mathrm{d}^2x\,. \tag{5.19}$$

We see that the expressions multiplying $(A\cdot t)^I$ and $(e\cdot t)_I$ are constraints on the canonical variables. That is, if we vary the action with respect to them, we have

$$\mathcal{D}_a\tilde{E}^a_I = 0\,, \tag{5.20}$$

which is the *Gauss law*, and

$$\tilde{\eta}^{ab}F^I_{ab} \mp \Lambda\epsilon^{IJK}\underset{\sim}{\eta}_{ab}\tilde{E}^a_J\tilde{E}^b_K = 0\,. \tag{5.21}$$

This constraint can be thought of as the *dynamical constraint.*

Algebra of Constraints

Notice that both constraints have a "free internal index", so in order to have constraint functions on Γ_G we can give smearing fields α_I and v^I and define,

$$F^{\pm}[\alpha] := \frac{1}{2G}\int_\Sigma \mathrm{d}^2x\,\alpha_I(\tilde{\eta}^{ab}F^I_{ab} \pm \Lambda\epsilon^{IJK}\underset{\sim}{\eta}_{ab}\tilde{E}^a_J\tilde{E}^b_K) \qquad \text{and} \qquad G[v] := \frac{1}{G}\int_\Sigma \mathrm{d}^2x\, v^I\mathcal{D}_a\tilde{E}^a_I. \tag{5.22}$$

These constraints functions generate infinitesimal symplectomorphisms on Γ_G via their Hamiltonian vector field. Accordingly, they generate infinitesimal change in the coordinates (A, E) as follows. The constraint function $F[\alpha]$ generates,

$$\begin{aligned} A^I_a &\rightarrow A^I_a \mp \epsilon(\Lambda\epsilon^{IJK}\underset{\sim}{\eta}_{ab}\tilde{E}^b_J\alpha_K) + O(\epsilon^2) \qquad \text{and} \\ \tilde{E}^a_I &\rightarrow \tilde{E}^a_I - \epsilon(\tilde{\eta}^{ab}\mathcal{D}_b\alpha_I) + O(\epsilon^2)\,. \end{aligned} \tag{5.23}$$

[1]To be precise, it is $\frac{1}{G}\tilde{E}^a_I$

Similarly, under the diffeomorphism generated by $G[v]$ we have,

$$\begin{aligned} A_a^I &\rightarrow A_a^I - \epsilon \mathcal{D}_a v^I + O(\epsilon^2) \quad \text{and} \\ \tilde{E}_I^a &\rightarrow \tilde{E}_I^a - \epsilon\{v, \tilde{E}^a\}_I + O(\epsilon^2)\,, \end{aligned} \tag{5.24}$$

where $\{v, \tilde{E}^a\}_I := \epsilon^K{}_{JI} v^J \tilde{E}^a_K$.

We can compute the Poisson bracket of the constraint functions $F[\alpha]$ and $G[v]$. We find that for the Gauss constraint,

$$\{G[v], G[w]\} = G\left[[v,w]\right]\,, \tag{5.25}$$

where $[v,w]^I = \epsilon^I{}_{JK} v^J w^K$ is the Lie bracket of v^I and w^J, so that $v^I \rightarrow G[v]$ is a representation of the Lie algebra $\mathbf{g}$. We also have that

$$\{G[v], F^{\pm}[\alpha]\} = -F^{\pm}[\{v,\alpha\}]\,, \tag{5.26}$$

where $\{v, \alpha\} = \epsilon^K{}_{JI} v^J \alpha_K$ is the "co-adjoint bracket" of v^I and α_J. Finally,

$$\{F^{\pm}[\alpha], F^{\pm}[\beta]\} = \mp \Lambda G[\epsilon(\alpha,\beta)]\,, \tag{5.27}$$

where $\epsilon(\alpha,\beta) := \epsilon^{IJK}\alpha_J \beta_K$. Therefore, the algebra of the constraints is closed under Poisson bracket, that is, they are of first class in Dirac's terminology. Furthermore, since (5.25), (5.26) and (5.27) do not involve any functions on Γ they form a Lie algebra. In fact, they form a representation of the $\Lambda-$*deformation*, ΛG, of the Lie group G [66]. If we start with the group SO$(2,1)$, its deformations are SO$(3,1)$ for $\Lambda > 0$ and SO$(2,2)$ for $\Lambda < 0$. For the Euclidean case, the deformations of SO(3) are SO(4) for $\Lambda > 0$ and SO$(3,1)$ for $\Lambda < 0$. Note that the algebra of the constraints coincide for Lorentzian gravity with a *positive* cosmological constant and for Euclidean gravity with a *negative* cosmological constant.

Recall that for a system with first class constraints, the reduced phase space $\hat{\Gamma}$ is the space of gauge orbits generated by the constraints on the constraint surface. A naive counting of the local degrees of freedom of the system gives zero. This is because we have 12 local degrees of freedom from the pair $(A_a^I, \tilde{E}_J^a)$, minus 6 constraints per point. Since they are first class they generate gauge symmetries, so we have: Local degrees of freedom = (12 variables - 6 constraints - 6 gauge orbits) = 0. However, it is not straightforward to give a characterization of the reduced phase space $\hat{\Gamma}$ in terms of this variables as it is the case for $\Lambda = 0$. If we choose the connection A_a^I as configuration variable, we can identify Γ with the cotangent bundle T^*C over the configuration space C, where $\tilde{E}$ is a coordinate for the momenta. The Gauss constraint is linear in momenta so its action can be easily understood. The other constraint is however *quadratic* in momenta so there is no clean geometric interpretation for it. Recall that in the case of vanishing cosmological constant, $\hat{\Gamma}$ could be identified as the cotangent bundle $T^*\hat{C}$ over a *reduced* configuration space $\hat{C}$. We have lost this interpretation in our case. In the next part, we will see that by performing a complex transformation of the basic variables, we will gain a better understanding of the geometry of the constraints.

5.2.3 Complex Transformation

As we saw in the last part, in terms of the variables (A, E) it is not easy to treat the constraints (5.21) and (5.20) and therefore we do not have a characterization of the reduced phase space $\hat{\Gamma}$. Consider the new set of variables,

$$\mathbf{A}_a^I := A_a^I + i\sqrt{|\Lambda|}e_a^I, \quad \text{and} \quad \overline{\mathbf{A}_a^I} := A_a^I - i\sqrt{|\Lambda|}e_a^I. \tag{5.28}$$

Where $\overline{\mathbf{A}_a^I}$ is the complex conjugate of $\mathbf{A}_a^I$. Recall that e_a^I is dimension-less and that the dimensions of A_a^I are L^{-1}. We have at our disposal the quantity $\sqrt{|\Lambda|}$ that has the right dimensions of L^{-1}. Let us call $\lambda := \sqrt{|\Lambda|}$.

The phase space will now be coordinatized by the pair $(\mathbf{A}_a^I, \overline{\mathbf{A}_a^I})$. The symplectic structure Ω_G takes the form,

$$\Omega_G = \frac{-i}{4G\lambda}\int_\Sigma \mathrm{d}^2x \left(\mathbf{dA}_a^I \wedge \mathbf{dA}_b^J - \mathbf{d}\overline{\mathbf{A}_a^I} \wedge \mathbf{d}\overline{\mathbf{A}_b^J}\right)\tilde{\eta}^{ab}k_{IJ}\,, \tag{5.29}$$

The Poisson bracket of the new variables is,

$$\begin{aligned} \{\mathbf{A}_a^I(x), \mathbf{A}_b^J(y)\} &= 2\,i\,G\lambda\,\underset{\sim}{\eta}_{ab}k^{IJ}\delta^3(x,y) \\ \{\overline{\mathbf{A}_a^I}(x), \overline{\mathbf{A}_b^J}(y)\} &= -2\,i\,G\lambda\,\underset{\sim}{\eta}_{ab}k^{IJ}\delta^3(x,y) \\ \{\mathbf{A}_a^I(x), \overline{\mathbf{A}_b^J}(y)\} &= 0\,. \end{aligned} \tag{5.30}$$

Note that the only thing we are doing is to define complex coordinates on the phase space Γ and to re-write the *real* symplectic structure (5.17) in the new coordinates. It is very much like going from (q,p) to $(z,\overline{z})$ coordinates for the harmonic oscillator. The difference is however, that the transformation we have defined is not canonical, in the sense that the new coordinates $\mathbf{A}_a^I$ do not Poisson commute among themselves.

Constraints

We will now consider the constraints, when written in terms of the new variables. In order to do that, let us write down the curvature of the new complex connection,

$$\begin{aligned} \mathbf{F}_{ab}^I &:= 2\partial_{[a}\mathbf{A}_{b]}^I + \epsilon^I{}_{JK}\mathbf{A}_a^J\mathbf{A}_b^K \\ &= \left(F_{ab}^I - |\Lambda|\epsilon^I{}_{JK}e_a^Je_b^K\right) + i\,2\mathcal{D}_{[a}e_{b]}^I. \end{aligned} \tag{5.31}$$

We notice that if $\Lambda > 0$, the real part of the curvature of the new connection is the constraint (5.21) for *Lorentzian* gravity with *positive* cosmological constant. On the other hand, if $\Lambda < 0$, it corresponds to the constraint (5.21) for *Euclidean* signature with a *negative* cosmological constant. The Imaginary part of (5.31) is in either case (two times) the Gauss constraint (5.20). Thus, imposing the two constraints (5.21) and (5.20) is equivalent to imposing,

$$\mathbf{F}_{ab}^I = 0 \quad \text{and} \quad \overline{\mathbf{F}_{ab}^I} = 0. \tag{5.32}$$

That is, for the above mentioned cases $[(so(2,1), \Lambda > 0)$ and $(so(3), \Lambda < 0)]$ the only constraint is that the complex connection should be *flat*. The next step is to analyze the infinitesimal symplectomorphisms generated by the constraints (5.32) with respect to the symplectic structure (5.29). We can construct constraint functions by smearing with respect to (complex fields),

$$\mathbf{F}[\alpha] := \frac{1}{4G}\int_\Sigma \mathrm{d}^2x\, \tilde{\eta}^{ab}\alpha_I \mathbf{F}^I_{ab} \qquad \text{and} \qquad \overline{\mathbf{F}}[\beta] := \frac{1}{4G}\int_\Sigma \mathrm{d}^2x\, \tilde{\eta}^{ab}\beta_I \overline{\mathbf{F}}^I_{ab}. \tag{5.33}$$

We notice that the Poisson bracket of the constraint functions with the coordinates $\mathbf{A}$ and $\overline{\mathbf{A}}$ is given by:

$$\{\mathbf{A}^I_a, \mathbf{F}[\alpha]\} = i\lambda\, \mathbf{D}_a\alpha^I \quad ; \quad \{\overline{\mathbf{A}}^I_a, \overline{\mathbf{F}[\beta]}\} = -i\lambda\, \overline{\mathbf{D}}_a\beta^I \tag{5.34}$$

and,

$$\{\mathbf{A}^I_a, \overline{\mathbf{F}[\beta]}\} = \{\overline{\mathbf{A}}^I_a, \mathbf{F}[\alpha]\} = 0\,. \tag{5.35}$$

where $\mathbf{D}$ is the new generalized covariant derivative,

$$\mathbf{D}_a k^I := \partial_a k^I + \epsilon^I{}_{JM}\mathbf{A}^J_a k^M. \tag{5.36}$$

Under the infinitesimal symplectomorphism generated by the constraint function $\mathbf{F}[\alpha]$ the connection transforms as,

$$\mathbf{A}^I_a \to \mathbf{A}^I_a - i\lambda\,\epsilon\, \mathbf{D}_a\alpha^I \tag{5.37}$$

Finally, the Poisson Bracket of two constraints is given by,

$$\{\mathbf{F}[\alpha], \mathbf{F}[\beta]\} = i\,\lambda \mathbf{F}[\epsilon(\alpha,\beta)]. \tag{5.38}$$

And similarly for the constraint function $\overline{\mathbf{F}}[\beta]$. Thus, the $\mathbf{A}$ and $\overline{\mathbf{A}}$ sectors of the phase space "decouple" from each other, and are also not independent (they are the complex conjugate of each other). We can, therefore, focus only the $\mathbf{A}$ coordinate from now on. Notice that the constraint functions $\mathbf{F}[\alpha]$ are of first class and generate *gauge transformation* on the connection $\mathbf{A}$. With the new coordinates we have a complete characterization of the reduced phase space: $\hat{\Gamma}$ is the space of *flat* connections modulo *gauge* transformations, or as it is normally known, the *moduli space* of flat connections.

Let us summarize our results. We started with coordinates $(A, \tilde{E})$ for Γ where A is a Lie-algebra valued connection where the gauge group is $\mathrm{SO}(2,1)$ for Lorentzian gravity and $\mathrm{SO}(3)$ for Euclidean. We saw that by performing a complex transformation, in the case in which the cosmological constant is positive for $\mathrm{SO}(2,1)$ and negative for $\mathrm{SO}(3)$, we could simplify the constraints considerably. We arrived at a characterization of the reduced phase space in terms of moduli space of flat connection. The question that now arises is: what connection are we talking about? is it the same reduced phase space $\hat{\Gamma}_G$ for the euclidean and Lorentzian cases? We will shortly give answers to those questions.

First, recall that the algebra of the original constraints (5.25)-(5.27), corresponding to the Λ-deformation of the gauge group coincided for the cases we are considering; for both SO(2,1) with $\Lambda > 0$ and SO(3) with $\Lambda < 0$, the Λ-deformed group is SO(3,1). How is the Lorentz group in four dimension related to the complex connections we defined? Consider for instance the case of SO(2,1) to begin with. When we perform the complex transformation (5.28), we are defining a new connection taking values in the *complexification* of the Lie algebra $so(2,1)$. The gauge group is therefore a suitable complexification of the original one. It turns out that the complexified Lie algebra of $so(2,1)$ and $so(3)$ *coincide* and correspond to $sl(2,\mathbb{C})$. The complex gauge group is now $\mathrm{SO}(3,\mathbb{C})$ which can be recovered from its double cover $\mathrm{SL}(2,\mathbb{C})$, by taking the quotient over a $\mathbb{Z}_2$ action, in the same way that SU(2) projects down to SO(3). The relation between (the connected component of the identity of) the Lorentz group SO(3,1) and $\mathrm{SL}(2,\mathbb{C})$ is given at two levels. First, at the Lie-algebra level by an isomorphism between $sl(2,\mathbb{C})$ and the *self-dual* part of the (complexified) $so(3,1)$. At the group level, there is a 2 to 1 mapping from $\mathrm{SL}(2,\mathbb{C})$ to SO(3,1), extensively used in the representation theory of the Lorentz group. Therefore, there is a precise sense in which one can identify $\mathrm{SO}(3,\mathbb{C})$ with the restricted Lorentz group $\mathrm{SO}^{\uparrow}_{+}(3,1)$. We shall make use of this identification in next section.

How can we have the *same* reduced phase space $\hat{\Gamma}$ for both $(\mathrm{SO}(3), \Lambda < 0)$ and $(\mathrm{SO}(2,1), \Lambda > 0)$? Each point of $\hat{\Gamma}$ represents a physically distinct configuration, which in the space-time picture means a diffeomorphism equivalence class of space-time metrics $[g_{ab}]$. If both the Lorentzian and Euclidean theories are to share the same reduced phase space, then we should be able to produce, given any point of $\hat{\Gamma}$, a Euclidean space-time with negative constant curvature and a Lorentzian space-time with positive curvature. We shall see in the last Section that in the case of spatial slices diffeomorphic to the torus, this is indeed possible, and both space-times can be explicitly constructed.

5.3 The Torus T^2

In this section we restrict our attention to the case in which the hyper-surface Σ is diffeomorphic to the 2-torus, T^2. We parametrize explicitly the reduced phase space $\hat{\Gamma}$ in terms of two complex coordinates and show that it has the structure of a cotangent bundle.

In Section 5.2 we arrived at the conclusion that the reduced phase space $\hat{\Gamma}$, for either signature of the space-time, corresponds to the moduli space of flat $\mathrm{SO}(3,\mathbb{C})$ connections on Σ.

There is a very convenient way of characterizing the space $\hat{\Gamma}$, in terms of "almost" gauge invariant objects on Σ. For that, fix a point p on Σ. Consider the holonomies around all possible loops γ that have p as their base point. As G-valued functions (assuming a trivial bundle) on the phase space they are invariant under gauge transformation that are the identity at p (one can, of course take all possible gauge transformation, but then the holonomies are only gauge covariant: one has to get rid of a trivial conjugation at the base point p). If we now restrict to points on the constraint surface, that is, to *flat* connections, then we have a further simplification. For any two loops that are homotopic to each other, the holonomy of a flat connection is the same, so the only property of the loop we are interested in is its homotopy class $[\gamma]$. Therefore, we have,

$$\hat{\Gamma} \simeq \frac{\Pi_1(\Sigma) \xrightarrow{\rho} G}{\text{Adj}(G)}\,. \tag{5.39}$$

That is, $\hat{\Gamma}$ is in one to one correspondence with all homomorphisms ρ from the fundamental group of Σ to the gauge group G, modulo the adjoint action of the group G at the base point p. In our case, the reason why we are able to give a simple description for the reduced phase space $\hat{\Gamma}$ is that the fundamental group of the torus is Abelian. In a sense, the genus-one surface is "singular" compared to the higher genus surfaces. In fact $\Pi_1(\text{T}^2) = \mathbb{Z} \oplus \mathbb{Z}$, so we have to find a representation of the fundamental group into $\text{SO}(3,\mathbb{C})$. At this point it is convenient to consider the double coverture of the group, that is, we will for simplicity work with $G = \text{SL}(2,\mathbb{C})$ and later we shall project down to $\text{SO}(3,\mathbb{C})$ (and $\text{SO}^{\uparrow}_{+}(3,1)$).

We know that the fundamental group has two generators γ_1 and γ_2, and that the holonomies corresponding to them $H(\mathbf{A},\gamma_i)$ must commute. We are therefore restricted to an Abelian subgroup of $\text{SL}(2,\mathbb{C})$, modulo conjugation. The elements of the maximal Abelian subgroup, up to conjugation, are diagonal matrices,

$$\begin{pmatrix} z & 0 \\ 0 & z^{-1} \end{pmatrix} \tag{5.40}$$

where z is a complex number, and $z \neq 0$. There is still one further identification given by conjugation by elements of the form

$$\begin{pmatrix} 0 & -d \\ d^{-1} & 0 \end{pmatrix} \tag{5.41}$$

so that we have to identify,

$$\begin{pmatrix} z & 0 \\ 0 & z^{-1} \end{pmatrix} \sim \begin{pmatrix} z^{-1} & 0 \\ 0 & z \end{pmatrix} \tag{5.42}$$

Note that we can coordinatize $\hat{\Gamma}$ by two complex numbers (z_1, z_2), where $z_i \in \mathbb{C}^*$ characterizes the holonomy along the "i" generator. We still have the identification $(z_1, z_2) \sim (z_1^{-1}, z_2^{-1})$[2]

We can conclude that the reduced phase space $\hat{\Gamma}$ is homeomorphic to,

$$\hat{\Gamma} = \frac{(\mathbb{C}^* \times \mathbb{C}^*)}{\mathbb{Z}_2}\,, \tag{5.43}$$

which is also homeomorphic to $(\text{T}^2 \times R^2)/\mathbb{Z}_2$. As we shall see below, it has the structure of a cotangent bundle over the torus. Note that this "new" torus is *not* the space-like manifold Σ we started with.

We can give real coordinates to $\hat{\Gamma}$ by noting that we can write $z_i \in \mathbb{C}^*$ as,

$$z_j := \exp[s_j + i a_j], \tag{5.44}$$

[2] Just as in the case of the $(\mathbf{A}^I_a, \overline{\mathbf{A}}^I_a)$ coordinates for the phase space, we can consider the complex conjugates $(z_1, \overline{z}_1, z_2, \overline{z}_2)$ to coordinatize $\hat{\Gamma}$. To simplify the discussion, we shall consider only the coordinates z_i in the remaining of this section.

where $s_j \in (-\infty, \infty)$ and $a_j \in [0, 2\pi)$. Therefore, we can also parametrize $\hat{\Gamma}$ by four real coordinates (s_1, s_2, a_1, a_2) with the identifications $(s_1, s_2, a_1, a_2) \sim (-s_1, -s_2, -a_1, -a_2)$.

It is a well known fact from the theory of reduction of symplectic systems that the reduced phase space $\hat{\Gamma}$ inherits a symplectic structure $\hat{\Omega}$ from the original phase space Γ. In order to see what the induced symplectic structure is, we will find a representative $\mathbf{A}_a^I$ of the gauge equivalence class given by (z_1, z_2) and from there we shall find the induced symplectic form. Let us consider periodic coordinates (x_1, x_2) on the hyper-surface Σ, where $x_i \in [0, 1)$. Since these are in a sense, "globally defined coordinates", we can find a gauge such that the connection $\mathbf{A}_a^I$ is a constant and can be written as,

$$\mathbf{A}_a := (\mathbf{A}_1^I (\mathrm{d}x_1)_a + \mathbf{A}_2^I (\mathrm{d}x_2)_a)\tau_I \,, \tag{5.45}$$

where τ_I is a basis for $sl(2, \mathbb{C})$. We can choose $\mathbf{A}_a^I$ to be of the form,

$$\mathbf{A}_a = \tau_3 \left[(\zeta_1 - i\eta_1)(\mathrm{d}x_1)_a + (\zeta_2 - i\eta_2)(\mathrm{d}x_2)_a \right] \,. \tag{5.46}$$

Note that in this particular gauge, the connection takes values in a "constant" direction in the Lie algebra, and therefore, the path order exponential will be just the ordinary exponential of the Lie algebra element, so the holonomy around the class $[\gamma]$ is,

$$H(\mathbf{A}, [\gamma]) = \begin{pmatrix} \exp\left(\frac{i}{2}[n(\zeta_1 - i\eta_1) + m(\zeta_2 - i\eta_2)]\right) & 0 \\ 0 & \exp\left(\frac{-i}{2}[n(\zeta_1 - \eta_1) + m(\zeta_2 - i\eta_2)]\right) \end{pmatrix} \tag{5.47}$$

where (n, m) denotes the number of times that $[\gamma]$ winds around the generators γ_1 and γ_2 respectively. Note however that we can not identify the coordinates (a_i, s_1) with (ζ_i, η_i). In fact, they are related by a factor of 2. This is precisely the factor that defines the 2 to 1 projection from $\mathrm{SL}(2, \mathbb{C})$ to $\mathrm{SO}(3, \mathbb{C})$ where the mapping is of the form $\theta \to 2\theta$.

It is important to distinguish at this point between both groups. Since the connections we started with were either in $\mathrm{SO}(2,1)$ or $\mathrm{SO}(3)$, we expect the holonomies of the complex connection to be in $\mathrm{SO}(3, \mathbb{C})$, rather than in $\mathrm{SL}(2, \mathbb{C})$. This means that in (5.45) we should write the generator iJ_3 of, say, rotations in the z direction, instead of τ_3,

$$\mathbf{A}_a = iJ_3 \left[(a_1 - is_1)(\mathrm{d}x_1)_a + (a_2 - is_2)(\mathrm{d}x_2)_a \right] \,, \tag{5.48}$$

where now the parameters labeling the connection are precisely (a_i, s_i) as defined above.

Denoting $w_j := a_j - is_j$ so that $z_j = \exp(iw_j)$, we can define "the trace of the holonomy" or *Wilson loop* along $[\gamma] \simeq (n, m)$ has the form,

$$\begin{aligned} \mathrm{T}(n, m) := \frac{1}{2} \mathrm{Tr} H(\mathbf{A}, [\gamma]) &= \frac{1}{2} \left(z_1^n z_2^m + z_1^{-n} z_2^{-m} \right) \\ &= \cos\left(n w_1 + m w_2 \right) \,. \end{aligned} \tag{5.49}$$

Note that as we would expect, these functions $\mathrm{T}(n, m)$ are well defined on $\hat{\Gamma}$ since they are invariant under the inversion $(z_1, z_2) \to (z_1^{-1}, z_2^{-1})$. We have chosen to work with this functions for simplicity even though they do *not* correspond to traces of the $\mathrm{SO}(3, \mathbb{C})$ holonomies (in the defining representation).

Finally we have to find an expression for the induced symplectic structure $\hat{\Omega}$ in terms of our coordinates. We can see that equation (5.17) together with (5.48) yield,

$$\hat{\Omega} = -\frac{1}{G\lambda}\left(\mathrm{d}a_1 \wedge \mathrm{d}s_2 + \mathrm{d}s_1 \wedge \mathrm{d}a_2\right) , \tag{5.50}$$

which in terms of the coordinates w_j reads,

$$\hat{\Omega} = -\frac{i}{2G\lambda}\left(\mathrm{d}w_1 \wedge \mathrm{d}w_2 - \mathrm{d}\overline{w}_1 \wedge \mathrm{d}\overline{w}_2\right) . \tag{5.51}$$

Note that the symplectic structure contains the parameter λ. Just as in the case of the full phase space, the symplectic form remains real, but is just rewritten in terms of complex coordinates.

In the reminder of this section we shall recall the reduced phase space description for Euclidean $2+1$ gravity without cosmological constant, and compare it to our case[3][67]. We have the same Gauss constraint (5.20), but now the other constraint reads

$$\tilde{\eta}^{ab} F^I_{ab} = 0. \tag{5.52}$$

This means that the SO(3) connection A^I_a is flat. The reduced phase space $\hat{\Gamma}_{\Lambda=0}$ is now the cotangent bundle over the reduced configuration space $\hat{C}$ consisting of flat SO(3) connection modulo gauge transformations. It is possible to characterize $\hat{C}$ just as we did previously, by considering the holonomies along the generators of $\Pi_1(\Sigma)$, modulo conjugation. We can parametrize this space by two angles (θ_1, θ_2) with the identification $(\theta_1,\theta_2) \sim (-\theta_1, -\theta_2)$. We can also give coordinates (p_i, p_2) for the "momenta". We arrive therefore to a space that is mathematically identical to the space $\hat{\Gamma}$ of the previous subsection, that is,

$$\hat{\Gamma}_{\Lambda=0} \simeq \frac{(\mathrm{T}^2 \times \mathbb{R}^2)}{\mathbb{Z}_2} , \tag{5.53}$$

with the above mentioned identifications.

There are however, important differences when we consider the 'physics" of both theories. Let us consider for instance the functions $\cos(n\theta_1 + m\theta_2)$. They are well defined functions on both reduced phase spaces, but it is only in the $\Lambda = 0$ case that they have a simple geometrical interpretation in terms of the full phase space. Recall that in that case the Wilson loops of the SO(3) connection A^I_a are full observables in the sense of Dirac (they Poisson commute with all constraints) and therefore descend to $\hat{\Gamma}_{\Lambda=0}$. They correspond precisely to the functions we are considering,

$$\mathrm{T}(n,m)_{\Lambda=0} = \cos(n\theta_1 + m\theta_2). \tag{5.54}$$

On the other hand, when the cosmological constant is not zero, Wilson loops of the SO(3) connection are no longer observables, and therefore there is no function on $\hat{\Gamma}$ corresponding to them. Reciprocally, the functions $\cos(n\theta_1 + m\theta_2)$ *are* well defined observables on $\hat{\Gamma}$, but they do not have a simple geometric interpretation in terms of the coordinates (A^I_a, e^I_a).

We shall see in next section, when we consider the reduced phase space quantization of the theory, that there is a similar discussion for the quantum theories of both systems.

[3]The discussion that follows also applies to the Lorentzian case in the "time-like sector"

5.4 Quantization

In this section we will quantize the system that we studied in Sec. 5.3. Recall that there are in general two main approaches to the quantization of constrained systems. They differ, roughly speaking, on the order in which one reduces and quantizes the system. In the first approach, one reduces the classical system from Γ to the reduced phase space $\hat{\Gamma}$ and then quantizes the reduced system. This method is called *reduced phase space quantization* (RPSQ). Within the second approach, one quantizes the total phase space Γ and then imposes the constraints, that are now quantum operators, as conditions on the physically admissible wave functions. This is known as *Dirac quantization* (DQ).

Three dimensional gravity has the peculiar property that the only physical degrees of freedom are topological, that is, one starts with a field theory with infinite number of degrees of freedom, and after reduction the system has only a *finite* number of degrees of freedom. This property makes the first approach to quantization "trivial" in the sense that we are left with a finite dimensional phase space. In practice, we are doing quantum mechanics. On the other hand, on the Dirac side, we have a field theoretic quantization with a set of constraints to be imposed. One expects the solutions of the quantum constraints to be in correspondence with the quantum states from RPSQ.

In this section we shall consider in detail the quantization of the reduced phase space of Sec. 5.3.

Recall that the reduced phase space $\hat{\Gamma}$ can be seen as the cotangent bundle over a torus T^2, with certain identifications. We can give it complex coordinates (z_1, z_2) where $z_i \in \mathbb{C}^*$ and we have the identification $(z_1, z_2) \sim (1/z_1, 1/z_2)$. We shall construct a quantum theory on $\hat{\Gamma}$ seen as a complex space and define holomorphic functions thereon, in close analogy with the Segal-Bargmann quantization of the harmonic oscillator.

We want to construct the quantum theory in a way that the geometrically motivated observables such as the Wilson loops for the generators $[\gamma_i]$ of the fundamental group are well defined operators. The strategy is to follow the algebraic approach outlined in [35, 39] (see also Chapter 3. The first step it to choose a set $\mathcal{S}$ of elementary observables in $\hat{\Gamma}$. The elements of $\mathcal{S}$ should be enough in order to generate any function on $\hat{\Gamma}$ by taking linear combinations of products on them. In other words, $\mathcal{S}$ should be complete. In order to have a uniquely defined quantum operator for each element of the set $\mathcal{S}$, we should ask it to be closed under Poisson bracket. Let us choose the following coordinate functions,

$$Z_1 := \exp[s_1 - ia_2] \qquad ; \qquad Z_2 := \exp[s_2 + ia_1]\,. \tag{5.55}$$

Note that they do *not* correspond to the coordinates (z_i, z_2) we had defined before. We can now choose as our set of elementary observables,

$$\mathcal{S} = \left\{Z_1, Z_2, Z_1^{-1}, Z_2^{-1}, s_1, s_2, 1\right\}\,. \tag{5.56}$$

The symplectic structure (5.17) defines the following Poisson brackets,

$$\{Z_1, Z_2\} = 0 \tag{5.57}$$

$$\{Z_i, s_j\} = iG\lambda\delta_{ij}Z_i \tag{5.58}$$

$$\left\{Z_i^{-1}, s_j\right\} = -iG\lambda\delta_{ij}Z_i^{-1}\,. \tag{5.59}$$

Therefore it is closed under Poisson bracket. Note that it is, however, not closed under complex conjugation. That is, the functions $\overline{Z}_i$ are *not* in $\mathcal{S}$. This is so because the Poisson bracket $\{\overline{Z}_1, Z_1\} = 2i\lambda\overline{Z}_1 Z_1$ takes us out of $\mathcal{S}$. If we were to include them, then we would have to include the function $\overline{Z}_1 Z_1$ in $\mathcal{S}$, but then $\mathcal{S}$ would have to contain almost all functions on $\hat{\Gamma}$!

The next step is to construct the abstract algebra $\mathcal{A}$ of operators and impose the commutation relations,

$$\begin{aligned} \left[\hat{Z}_i, \hat{s_j}\right] &= -G\hbar\lambda\delta_{ij}\hat{Z}_j\,, \\ \left[\widehat{Z_i^{-1}}, \hat{s_j}\right] &= G\hbar\lambda\delta_{ij}\widehat{Z_j^{-1}}\,, \\ \left[\widehat{Z_j^{-1}}, \hat{Z}_k\right] &= 0\,. \end{aligned} \tag{5.60}$$

There is only one over-completeness on our algebra $\mathcal{A}$, namely we should ask that $\widehat{Z_j^{-1}} = (\hat{Z}_j)^{-1}$.

The next step it to represent the algebra $\mathcal{A}$ as operators in a vector space V. We will choose as the vector space V the space of holomorphic functions $\Psi(Z_1, Z_2)$ on $\mathbb{C}^* \times \mathbb{C}^*$. Let us represent the operators,

$$\hat{s_j}\,\Psi(Z_1, Z_2) := G\hbar\lambda Z_j\,\frac{\partial\Psi(Z_1, Z_2)}{\partial Z_j}\,, \tag{5.61}$$

$$\hat{Z}_j\,\Psi(Z_1, Z_2) := Z_j\,\Psi(Z_1, Z_2)\,, \tag{5.62}$$

$$\widehat{Z_j^{-1}}\Psi(Z_1, Z_2) := Z_j^{-1}\,\Psi(Z_1, Z_2)\,. \tag{5.63}$$

Finally, we would like to define an inner product in the vector space V in order to construct the Hilbert space $\mathcal{H}$. The strategy at this point is to use the complex conjugate (or $*$-relations) on the algebra $\mathcal{A}$ and impose them as Hermiticity conditions. The problem we have with our construction is that we do not have $*$-relations! The set $\mathcal{S}$ is not closed under complex conjugation. We only have one reality condition, namely that s_i should be real. We should then impose that $\hat{s_i}$ be Hermitian, that is $\langle\Psi|\hat{s_i}\Phi\rangle = \langle\hat{s_i}\Psi|\Phi\rangle$. In other words, if we define the inner product to be of the form,

$$\langle\Psi|\Phi\rangle := \int_{\mathbb{C}^*\times\mathbb{C}^*} \mathrm{d}\mu(Z_1, Z_2, \bar{Z}_1, \bar{Z}_2)\,\overline{\Psi}\,\Phi, \tag{5.64}$$

then the condition that the operators $\hat{s_i}$ be Hermitian implies that the measure μ is of the form $\mu = \mu\left(|Z_1|^2 + |Z_2|^2\right)$

We would like to have a condition on $\hat{Z_i}^\dagger$ in order to impose it on μ. The difficulty, already mentioned above, is that we do not have an explicit expression for the operator $\hat{\bar{Z}}_i$ so that we can ask $\hat{\bar{Z}}_i = \hat{Z_i}^\dagger$. It seems that we have a problem!

One possibility would be to look at the classical expression for $\bar{Z}_i$ in terms of the elements of $\mathcal{S}$ and try to define the "corresponding quantum operator". In the case of the harmonic oscillator this procedure works because the dependence is linear so there is no ambiguity in the definition of the corresponding operator[4].

[4]Recall that in that case we have $\mathcal{S} = \{z = p - iq, q, 1\}$, so $\bar{z} = z + 2iq$, and therefore we can define $\hat{\bar{z}} := \hat{z} + 2i\hat{q}$.

Classically we have,

$$\bar{Z}_i = e^{2s_i} Z_i^{-1}, \tag{5.65}$$

so one would like to define the corresponding operator. The problem is that $e^{\hat{s}_i}$ and $\widehat{Z_i^{-1}}$ do not commute. In fact, since $e^{\hat{s}_i}$ is a very "complicated" function of s_i there is a great ambiguity in the factor ordering in order to define $\hat{\bar{Z}}_i$.

One possible strategy in order to resolve this difficulty would be to *postulate* the inner product and then *define* the Hermitian conjugates to be $\hat{\bar{Z}}_i := \hat{Z}_i^{\dagger}$. This apparent freedom is, however, severely limited by the fact that the operators $\hat{\bar{Z}}_i$ should go, in the classical limit, to the functions $\bar{Z}_i$. We need some extra input in order to choose the "correct" inner product.

Recall that a Holomorphic quantum theory for the harmonic oscillator can be constructed using a transform from the Hilbert space of square integrable functions on the real line to the space of holomorphic functions on $\mathbb{C}$ with a Gaussian measure. Luckily, there exists a similar transform relevant for our system. It is a generalization of the Segal-Bargmann transform and maps the Hilbert space of normalized functions on a compact, connected, Lie-group K with respect to the (left and right invariant) Haar measure to the Hilbert space of Holomorphic functions on the complexification of the Lie-group [68]. The transform gives also a prescription for the measure in the complexified group and provides an isometric isomorphism of Hilbert spaces. We can use this "Hall transform" to our problem since the phase space $\hat{\Gamma}$ can be seen as the cotangent bundle over $\mathrm{T}^2 = S^1 \times S^1$ that is a compact connected Lie-group and we can define a Hilbert space on it, namely the space of square integrable functions with respect to the Homogeneous measure on T^2.

We will give basic facts about the Hall transform for U(1). For details see Appendix D. Let us denote by K the compact Lie-group and $x \in K$ an element of it. An element of G, the complexified group will be denoted by g. Let us now define the mapping $C_t : L^2(K, \mathrm{d}x) \to \mathcal{H}(G)$ as follows,

$$C_t(\phi)(g) = \int_K \phi(x)\rho_t(x^{-1}g)\,\mathrm{d}x \qquad \phi \in L^2(K, \mathrm{d}x),\ g \in G \tag{5.66}$$

where $\rho_t(x)$ is the solution to the heat equation in K (at "time t") and $\rho_t(x^{-1}g)$ is its analytic continuation. The measure $\mathrm{d}x$ is the Haar measure. The transform is an isometric isomorphism onto the Hilbert space of holomorphic functions in $L^2(G, \nu_t)$. The measure ν_t is constructed from the heat kernel measure μ_t^c on G by averaging over the left action of K.

If we denote the transformed wave-function by $\phi^c := C\phi$, the "transformed operators" will be $\hat{A}^c := C\hat{A}C^{-1}$. Note also that we have the freedom in the choice of the parameter t. For each value of $t > 0$ we have a well defined holomorphic quantum theory. Let $x \in K$ be of the form $x = e^{i\theta}$. We can define an element of the complex group $G = \mathbb{C}^*$, by exponentiating an element of the complexified Lie algebra. Let us identify $\theta \in \mathbb{R}$ with the real Lie algebra of U(1) and define the complexification as $\theta \to \tilde{\theta} := \theta - ip$, in such a way that if $x = e^{i\theta}$ then $g \in \mathbb{C}^*$ will be $g := e^{i\tilde{\theta}} = e^{p+i\theta}$.

We can write any function $\phi(x) \in L^2(S^1, \mathrm{d}\theta)$ as,

$$\phi(x) = \sum_m a_m\, x^m = \sum_m a_m e^{im\theta}\,, \tag{5.67}$$

with the condition $\sum_m |a_m|^2 < \infty$. The transformed function will be then,

$$\phi_t^c(g) := C_t(\phi) = \sum_m a_m \, e^{-\frac{t}{2}m^2} g^m \,. \tag{5.68}$$

The averaged heat kernel measure is,

$$\nu_t = \frac{1}{\sqrt{\pi t}} \, e^{-\frac{p^2}{t}} \,, \tag{5.69}$$

so that the inner product in $L^2(\mathbb{C}^*, \nu_t)$ is,

$$\langle\psi|\phi\rangle = \int_{\mathbb{C}^*} \mathrm{d}p \, \mathrm{d}\theta \exp\left[-\tfrac{p^2}{t}\right] \overline{\psi(g)} \, \phi(g) \,. \tag{5.70}$$

We can very easily generalize the construction from S^1 to our case, namely for $S^1 \times S^1$. The transformation in this case takes functions $\phi(a_1, a_2)$ to functions of our new coordinates (Z_1, Z_2), so that if we have a function

$$\Phi(a_1, a_2) = \sum_{m,n} a_{mn} e^{(ima_1 + ina_2)} \,. \tag{5.71}$$

Then the transformed function will be

$$\Phi_t^c(Z_1, Z_2) := C_t(\Phi) = \sum_{m,n} a_{mn} e^{-\frac{t}{2}(m^2+n^2)} Z_2^m Z_1^n \,. \tag{5.72}$$

Note that we have a well defined transform for each value of the parameter $t > 0$. In order to make contact with the algebraic program, we should ask that the transformed operators corresponding to $\hat{s_i}$ are given by the expression (5.62), that is, $C \hat{s_i} \, C^{-1} = G\hbar\lambda Z_i \frac{\partial}{\partial Z_i}$. This condition selects the value of the parameter $t = G\hbar\lambda$[5].

Using the transform (5.72), we can find the expression for the Hermitian adjoint of $\hat{Z}_j$, namely,

$$\hat{Z_j}^{\dagger} = e^{G\hbar\lambda} e^{2\hat{s_j}} \widehat{Z_j^{-1}} \,. \tag{5.73}$$

Note that this is consistent with the reality conditions in the sense the in the limit $\hbar \to 0$, the operator (5.73) goes to the expression for the complex conjugate $\bar{Z}_i = e^{2s_i} Z_i^{-1}$. It is not very surprising that the inner product given by the Hall transform realizes the Hermiticity-reality conditions (5.73). The inner product is given by,

$$\langle\Phi|\Psi\rangle := \frac{1}{\pi G\hbar\lambda} \int_{\mathbb{C}^* \times \mathbb{C}^*} \mathrm{d}s_1 \mathrm{d}s_2 \mathrm{d}a_1 \mathrm{d}a_2 \, \exp\left[-\frac{1}{G\hbar\lambda}(s_1^2 + s_2^2)\right] \overline{\Phi} \, \Psi \,. \tag{5.74}$$

We can now give a characterization of the Hilbert space $\mathcal{H}$. Note that the monomials of the form $Z_1^m Z_2^n$ are orthogonal, and satisfy,

$$\langle Z_1^m Z_2^n | Z_1^m Z_2^n \rangle = e^{G\hbar\lambda(m^2+n^2)} \,. \tag{5.75}$$

[5]Note that in the units we are using, where $c = 1$, the parameter $t = G\hbar\lambda$ is dimension-less.

Therefore, an orthonormal basis for the space $\mathcal{H}$ is given by

$$\Phi_{mn} := e^{-\frac{G\hbar\lambda}{2}(m^2+n^2)} Z_1^m Z_2^n \,. \tag{5.76}$$

A general function on the Hilbert space will be given by the series,

$$\Psi(Z_i) = \sum_{m,n=-\infty}^{\infty} A_{mn}\Phi_{nm} \,. \tag{5.77}$$

So far, we have ignored the fact that in the reduced phase space $\hat{\Gamma}$ we should identify the points $(Z_1, Z_2) \sim (Z_1^{-1}, Z_2^{-1})$. Instead of defining the "physical" phase space to be the quotient of $\hat{\Gamma}$ under the equivalence relation, we have kept the total phase space and defined holomorphic functions thereon. We shall now ask that the *physical* Hilbert space $\mathcal{H}_{\rm phys}$ be given by wave-functions that are invariant under the identification. That is,

$$\mathcal{H}_{\rm phys} := \left\{\Psi \in \mathcal{H}/\Psi(Z_1, Z_2) = \Psi(Z_1^{-1}, Z_2^{-1})\right\} \,. \tag{5.78}$$

Therefore, in terms of the expansion (5.77), physical wave-functions will be those that are "even", that is, $A_{mn} = A_{(-m)(-n)}$. The vacuum vector $|0\rangle$ will be represented by $\Phi_0(Z) = \langle Z|0\rangle = 1$. Note that the operators $\hat{Z}_i$ are not *physical* operators since they throw us out of the physical Hilbert space. We need to consider even combinations like $\hat{Z}_i + \widehat{Z_i^{-1}}$.

As we mentioned before, we would like to have a quantization for $\hat{\Gamma}$ such that the traces of Holonomies be well defined operators. This is an essential requirement in the quantization of $3+1$ based on loop variables [35]. A fundamental difference with the $3+1$ case is that the phase space $\hat{\Gamma}$ is finite dimensional, so we would like to choose from the *over-complete* set of Wilson loops a minimum set that "covers" in a sense the phase space. An appropriate choice of such coordinates has been the subject of extensive study [69, 70]. For the case of the torus a standard choice is the set $\{\mathrm{T}(1,0), \mathrm{T}(0,1), \mathrm{T}(1,1)\}$. Another important difference with the $3+1$ case is that now the Wilson loops are phase space functions are therefore have a non-trivial Poisson algebra amongst themselves (whereas in the $3+1$ case and $2+1$ with $\Lambda = 0$ they Poisson commute). In terms of our fundamental function in $\mathcal{S}$ they have the form,

$$\mathrm{T}(1,0) = \frac{1}{2}\left(e^{s_1}e^{-s_2}Z_2 + e^{-s_1}e^{s_2}Z_2^{-1}\right) \,, \tag{5.79}$$

$$\mathrm{T}(0,1) = \frac{1}{2}\left(e^{s_2}e^{s_1}Z_1^{-1} + e^{-s_2}e^{-s_1}Z_1\right) \,, \tag{5.80}$$

$$\mathrm{T}(1,1) = \frac{1}{2}\left(\overline{Z_1}Z_2 + \overline{Z_1^{-1}}Z_2^{-1}\right) \,. \tag{5.81}$$

These functions turn out to have exact operators in the space $\mathcal{H}_{\rm phys}$ *without* any factor ordering ambiguity,

$$\hat{\mathrm{T}}(1,0) = \frac{1}{2}\left(e^{-\frac{G\hbar\lambda}{2}}e^{\hat{s}_1}\hat{Z}_2e^{-\hat{s}_2} + e^{\frac{G\hbar\lambda}{2}}e^{-\hat{s}_1}e^{\hat{s}_2}\widehat{Z_2^{-1}}\right) \,, \tag{5.82}$$

$$\hat{\mathrm{T}}(0,1) = \frac{1}{2}\left(e^{-\frac{G\hbar\lambda}{2}}\widehat{Z_1^{-1}}e^{\hat{s}_2}e^{\hat{s}_1} + e^{\frac{G\hbar\lambda}{2}}e^{-\hat{s}_1}\hat{Z}_1e^{-\hat{s}_2}\right) \,, \tag{5.83}$$

$$\hat{\mathrm{T}}(1,1) = \frac{1}{2}\left(e^{-G\hbar\lambda}\hat{Z}_2e^{-2\hat{s}_1}\widehat{Z_1^{-1}} + e^{G\hbar\lambda}\widehat{Z_2^{-1}}\hat{Z}_1e^{2\hat{s}_1}\right) \,. \tag{5.84}$$

In summary, we started with a closed set of functions $\mathcal{S}$ to be quantized. We found a representation on the space of holomorphic functions and found the inner product using the reality conditions together with the Hall transform. Then, we considered the subspace of physical states and found that the Wilson loops along the simplest loops are well defined operators thereon.

Finally, note that we did *not* try to quantize and find representations of the algebra generated by the Wilson loops $\mathrm{T}(n,m)$, in contrast to the program of Nelson and collaborators [70, 71, 72].

5.5 Relation with the Space-Time Picture

Recall from Sec. 5.3 that we have one phase space $\hat{\Gamma}$ for both the Lorentzian theory with $\Lambda > 0$ and Euclidean theory with $\Lambda < 0$. The goal of this section is to show that given a point q on $\hat{\Gamma}$ there is (at least) one Lorentzian space-time with constant *negative* curvature and a Euclidean space-time with *positive* curvature. Let q have coordinates (a_1, a_2, s_1, s_2). Let us recall that the characterization of the reduced phase space came by considering the holonomies along the two generators $[\gamma_1]$ and $[\gamma_2]$ of $\Pi_1(\mathrm{T}^2)$ of an $\mathrm{SO}(3,\mathbb{C})$ connection. To each generator we associate two parameters,

$$[\gamma_j] \leftrightarrow z_j = \exp[i(a_j - is_j)]\,, \tag{5.85}$$

where $j = 1, 2$. Therefore, we should expect intuitively that in the space-time M_q associated to point q, the parameters (a_j, s_j) will have the information about the holonomy along γ_j.

We know that the Lie-algebra $sl(2,\mathbb{C})$ is isomorphic to $so(3,\mathbb{C})$ and that it can be identified with the self-dual part of the complexified $so(3,1)$. The holonomies are however, elements of the group, not the Lie algebra. To understand the relation we shall use a peculiar feature of the Lorentz Lie-algebra in 4 dimensions. If we have a generator of, say, a rotation in the (X, Y) plane, then the dual element is a generator of boosts in the (Z, T) plane. These two generators commute, so the finite transformations that they generate, that is, the rotation and the boost, also commute. When we were constructing the reduced phase space we had to take elements of the Abelian subgroup up to conjugation. The fact that we are moding out by a conjugation means that any "direction" for the rotation (or in other words, the invariant plane) characterizing the $\mathrm{SO}(3,1)$ element of the holonomy is "lost". The only "gauge invariant" information are the parameters of the "rotation" and the parameter of the "boost" (in the dual direction).

We can conclude from this discussion that we expect the parameters (a_1, a_2) to correspond to some "rotation parameters" in the holonomies, and the numbers (s_1, s_2) should correspond to "boost parameters". The question that arises now is the following. If we are going to construct a 3-dimensional space-time, that has as its natural "gauge" group the Lorentz group in 3-dimensions $\mathrm{SO}(2,1)$ (or $\mathrm{SO}(3)$), how does the 4-dimensional Lorentz group come into the picture? How are we going to associate an $\mathrm{SO}(3,\mathbb{C})$ or $\mathrm{SO}(3,1)$ element to a space-like holonomy in 3-dimensions? We shall answer this question in what follows.

Finally, we know that the parameters (a_1, a_2, s_1, s_2) are "good" coordinates on $\hat{\Gamma}$ in the sense that they come from gauge invariant quantities on the full phase space, and therefore they will correspond to "constants of the motion" for the space-time M_q.

5.5.1 Lorentzian Space-time

With all this information at hand, we shall construct in this subsection the Lorentzian space-times satisfying all the properties mentioned above. The general strategy will be as follows. We know that there exists a "canonical" 3-dimensional Lorentzian space-time of constant positive curvature, namely, the 3-dimensional De Sitter space dS_3^L. There is a natural way of defining it as an embedding in a fixed 4-dimensional Minkowski space-time M^4, with metric diag$(-+++)$. The strategy is to construct our space-time M_q^L with topology $\mathrm{T}^2\times\mathbb{R}$ as embedded in dS_3^L by taking some proper identifications. Let us recall that De Sitter space-time is given by the points in Minkowski space-time M^4, with coordinates (T,X,Y,Z), satisfying the condition,

$$-T^2+X^2+Y^2+Z^2=\frac{1}{|\Lambda|}\,. \tag{5.86}$$

Note that slices of constant coordinate $T=k$ correspond to 2-spheres of radius $r^2=\frac{1}{|\Lambda|}+k^2$. Let us define the following embedding from a 3-dimensional space with coordinates (t,θ,ϕ) into dS_3^L as follows [73],

$$T=\frac{1}{\lambda}\sinh t\cosh\theta \qquad ; \qquad X=\frac{1}{\lambda}\sinh t\sinh\theta \tag{5.87}$$

$$Y=\frac{1}{\lambda}\cosh t\cos\phi \qquad ; \qquad Z=\frac{1}{\lambda}\cosh t\sin\phi \tag{5.88}$$

where $\theta\in(-\infty,\infty)$, $t\in(0,\infty)$ and $\phi\in[0,2\pi)$. Note that this mapping does not cover all of De Sitter space, but only a region of it with topology $S^1\times\mathbb{R}^2$. It is in this region of dS_3^L that we will construct our space-time M_q^L.

This coordinate system is particularly suited for constructing the space-time M_q^L in a $2+1$ fashion in the "York gauge", that is, with hyper-surfaces of constant extrinsic curvature. Such slices will be given by points of constant coordinate t, so for $t=\tau$=constant, we have,

$$T^2-X^2=\frac{1}{\Lambda}\sinh^2\tau(\cosh^2\theta-\sinh^2\theta)=\frac{1}{\Lambda}\sinh^2\tau\,, \tag{5.89}$$

therefore,

$$Y^2+Z^2=\frac{1}{\Lambda}(1+\sinh^2\tau)=\frac{1}{\Lambda}\cosh^2\tau\,. \tag{5.90}$$

That is, in the (T,X) plane the region lies in the hyperboloid with parameter $\frac{1}{\Lambda}\sinh^2\tau$. Note that as $\tau\to 0$ the hyperboloid becomes degenerate. This is a common feature of the York slicing in $2+1$ gravity. In the (Y,Z) plane, each point of the hyperboloid is a circle with radius $\frac{1}{\lambda}\cosh\tau$. Therefore, slices of constant $t=\tau$ have the topology $S^1\times\mathbb{R}$ and are coordinatized by (θ,ϕ). A key observation in this construction is to note that the coordinate θ can be regarded as an affine parameter for boosts in the (T,X) plane, and ϕ as a parameter for rotations in the (Y,Z) plane. As elements of the group $\mathrm{SO}(3,1)$ acting on the 4-d Minkowski space, they are dual to each other and therefore, commute.

We have now the stage ready for constructing the space-time M_q^L with topology $\mathrm{T}^2\times\mathbb{R}$. The strategy is to find four points and four links joining them in the "plane" (θ,ϕ) that are to be identified in order to form the torus. The condition that we have to satisfy is that the "parallel transport" or holonomy along the generators should have as "labels" the coordinates (s_i, a_i) of the point q in our phase space $\hat{\Gamma}$. We knew intuitively that these parameters (s_i, a_i) were to be associated to a boost and a rotation, so now the construction procedure is obvious. Lets take the point $(\theta = 0, \phi = 0)$ as our origin, and choose as the other vertices of the torus the points with coordinates $v_1 = (s_1, a_1)$ and $v_2 = (s_2, a_2)$. How are we going to choose the "paths" from the origin to the vertices? Let us give a simple prescription: take the straight lines from $(0, 0)$ to the points $v_1 = (s_1, a_1)$ and $v_2 = (s_2, a_2)$ and identify the points (v_1, v_2) with the origin. The straight lines are therefore the generators (γ_1, γ_2) of the torus. We then close the "parallelogram" in the obvious way.

Have we achieved what we wanted? It is true that we have used the parameters (s_i, a_i) in our construction of the torus, but in what sense can we say that the holonomies along γ_i should be related to (s_i, a_i)? We want to argue that if we parallel transport a 4-dimensional vector from the origin to say, the vertex v_1, the original vector and the parallel transported one will differ by a rotation in the (Y, Z) plane with parameter a_1 composed with a boost in the (T, X) plane with parameter s_1. A similar statement would be true for vertex v_2. How can we say that the parallel transported vector will be different that the original one when we are using the flat connection in M^4? It is true that when we parallel transport a 4-vector from the origin to v_1 it does not "change" (its components in the (T, X, Y, Z) coordinate system remain unchanged). But, since we are also identifying v_1 with the origin, we are at that step "push-forwarding" the vector back to the origin. We are using an element $U(a_1, s_1)$ of $SO(3, 1)$ for that, and since it is a linear map, the derivative $U_*(a_1, s_1)$ that sends tangent vector to tangent vectors coincides with $U(a_1, s_1)$. Therefore the "parallel transported" vector has as its holonomy precisely the group element in SO(3, 1) with parameters (s_1, a_1).

The metric induced on dS_3^L by the mapping (5.87)-(5.88) is given by,

$$\mathrm{d}s^2 = \frac{1}{\Lambda}(-\mathrm{d}t^2 + \sinh^2 t\,\mathrm{d}\theta^2 + \cosh^2 t\,\mathrm{d}\phi^2)\,. \tag{5.91}$$

The next step is to define the coordinates (x_1, x_2) along the generators in such a way that they take values from 0 to 1. Thus, we want the vertices to have coordinates $v_1 \sim (1, 0)$, $v_2 \sim (0, 1)$. We are looking for constants $\alpha, \beta, \gamma, \delta$ that define the transformation,

$$x_1 = \alpha\phi + \beta\theta \qquad ; \qquad x_2 = \gamma\phi + \delta\theta\,, \tag{5.92}$$

such that,

$$1 = \alpha a_1 + \beta s_1 \qquad ; \qquad 0 = \gamma a_1 + \delta s_1$$

and

$$0 = \alpha a_2 + \beta s_2 \qquad ; \qquad 1 = \gamma a_2 + \delta s_2\,,$$

The mapping is clearly,

$$x_1 = \frac{1}{s_1 a_2 - s_2 a_1}(-s_2\phi + a_2\theta)\,, \tag{5.93}$$

$$x_2 = \frac{1}{s_1 a_2 - s_2 a_1}(s_1\phi - a_1\theta)\,. \tag{5.94}$$

The inverse map is

$$\phi = a_1 x_1 + a_2 x_2 \qquad ; \qquad \theta = s_1 x_1 + s_2 x_2 . \tag{5.95}$$

Note that this change of coordinates will only work when $s_1 a_2 - s_2 a_1 \neq 0$. The condition $s_1 a_2 = s_2 a_1$ means that both vertices v_1, v_2 lie on the same line in the (θ, ϕ) plane, so there is no "interior region" to define the torus, or in other words, it is degenerate.

The line element (5.91) takes now the form,

$$\begin{aligned} ds^2 = \frac{1}{\Lambda}[-dt^2 \; &+ \; (s_1^2 \sinh^2 t + a_1^2 \cosh^2 t)\, dx_1{}^2 + 2\,(s_1 s_2 \sinh^2 t + a_1 a_2 \cosh^2 t)\, dx_1\, dx_2 \\ &+ \; (s_2^2 \sinh^2 t + a_2^2 \cosh^2 t)\, dx_2{}^2] \end{aligned} \tag{5.96}$$

We want to make contact with the line element in the York gauge coming from the ADM description, that is, we would like the line element to have the form,

$$ds^2 = -\frac{1}{\Lambda} dt^2 + \frac{e^{2l}}{m_2} \left[(dx_1 + m_1 dx_2)^2 + (m_2 dx_2)^2 \right] \tag{5.97}$$

where m_1, m_2 are the modulus of the torus. The transformation equations are given by,

$$e^{2l} = \frac{\sinh t \cosh t}{\Lambda}(a_1 s_2 - s_1 a_2)\,, \tag{5.98}$$

$$m_1 = \frac{a_1 a_2 + s_1 s_2 \coth^{-2} t}{a_1^2 + s_1^2 \coth^{-2} t}\,, \tag{5.99}$$

$$m_2 = \frac{(a_1 s_2 - s_1 a_2) \coth^{-1} t}{a_1^2 + s_1^2 \coth^{-2} t}\,. \tag{5.100}$$

5.5.2 Euclidean Space-time

As we mentioned in last section, the reduced phase space $\hat{\Gamma}$ is the same for the Lorentzian and Euclidean signatures, so given a point q in $\hat{\Gamma}$ we should be able to produce space-times with both signatures. In the last part we constructed a Lorentzian space-time by making some identifications in De Sitter space as embedded in 4-d Minkowski M^4. Recall that De Sitter space-time was the "hyperboloid" with signature $(- + +)$. There is another hyperboloid in M^4, but with signature $(+ + +)$, namely the point with coordinates (T, X, Y, Z) satisfying

$$-T^2 + X^2 + Y^2 + Z^2 = -\frac{1}{|\Lambda|}\,. \tag{5.101}$$

In fact, this region is disconnected, so we will concentrate only on one of the components, namely the one with $T > 0$. Just as in the previous case, we can define coordinates (t, θ, ϕ) and embedding,

$$T = \frac{1}{\lambda} \cosh t \cosh \theta \qquad ; \qquad X = \frac{1}{\lambda} \cosh t \sinh \theta \tag{5.102}$$

$$Y = \frac{1}{\lambda} \sinh t \cos \phi \qquad ; \qquad Z = \frac{1}{\lambda} \sinh t \sin \phi \tag{5.103}$$

For $T=k$=constant, we have

$$X^2+Y^2+Z^2=k^2-\frac{1}{|\Lambda|},$$

so that $T>\frac{1}{|\Lambda|}$.

Hyper-surfaces of constant t shall again provide space-times in the York gauge. Lets us take $t=\tau$, then

$$-T^2+X^2=-\frac{1}{|\Lambda|}\cosh^2\tau\,,$$

and

$$Y^2+Z^2=\frac{1}{|\Lambda|}\sinh^2\tau.$$

The induced metric on the 3-dimensional space-time dS_3^E is then,

$$\mathrm{d}s^2=\frac{1}{|\Lambda|}\left[\mathrm{d}t^2+\cosh^2 t\mathrm{d}\theta^2+\sinh^2 t\mathrm{d}\phi^2\right]\,. \tag{5.104}$$

The construction of the space-time M_q^E with topology $\mathrm{T}^2\times\mathbb{R}$ follows the same steps as the Lorentzian one, so we shall not repeat them here. We shall only comment on the fact that the same element $U(a_1,s_i)$ of $\mathrm{SO}(3,1)$ that we used in order to identify the vertices to the origin in dS_3^L, will be used for the Euclidean construction on dS_3^E. Recall that any element of $\mathrm{SO}(3,1)$ is an isometry of both dS_3^L and dS_3^E.

To conclude this section, let us compare the geometries of the $t=\tau$ slices for both space-times. Consider the proper length L_i along the γ_i generator. We have,

$$\mathrm{L}_i = \sqrt{a_i^2\cosh^2\tau+s_i^2\sinh^2\tau}\qquad \text{Lorentzian,} \tag{5.105}$$

$$\mathrm{L}_i = \sqrt{s_i^2\cosh^2\tau+a_i^2\sinh^2\tau}\qquad \text{Euclidean.} \tag{5.106}$$

Note that there is a "duality" between the two cases under the permutation $a_i\leftrightarrow s_i$.

In this section we have seen that the reduced phase space $\hat{\Gamma}$ of the connection-dynamics formulation of Sec. 5.3, in the case where the hyper-surface is a genus one surface can be parametrized by two complex numbers (z_1,z_2) or alternatively by four real numbers (a_1,a_2,s_1,s_2). For each point q on $\hat{\Gamma}$ we constructed a Lorentzian space-time M_q^L and a Euclidean one M_q^E that could be embedded in Minkowski space M^4. We saw that, associated with a point in $\hat{\Gamma}$, there is countable number (parametrized by two integers) of space-times of both signatures that can be obtained by performing a large gauge transformation. The failure of $\hat{\Gamma}$ to distinguish these sectors of the theory can be traced back to the fact that holonomies are invariant under all gauge transformations. This space-times can not be simply embedded into M^4.

5.6 Discussion

We saw in Sec. 5.4 that we could construct a holomorphic quantum theory of the reduced phase space of Sec. 5.3. We could successfully incorporate the reality conditions in such a way that real observables are self adjoint operators of the theory. Furthermore, we could define the Wilson loops of the basic loops as unambiguous operators without the need to introduce any factor ordering choice. One might question the relevance of the holomorphic quantization given that the phase space has an alternative real coordinatization in terms of the a_i and s_i parameters. We could, after all, have defined the quantum theory as functions of the real parameters a_i. From a geometric point of view, in terms of a space-time interpretation, the discussion of Sec. 5.5 leads us to conclude that points of the phase space $\hat{\Gamma}$ are much more relevant than pairs (a_1, a_2). In the case of Euclidean $2+1$ without cosmological constant these are coordinates of the reduced configuration space with direct geometric interpretation, but in our case they do not have any relevance. Therefore, a holomorphic quantization, naturally defined over phase space is much more natural. It is well known that holomorphic quantizations allow us for constructing coherent states, labeled by points on $\hat{\Gamma}$, that are of physical relevance for approximating semi-classical states.

The result, surprising at first, that there is a common description for two physically distinct systems, namely space-times with different signatures and curvatures, could be understood in a natural, geometrical construction given in Sec. 5.5. The key observation was that there is a common 4-dimensional space-time in which both space-times can be embedded. A point in the phase space $\hat{\Gamma}$ defines then an isometry on both hyper-surfaces allowing for a simultaneous quotient construction, and at the same time, given a natural isomorphism between this two systems. The moral that can be drawn from this model is that we might need to give, in the full $3+1$ gravity, some external input to the solutions of the quantum theory in order to recover the macroscopic, semiclassical limit of the theory with the right signature that we want to describe.

Finally, let us mention a difficulty that arises in the particular description, in terms of connections, that we have used. Note that in (5.99)-(5.100) of Sec. 5.5.1 we could add to, say a_1, a multiple of 2π, that is, $a_1 \to a_1 + 2k\pi$ with $k \in Z$ and we would get a *different* metric (5.97). On the other hand, as a points of $\hat{\Gamma}$, (a_1, a_2, s_1, s_2) and $(a_1 + 2k\pi, a_2, s_1, s_2)$ are the *same* point. Therefore, there is a many-to-one mapping from the space of equivalence class of metrics $\mathcal{M}^{(k,l)}$ to the phase space $\hat{\Gamma}$. All this space-times, with labels (k, l) have the same "holonomies" around the generators. This problem has already been discussed in [74, 75] and in more detail in [73]. Note that in the transformation (5.95) if we change a_1 to $a_1 + 2\pi$, the range of ϕ will go off the interval $[0, 2\pi)$ and therefore, the embedding we used (5.87)-(5.88) into M^4 will break down. It is only for values of $a_i \in [0, 2\pi)$ that we have the interpretation of M_q as embedded in dS_3^L. Thus, for values of $k, l > 1$, we would have to consider coverings space of the region of De Sitter space covered by the embedding.

In our construction of the reduced phase space $\hat{\Gamma}$, we used holonomies as gauge invariant quantities on phase space Γ and later on, we restricted ourselves to the constraint surface of flat connections. However, there is a subtlety that one has to worry about at some point. It is related to the fact that holonomies are invariant under *all* gauge transformations. That is, they do not distinguish between large and small transformations. However, the canonical analysis of constrained systems tells us that one should mode out by *small* gauge transformations only, that is, those transformations generated by the constraints.

Consider the connection (5.48) and the gauge transformation generated by $g = g(x_1, x_2)$ given by

$$g = \exp(2\pi i J_3\, x_1)\,. \tag{5.107}$$

The gauge related connection $\mathbf{A}^g_a$ is

$$\mathbf{A}^g_a = iJ_3\left[((a_1 + 2\pi) - is_1)\,(\mathrm{d}x_1)_a + (a_2 - is_2)\,(\mathrm{d}x_2)_a\right]\,. \tag{5.108}$$

We see that we can generate all non-isometric space times by applying gauge transformations of the form $g_{(k,l)} = \exp(2k\pi i J_3\, x_1 + 2l\pi i J_3\, x_2)$, where (k, l) label the *winding number* of the mapping.

The question of whether $g_{(k,l)}$ are large transformation can now be answered. Recall that there is a $\mathbb{Z}_2$ surjective mapping from $\mathrm{SL}(2,\mathbb{C})$ to $\mathrm{SO}(3,\mathbb{C})$, that is $\mathrm{SO}(3,\mathbb{C}) = \mathrm{SL}(2,\mathbb{C})/\mathbb{Z}_2$. The group $\mathrm{SL}(2,\mathbb{C})$ is simply-connected, since it is homeomorphic to $S^3 \times R^3$. Therefore, the fundamental group $\Pi_1(\mathrm{SO}(3,\mathbb{C}))$ is $\mathbb{Z}_2$. This means that there are closed loops in $\mathrm{SO}(3,\mathbb{C})$ that cannot be continuously deformed to the identity. Now, since the gauge transformation (5.107) does not depend on the coordinate x_2, it can be effectively seen as a mapping from S^1 to $\mathrm{SO}(3,\mathbb{C})$. The question of whether it is a small or large transformation reduces to asking which homotopy class in $\Pi_1(\mathrm{SO}(3,\mathbb{C}))$ it belong to. It is easy to see that it belongs to the class not homotopic to the identity, therefore it *is* a large gauge transformation. We see that this homotopically inequivalent loops give us four possible "sectors" of the theory (two for each generator). Thus, we can not account for the countable number of space-times with the same holonomy in terms of large gauge transformations.

Let us end with a remark. It is possible to understand this discrepancy between the geometrodynamical approach and the gauge theoretic one that we have used here. The basic distinctions comes from the fact that the co-triad e^I_a, taken to be *non-degenerate* in the geometrodynamical approach, looses this virtue in the gauge theoretic viewpoint since one is allowing for *arbitrary* gauge transformations mapping non-degenerate e to degenerate ones. Effectively, we are working on different phase spaces. For vanishing and negative cosmological constant (with Lorentzian signature), where the gauge theoretical description is always real, this multiplicity phenomenon is not present. One might then speculate that the reason for the existence of this problem can be traced back to the complex nature of the gauge group. We leave the resolution of this matters for a forthcoming publication [76].

A

Bundles, Connections, Gauge Fields and all that

The notion of fiber bundle is at the foundation of gauge theories. In this appendix we shall give a brief introduction to this topic, stating the main definitions and facts that are used throughout this thesis. In the first part we define fiber bundles. In the second one, the notion of connection on a bundle is introduced. In the third part, the curvature of a connection is considered. A more detailed treatment can be found in [12, 77, 78].

A.1 Fiber Bundles

A *bundle* is a triplet (E, π, M), where E and M are taken to be C^∞-manifolds and $\pi : E \to M$ is surjective and smooth. E is called the *total space* or *bundle space* and M is the *base space*. The map π is called the *projection map*. The inverse image $\pi^{-1}(x)$, $x \in M$, is the *fiber*, F_x, over x. Fiber bundles are those bundles whose fibers over all of M are homeomorphic to a common space F, the *typical fiber*. Thus, F is known as the *fiber* of the bundle.

A *cross section* of a bundle (E, π, M) is a map $s : M \to E$ such that the image of each point $x \in M$ lies in the fiber F_x over x, i.e. $\pi \circ s = \mathrm{id}_M$. The space of sections of a bundle is denoted by $\Gamma(E)$.

Most of the bundles of interest are *locally trivial*. This means that for each $x \in M$, there is a neighborhood U of x and an isomorphism,

$$\phi : \pi^{-1}(U) \longrightarrow U \times F \,, \tag{A.1}$$

sending each fiber $\pi^{-1}(x)$ to $\{x\} \times F$. We call ϕ a *local trivialization*. Intuitively, a bundle looks locally as a product of the base manifold and the fiber. If the space E is globally $M \times F$ the bundle is said to be *trivial*. Therefore, a general bundle can be thought of as a sort of 'twisted product'. We need an extra structure in the definition of a fiber bundle, namely the notion of the group G of homeomorphisms of F onto itself.

Example. A simple example of a bundle is given by the *tangent bundle* TM. The total space TM is the space of pairs (x, v^a_x), $\forall\, x \in M$. Here $v^a_x \in T_xM$, the tangent space of M at x. For later convenience, we have introduced here a bit of extra notation. We denote by v^a an element of the tangent space, where the index 'a' is an *abstract index* in the sense of Penrose [79], [81]. The base space is the manifold M. The fiber $F_x = T_xM$. The typical fiber F is $\mathbb{R}^n$. The projection is given by $\pi : (x, v_x^a) \to x$. The structure group G is the group $GL(n, \mathbb{R})$ of linear automorphisms of $\mathbb{R}^n$. Then, a vector field on M is nothing but a section of this bundle $v^a : M \to TM$.

This example is a particular case of bundles that are of great interest. A *vector bundle* is a fiber bundle such that the typical fiber F is a vector space and the group G is the group of linear automorphisms of F. We can construct general tensor bundles over M by taking tensor products of the 'basic' bundles TM and T^*M, just in the same way that one constructs tensors (and forms) starting from a vector space V.

Another example of a fiber bundle is the *frame bundle* $B(M)$. The total space consists of the set of basis vectors v_i^a of the tangent space for all points over the manifold M. Here i is an index $i = 1, \dots, n$ labeling the n basis vectors v^a. There is a natural free right action by $GL(n, \mathbb{R})$ on $B(M)$ defined by,

$$(v_1^a, v_2^a, \dots, v_n^a)g := (v_j^a g_1^j, v_j^a g_2^j, \dots, v_j^a g_n^j)\,, \tag{A.2}$$

where g in $GL(n, \mathbb{R})$. We can also see the array of vectors $(v_1^a, v_2^a, \dots, v_n^a)$ as a $n \times n$ matrix with non-zero determinant. Thus, the typical fiber F is also $GL(n, \mathbb{R})$.

The frame bundle is an example of a second type of bundles which are of physical relevance. These are called *principal bundles.* In a principal fiber bundle (E, π, M, G) each fiber is diffeomorphic to the structure group G. However, the diffeomorphism is not canonical; it depends on the covering $\{U_i\}$ of M and on the choice of local trivializations $\phi_i : \pi^{-1}(U_i) \to U_i \times G$. Principal bundles are in a sense more fundamental than vector bundles, since one can always regard vector bundles as *associated bundles* to a particular principal bundle. In our examples, the tangent bundle TM is the associated vector bundle to the frame bundle $B(M)$.

We gave as examples of vector bundles, those generated by tensor products of the tangent and cotangent bundles. We could have considered, for instance, a bundle $W(M)$ whose typical fiber is R^n (for an n-dimensional M), but with no canonical isomorphism between W_x, the internal space, and the tangent space T_xM. Just as sections of the tangent bundle are denoted by v^a, we shall denote by k^I the sections of this 'gauge bundle' $W(M)$. Recall that, in the abstract index notation, the abstract index 'I' labels the 'valence' of the field. We can now take tensor products of these bundles, whose sections shall be denoted by objects having 'mixed' indices. As a concrete example consider the bundle whose fiber at point x is $T_x^*M \otimes W_x$. A section e of this bundle will have two abstract indices: e_a^I. Note that, at each point x, e_a^I is a mapping from the tangent space T_xM to the 'internal space' W_x. That is, $e_a^I : T_xM \to W_x$. If the mapping is an isomorphism, then the section e_a^I is called the *soldering form.*

A.2 Connections

We have seen that a fiber bundle looks locally as a 'bundle of fibers'. However, we do not have a way of 'connecting' points on different fibers. We have also defined sections of various vector bundles, but we do not have a derivative operator defined over them. A connection on a bundle solves this two problems; it gives a prescription in order to 'connect' the fibers and defines a covariant derivative on sections. This two aspects of a connection give complementary viewpoints to the formulation of gauge theories. Given a principal bundle, a curve C in the base manifold M and a point p in the fiber over the initial point of C, a connection specifies the *horizontal lift* $\hat{C}$ on E. We also say that the curve $\hat{C}$ defines the *parallel transport* of the point p.

There are at least three definitions of a connection on a principal bundle (see for instance [78]). We shall only give one and mention the relation with the other ones. We have already referred to one of them. The fact that a connection lifts a curve on the base manifold to the bundle means that it maps tangent vectors on T_xM to vectors on T_pE $(x = \pi(p))$.

A *connection on the principal fiber bundle* (E, π, M, G) is a 1-form w_μ on P with values in the vector space $\mathbf{g} := \mathrm{lie}(G)$, such that if $\tau^\mu \in T_pE$, we have,

1. $w_p(X_A) = A^i \quad \forall\, p \in E,\ A^i \in \mathcal{G}$,
2. $(\delta_g^* w)(\tau) = \mathrm{Ad}_{g^{-1}}(w_p(\tau)), \quad \forall\, \tau \in T_pE$,

where we have used the Lie algebra isomorphism ι between $\mathbf{g}$ and vector fields in E. $\iota : \mathbf{g} \to T_pE$, such that $X_A = \iota(A^i)$. The mapping δ_g is the right action of the group G on the fiber, and δ^* is the pull-back.

Now, we define a *horizontal vector* τ_h to be such that

$$w(\tau_h) = 0. \tag{A.3}$$

Note that a connection is effectively decomposing the tangent space at each point p into two components: $T_pE \sim V_pE \oplus H_pE$, where H_pE is the vector space of horizontal vectors and V_pE is the space of *vertical vectors*. The space of vertical vectors is 'tangent' to the fiber at each point and is, therefore, isomorphic to the Lie algebra $\mathbf{g}$. If we have two connections w_1 and w_2, the affine sum $(f \circ \pi)w_1 + (1 - f \circ \pi)w_2$ is also a connection, for $f \in C^\infty(M)$. Therefore, the space of connections is an *affine* space.

We shall now see how is it that a connection defines a 'covariant derivative' on section of vector bundles. Consider the vector bundle (E, π, M). Let us denote by $s \in \Gamma(E)$ a section of E. A *connection* $\mathcal{D}$ on M assigns to each vector field v^a on M a mapping $\mathcal{D}_v : \Gamma(E) \to \Gamma(E)$. satisfying the properties,

$$\begin{aligned}
\mathcal{D}_v(\alpha s) &= \alpha \mathcal{D}_v s \\
\mathcal{D}_v(s + t) &= \mathcal{D}_v s + \mathcal{D}_v t \\
\mathcal{D}_v(fs) &= v(f)s + f\mathcal{D}_v s \\
\mathcal{D}_{v+w}(s) &= \mathcal{D}_v(s) + \mathcal{D}_w(s) \\
\mathcal{D}_{fv}(s) &= f\mathcal{D}_v(s)
\end{aligned} \tag{A.4}$$

for all $v^a, w^a \in TM$, $s, t \in \Gamma(E)$, $f \in C^\infty(M)$ and all scalars α. Given any section s and vector field v^a, we call $\mathcal{D}_v s$ the *covariant derivative* of s in the direction v^a. From the two last properties we see that, if we regard the covariant derivative as 'eating' both a vector v^a and a section s and 'spitting' a section s', then it is a linear operator in the vector field v^a. This suggests that we define an object with a 1-form index $\mathcal{D}_a$ such that

$$\mathcal{D}_v s =: v^a \mathcal{D}_a s\,. \tag{A.5}$$

The object $\mathcal{D}_a$ will be simply called the *covariant derivative.*

The space of connection $\mathcal{D}_a$ is an affine space. This means that if we take the difference $\mathcal{D}_a - \mathcal{D}'_a$ of two connections there is a tensorial field $\mathcal{A}_a$ (a section of the appropriate bundle) such that,

$$(\mathcal{D}_a - \mathcal{D}'_a)(s) = \mathcal{A}_a(s)\,. \tag{A.6}$$

Let us be concrete and take as our bundle the tangent bundle TM. A section w^a is a vector field on M. The covariant derivative $\mathcal{D}_a w^b$ is therefore a section of $TM \otimes T^*M$ (or a $\binom{1}{1}$ tensor). Therefore, the difference of two covariant derivatives is

$$(\mathcal{D}_a - \mathcal{D}'_a) w^b = \mathcal{A}_a{}^b{}_c w^c\,. \tag{A.7}$$

If the section is a vector with 'internal index' k^I, the corresponding equation reads,

$$(\mathcal{D}_a - \mathcal{D}'_a) k^I = \mathcal{A}_a{}^I{}_J k^J\,. \tag{A.8}$$

Sections of bundles with mixed indices are called *generalized tensor fields* [82]. It is sometimes convenient to fix an 'origin' of the space of connections and reduce the affine space to a vector space. We do that by declaring an operator $\overset{\circ}{\mathcal{D}}_a$ with zero curvature (see below) as the origin and labeling any connection $\mathcal{D}$ by its corresponding generalized tensor field defined by $(\mathcal{D}_a - \overset{\circ}{\mathcal{D}}_a) k^I = A_a{}^I{}_J k^J$ for all k^I. We shall denote by $A_a{}^I{}_J$ the tensor field $\mathcal{A}_a{}^I{}_J$ whenever one of the connections, i.e., the origin, has zero curvature.

We have used the 'curly' $\mathcal{D}_a$ to denote covariant derivative. There are other possibilities. For instance, in chapter 4 we use the symbol ∇_a for the covariant derivative compatible with the metric g_{ab}, namely $\nabla_c g_{ab} = 0$. In chapter 5, the Lorentz connection in the 3-dimensional space-time is denoted by D_a.

A.2.1 Holonomy

We have seen the two manifestations of a connection, both as a splitting of the tangent space of a principal bundle and as a covariant derivative on the sections of a vector bundle. Furthermore, we can see that a connection defines a *parallel transport* in both cases. In a principal bundle, we have shown that, for a given path γ on M, $\gamma : [0,1] \to M$, with $x = \gamma(0)$ and $y = \gamma(1)$, a connection defines the lifting, $\hat{\gamma}$, of γ, such that $p = \hat{\gamma}(0)$, $x = \pi(p)$ and $q = \hat{\gamma}(1)$, $y = \pi(q)$. We say that q is the *parallel transport* of p along γ. If the path is closed, namely, if we have a *loop* α, then the point q lies in the same fiber as p. Since the group acts by right action on the fiber F_x, there is a $g \in G$ such that $q = pg$. We say that the group element g is the *holonomy* of the connection along the loop α. The holonomy is independent of the point p at which we started the parallel transport. It is sometimes denoted by $g = H(A, \alpha)$, where we 'label' the connection by A. It is important to note that it depends on the loop α and the connection A only.

Let us now consider vector bundles. The intuitive picture is very similar. We consider a base point x and a point p on the fiber over it. The point p is an element of the vector space $V = F_x$. For concreteness let us suppose that $E = TM$, so the point p is a tangent vector v^a at point x. Given a path $\gamma : [0,1] \to M$ in M, we want to define the parallel transport of v^a in the same way that we defined the parallel transport for a principal bundle. How can we use the covariant derivative defined in the last part? We say that a vector $v^a(t)$, $t \in [0,1]$, is parallel transported if and only if

$$\dot{\gamma}^b(t)\mathcal{D}_b v^a = 0\,, \tag{A.9}$$

for all t, where $\dot{\gamma} := \frac{\mathrm{d}\gamma(t)}{\mathrm{d}t}$. This equation, when written in a local coordinate system, is a first order equation of the form,

$$\dot{\gamma}^a(\partial_a v^b + A_a{}^b{}_c v^c) = \frac{\mathrm{d}}{\mathrm{d}t} v^b(t) + A_a{}^b{}_c \dot{\gamma}^a(t) v^c = 0\,. \tag{A.10}$$

In the last equation we have denoted by ∂_a the preferred flat connection $\overset{\circ}{\mathcal{D}}_a$. The solution is given by the *path ordered exponential*

$$\mathcal{P}\exp\left[-\int_0^t A_a\dot{\gamma}^a \mathrm{d}t\right] := \sum_{n=1}^{\infty} \int_0^t \mathrm{d}t_1 \cdots \int_0^{t_{n-1}} \mathrm{d}t_n A_{a_1}\dot{\gamma}^{a_1}(t_1)\cdots A_{a_n}\dot{\gamma}^{a_n}(t_n). \tag{A.11}$$

Note that the path order exponential has two indices $\mathcal{P}[\gamma, \mathcal{D}](t) = \mathcal{P}^b{}_c$. Therefore, the parallel transported v^a is given by

$$v^b(t) := \mathcal{P}\exp[\gamma, \mathcal{D}](t)^b{}_c\, v^c(0)\,. \tag{A.12}$$

In this case the holonomy is again in the group G. To see this recall that the group G acts on the fiber by left action. Given an element g of $GL(n, \mathbb{R})$ (the structure group of TM), we can define the action $\delta_g(v) := g^a{}_b v^b$. Thus, the group element connecting the two vectors $v^a(0)$ and $v^a(1)$ on the fiber F_x, is given by the path ordered exponential of the closed loop α:

$$H(\alpha, A) := \mathcal{P}\exp\left[-\oint_\alpha A_a \dot{\alpha}^a \mathrm{d}t\right]\,. \tag{A.13}$$

A.2.2 Gauge transformations

Just as in ordinary differential geometry one can view diffeomorphisms as active or passive, one can do the same thing with gauge transformations on fiber bundles. In the passive viewpoint, one looks at the same objects under "change of coordinates". In the bundle picture, this amounts to a change of local trivializations. For a point $x \in M$ in the intersection $U_i \cap U_j$ of two charts U_i and U_j, there correspond two local trivializations $\phi_i, \phi_j : \pi^{-1}(U_i \cap U_j) \to U_i \cap U_j \times F$. Given a connection $\mathcal{D}$, and a path $\gamma \in U_i \cap U_j$, we obtain a path $\hat{\gamma}$ on E, which under the trivializations will yield two different paths $\hat{\gamma}_i, \hat{\gamma}_j$ in $U_i \cap U_j \times F$. This two paths are related by a 1-parameter family of "transition functions": $\hat{\gamma}_i(t) = g_{ij}(t) \cdot \hat{\gamma}_j(t)$. The $G-valued$ function $g_{ij}(t)$ on $U_i \cap U_j$ is then a *gauge transformation.*

In the active viewpoint, gauge transformations are *vertical automorphisms* of the bundle E, i.e., $\mathcal{F} : \pi^{-1}(x) \to \pi^{-1}(x)$, $\mathcal{F}(pg) = \mathcal{F}(p)g$, $\forall\, p \in E$ where $pg = R_g p$ is the right action of G on F_x.

For a vector bundle in which the group G acts by endomorphisms of the fiber V, a gauge transformation is a one to one and onto linear transformation of the fiber F_x which varies smoothly with points x in M. In cases of interest the group G is required to be a Lie group. The gauge transformation $T : \Gamma(E) \to \Gamma(E)$ is such that $T(x)$ lives in G for all x in M. The set of gauge transformations is a group, with products and inverses given by,

$$\begin{aligned} (gh)(p) &= g(p)h(p)\,, \\ g^{-1}(p) &= g(p)^{-1}\,. \end{aligned}$$

It is sometimes denoted by $\mathcal{G}$. Let $\mathcal{D}$ be a connection on E and let $g \in \mathcal{G}$ be a gauge transformation. Then, there is a new connection D' on E such that,

$$\mathcal{D}'_v(gs) = g\mathcal{D}_v s\,, \tag{A.14}$$

for all vector fields v^a an M and sections s in E. The formula is given by,

$$\mathcal{D}'_v(s) = g\mathcal{D}_v(g^{-1}s)\,. \tag{A.15}$$

Using a local trivialization over U_i we can write the connection in the form $\mathcal{D}_a = \overset{\circ}{\mathcal{D}}_a + A_a$. Then the *vector potential* A'_a for $\mathcal{D}'$ is given by

$$A'_a = gA_a g^{-1} + g\,\overset{\circ}{\mathcal{D}}_a\, g^{-1}\,. \tag{A.16}$$

This is the formula usually found in books for gauge transformations. It is common to denote as $\mathcal{A}$ the space of connections on a bundle E and write $\mathcal{A}/\mathcal{G}$ for the space of gauge equivalent class of connections, or the space of *connections modulo gauge transformations.*

Let us consider the case of the simplest theory of connections: the electro-magnetic field. In this case the gauge group G is $U(1)$. For simplicity, lets assume that the bundle is trivial, that is $E = M \times \mathbb{C}$, so the fiber F_x over any point x is $\mathbb{C}$. A connection $\mathcal{D}$ on E can be described by its vector potential A. E becomes a $U(1)$ bundle if we think of its standard fiber $\mathbb{C}$ as the fundamental representation of the group $U(1)$. If the connection $\mathcal{D}$ is a $U(1)$ connection, the vector potential 1-form $\tilde{A}_a$ must live in $u(1) = \{i\tau : \tau \in \mathbb{R}\}$. So $\tilde{A}_a$ is equal to i times a real valued 1-form. One normally uses this real valued 1-form as the vector potential in electro-magnetism and denotes it by A_a. The gauge transformation (A.16) simplifies to,

$$iA'_a = iA_a + g\partial_a g^{-1}\,, \tag{A.17}$$

and if we can write $g = \mathrm{e}^{-if}$ for some real function f, we have $g\partial_a g^{-1} = i\partial_a f$, hence

$$iA'_a = iA_a + i\partial_a f\,. \tag{A.18}$$

This is the standard definition of gauge transformation in electromagnetism.

A.3 Curvature and Gauge Fields

In this section we will introduce the *curvature* F of the connection $\mathcal{D}$. The curvature is, in gauge theories, the 'gauge field'. Let E be a vector bundle, and let v^a, w^a be two vector fields in M. Then, the curvature $F(v,w)$ of a torsion-free connection is the operator on sections of E given by,

$$F(v,w)s = \mathcal{D}_v \mathcal{D}_w s - \mathcal{D}_w \mathcal{D}_v s - \mathcal{D}_{[v,w]} s\,. \tag{A.19}$$

That is, the curvature is a measure of the failure of covariant derivatives to commute. If the curvature is such that $F(v,w)s = 0$ for all vectors v^a and w^a and sections s, it is said to be *flat*. As an abstract operator we can write it as

$$F(v,w) = [\mathcal{D}_v, \mathcal{D}_w] - \mathcal{D}_{[v,w]}\,. \tag{A.20}$$

It has the following properties:

$$F(v,w) = -F(w,v)\,, \tag{A.21}$$

$$F(fv,w)s = F(v,fw)s = fF(v,w)s\,, \tag{A.22}$$

$$F(v,w)fs = fF(v,w)s\,. \tag{A.23}$$

The second equation tell us that the curvature is linear in both its arguments and therefore there should be a 2-form associated to it. The last equation tells us that F defines a $C^\infty(M)$-linear map from $\Gamma(E)$ to itself, i.e. a section of End(E). Thus, there is a End(E)-valued two-form F_{ab} corresponding to the curvature.

For instance, if the bundle in the tangent bundle TM, so section are vector fields on M, then the curvature is a $gl(n,\mathbb{R})$ valued two form $F_{ab}{}^c{}_d$ such that

$$2\mathcal{D}_{[a}\mathcal{D}_{b]}v^c = F_{ab}{}^c{}_d v^d\,. \tag{A.24}$$

Similarly, for a internal vector k^I, we have,

$$2\mathcal{D}_{[a}\mathcal{D}_{b]}k^I = F_{ab}{}^I{}_J k^J\,. \tag{A.25}$$

We can now write an expression for $F_{ab}{}^I{}_J$ in terms of the 1-form $A_a{}^I{}_J$. Recall that we can write

$$\mathcal{D}_a k^I = \overset{\circ}{\mathcal{D}}_a k^I + A_a{}^I{}_J k^J.$$

Therefore,

$$F_{ab}{}^I{}_J = 2\,\overset{\circ}{\mathcal{D}}_{[a} A_{b]}{}^I{}_J + [A_a, A_b]^I{}_J\,, \tag{A.26}$$

where $[A_a, A_b]^I{}_J$ is the Lie-bracket of A_a and A_b.

There is a simple geometrical identity satisfied by the curvature tensor. It is called the *Bianchi identity*. One form of it is the following:

$$[\mathcal{D}_u,[\mathcal{D}_v,\mathcal{D}_w]] + [\mathcal{D}_w,[\mathcal{D}_u,\mathcal{D}_v]] + [\mathcal{D}_v,[\mathcal{D}_w,\mathcal{D}_u]] = 0\,, \tag{A.27}$$

which can be rewritten in terms of the covariant derivative and F as,

$$\mathcal{D}_{[a}F_{bc]} = 0\,. \tag{A.28}$$

For the electro-magnetic field, The covariant derivative is given by

$$\mathcal{D}_a = \partial_a + iA_a\,, \tag{A.29}$$

where we have denoted by ∂ the fiducial flat connection. The curvature $\tilde{F}$ is given by,

$$\tilde{F}_{ab} = 2i\partial_{[a}A_{b]} := iF_{ab}\,, \tag{A.30}$$

since the connection is Abelian (The Lie-algebra of U(1) is Abelian). We can therefore regard the (real) F_{ab} as the curvature, or gauge field. The Bianchi identity is just $\partial_{[a}F_{bc]} = 0$. In this case, since A is a real 1-form, we have $F = \mathrm{d}A$ and $\mathrm{d}F = 0$, where d is the exterior derivative.

B

Gauss Linking Number

In this appendix we shall show the analytic equivalence between the intersection number $I(\alpha, S_\beta)$ of Eq (2.6) and the Gauss linking number of Eqs (2.1,2.2). We start by re-writing the intersection number,

$$\begin{aligned} I(\alpha, S_\beta) &= \{B[\alpha], E[\beta]\} \\ &= \int \mathrm{d}^3x\, F^a[\alpha, \bar{x}) w_a[\beta, \bar{x})\,, \end{aligned} \tag{B.1}$$

where $F^a[\alpha, \bar{x})$[1] is the so-called *form factor* of the loop α:

$$F^a[\alpha, \bar{x}) = \oint_\alpha \mathrm{d}\alpha^a \delta^3(\alpha, \bar{x})\,, \tag{B.2}$$

and $w_a[\beta, \bar{x})$ is given by

$$w_a[\beta, \bar{x}) = \int_{S_\beta} \mathrm{d}S_a\, \delta^3(\bar{x}, \bar{s}_\beta)\,. \tag{B.3}$$

Also, note that w_a is a potential for the form factor, since,

$$F^a[\alpha, \bar{x}) = \epsilon^{abc}\partial_b w_c[\alpha, \bar{x}) \tag{B.4}$$

There is an extra 'gauge freedom' in the definition of w since w_a and $w'_a = w_a + \partial_a f$ give rise to the *same* form factor F^a.

So far, the intersection number does not depend on any background structure and is therefore, topological in nature. The expression (2.1) is, however, written in terms of an Euclidean metric. Let us therefore introduce such a metric. Then,

$$\delta^3(\bar{x} - \bar{y}) = -\frac{1}{4\pi}\nabla^2_x \frac{1}{|\bar{x} - \bar{y}|} = -\frac{1}{4\pi}\partial_a \partial^a_x \left(\frac{1}{|\bar{x} - \bar{y}|}\right)\,. \tag{B.5}$$

[1] The notation $F^a[\alpha, \bar{x})$ with 'mixed' brackets means that F depends *functionally* on the loop α and as a function of the point $\bar{x}$.

We can now re-write the intersection number as,

$$
\begin{aligned}
I(\alpha, S_\beta) &= -\frac{1}{4\pi}\int \mathrm{d}^3x \int \mathrm{d}^3y\, F^a[\alpha,\bar{x})\, w_a[\beta,\bar{y})\, \partial_d \partial_y^d \left(\frac{1}{|\bar{x}-\bar{y}|}\right)\,, \\
&= \frac{1}{4\pi}\int \mathrm{d}^3x \int \mathrm{d}^3y\, F^a[\alpha,\bar{x})\, \partial_d w_a[\beta,\bar{y})\, \partial_y^d \left(\frac{1}{|\bar{x}-\bar{y}|}\right)\,.
\end{aligned}
\tag{B.6}
$$

where in the second step we have integrated by parts. Now,

$$
I(\alpha, S_\beta) = \frac{1}{4\pi}\int \mathrm{d}^3x F^a[\alpha,\bar{x}) \int \mathrm{d}^3y \left(2\partial_{[d}w_{a]}[\beta,\bar{y}) + (\partial_a w_d[\beta,\bar{y}))\partial_y^d\left(\frac{1}{|\bar{x}-\bar{y}|}\right)\right)\,. \tag{B.7}
$$

The last term can be again integrated by parts

$$
\int \mathrm{d}^3y(\partial_a w_d[\beta,\bar{y}))\partial_y^d\left(\frac{1}{|\bar{x}-\bar{y}|}\right) = -\int \mathrm{d}^3y\frac{1}{|\bar{x}-\bar{y}|}(\partial_a\partial^d w_d[\beta,\bar{y})) = 0\,, \tag{B.8}
$$

where we have used the gauge freedom to select w_a such that $\partial^a w_a = 0$.

Finally, we have,

$$
\begin{aligned}
I(\alpha, S_\beta) &= \frac{1}{4\pi}\int \mathrm{d}^3x \int \mathrm{d}^3y\, F^a[\alpha,\bar{x})F^b[\beta,\bar{y})\epsilon_{abc}\partial_y^c\left(\frac{1}{|\bar{x}-\bar{y}|}\right)\,, \\
&= \frac{1}{4\pi}\int \mathrm{d}^3x \int \mathrm{d}^3y\, F^a[\alpha,\bar{x})F^b[\beta,\bar{y})\epsilon_{abc}\frac{(x^c-y^c)}{|\bar{x}-\bar{y}|^3}\,.
\end{aligned}
\tag{B.9}
$$

Now, the definition of the form factor implies that: $\int \mathrm{d}^3x F^a[\alpha,\bar{x})f_a = \oint_\alpha \mathrm{d}\alpha^a f_a$. Hence, we have the desired equality:

$$
I(\alpha, S_\beta) = \frac{1}{4\pi}\oint_\alpha \mathrm{d}s \oint_\beta \mathrm{d}t\; \epsilon_{abc}\, \dot{\alpha}^a(s)\, \dot{\beta}^b(t)\frac{\alpha^c(s)-\beta^c(t)}{|\alpha(s)-\beta(t)|^3}\,. \tag{B.10}
$$

which is the form of the linking number introduced originally by Gauss.

C

Quantum Fields and Fock Representation

In this appendix we shall outline the quantization of linear fields in the Fock representation. This construction is relevant for both the Maxwell field of Chapters 2 and 3 and the Klein-Gordon equation of Chapter 4.

Fix a 4-dimensional globally hyperbolic space-time (M, g). The first step is to consider the vector space Γ of solutions of the equations of motion. One then constructs the algebra of fundamental observables to be quantized, which in this case consists of suitable linear functionals on Γ. The next step is to construct the so called *one-particle Hilbert space* $\mathcal{H}$ from the space Γ. This amounts to defining a complex structure on Γ compatible with the naturally defined symplectic structure thereon. From the Hilbert space $\mathcal{H}$ one constructs its symmetric (since we are considering bose fields) Fock space $\mathcal{F}_s(\mathcal{H})$. The final step is to represent the algebra of observables in the Fock space as suitable combinations of creation and annihilation operators.

We will construct in detail the quantization outlined above for the case of the Maxwell field since in our opinion, an unified construction (although completely elementary) is not available elsewhere. In the first part we shall consider the classical description of the Maxwell field and in the second part the Fock quantization outlined above.

C.1 Maxwell Field

In the construction of the quantization of the Maxwell field there are two equivalent but complementary viewpoints, namely the *covariant* and the *canonical* formalisms. In what follows we shall develop both approaches and show their equivalence.

C.1.1 Covariant Phase Space

The starting point for the construction of the covariant phase space is the identification of the symplectic vector space Γ of solutions to the equations of motion. Let us start the by writing down the action for the free Maxwell theory:

$$S_{\rm M} \;:=\; -\frac{1}{4}\int_M F^{ab}F_{ab}\,\sqrt{|g|}\,{\rm d}^4x\,,$$
$$\;=\; -\frac{1}{2}\int_M F^{ab}\nabla_{[a}A_{b]}\,\sqrt{|g|}\,{\rm d}^4x\,. \qquad \text{(C.1)}$$

The variation of the action is given by,

$$\delta S_{\rm M} = \int_M (\nabla_a F^{ab})\delta A_a\,\sqrt{|g|}\,{\rm d}^4x - \int_{\partial M} F^{ab}\delta A_b\,{\rm d}\Sigma_a\,. \qquad \text{(C.2)}$$

The volume term tells us that the action is extremized when $\nabla_a F^{ab}=0$. Since we are assuming that there exists a connection A_a such that its curvature is the Maxwell field $F_{ab}:=2\nabla_{[a}A_{b]}$, the equation $\nabla_{[a}F_{bc]}=0$ is automatically satisfied (the Bianchi identity). Therefore we have the full set of Maxwell equations. The second term in Eq (C.2), the boundary term, is often referred to as the *symplectic current*. It can be interpreted as a 1-form on the space $\bar{\Gamma}$ of solutions to the equations of motion. It is acting on the vector δA_a and producing a number. We can take now another 'variation' of this term in order to get the conserved (pre)-symplectic structure $\omega(\cdot,\cdot)$,

$$\omega(\delta A,\widetilde{\delta A}) := \int_\Sigma (\delta F^{ab}\delta\tilde{A}_b - \delta\widetilde{F}^{ab}\delta A_b)\,{\rm d}\Sigma_a\,, \qquad \text{(C.3)}$$

where Σ is any Cauchy surface in the space-time M. We have not been very precise about functional analytic issues. We are just requiring falloff conditions (on any Σ) such that the symplectic form at spatial ∞ vanishes. If, in particular, we restrict ourselves to solutions of the Maxwell equations that induce data of *compact support* an any Cauchy surface, that conditions will be satisfied. This bilinear mapping is, however, degenerate. There are tangent vectors X_a such that $\omega(X,Y)=0,\ \forall\, Y\in\bar{\Gamma}$. This are the degenerate directions of ω. Consider vectors of the type $X_a=\nabla_a\lambda$ for some function λ. Then, using the fact that $\nabla_{[a}X_{b]}=0$ we have,

$$\omega(X,\delta A) \;=\; \int_\Sigma -\delta F^{ab}\nabla_b\lambda\;{\rm d}\Sigma_a$$
$$\;=\; \int_\Sigma \lambda\nabla_b(\delta F^{ab})\;{\rm d}\Sigma_a = 0\,.$$

since we are restricting ourselves to the space $\bar{\Gamma}$, the tangent vectors satisfy the *linearized* equation of motion, that in this case coincide with the Maxwell equations. We can conclude that the degenerate directions of ω are of the form $\nabla_a\lambda$. This is the manifestation, in the covariant phase space approach, of the "gauge freedom" present in electro-magnetism. In order to get a true symplectic space, we should take the quotient of $\bar{\Gamma}$ by the degenerate directions of ω to get Γ, the (reduced) phase space of the theory. Note that Γ can be equivalently parametrized by the equivalence class of gauge potentials $[A_a]$, where $A\sim\bar{A}$ iff $A_a=\bar{A}_a+\nabla_a\lambda$, or alternatively, by the gauge fields F_{ab}, satisfying Maxwell equations.

We can now write the (weakly non-degenerate) symplectic form on Γ:

$$\omega(F,\tilde{F}) = \int_\Sigma (F^{ab}\tilde{A}_b - \tilde{F}^{ab}A_b)\;{\rm d}\Sigma_a\,. \qquad \text{(C.4)}$$

Note that it is well defined on Γ since it does not depend on the representative of the equivalence class $[A]$. Note that in writing (C.4) we have used the fact that Γ is a linear space and therefore we can identify points in Γ with tangent vectors.

The next step is to construct observables of the theory, namely, real valued functions on Γ. A natural strategy is to use the symplectic form in order to construct such functions. Let h_a be a "test 1-form". The observable $\mathcal{O}[h] : \Gamma \to \mathbb{R}$, labelled by h, is defined by,

$$(\mathcal{O}[h])(F) := \omega(F,T) = \int_\Sigma (F^{ab} h_b - T^{ab} A_b)\ \mathrm{d}\Sigma_a\,, \tag{C.5}$$

where $T_{ab} := 2\nabla_{[a} h_{b]}$. We need it to be a well defined function on Γ, so $\mathcal{O}[h]$ should be invariant under gauge transformations $A_a \to A_a + \nabla_a \lambda$. Thus,

$$\int_\Sigma T^{ab} \nabla_b \lambda\ \mathrm{d}\Sigma_a = 0\,, \tag{C.6}$$

which implies $\nabla_a T^{ab} = 0$. Therefore, an element h_a of Γ defines by itself a linear observable. In the quantum theory, to each of this observables there will correspond a quantum operator, making the correspondence between solutions to Maxwell equations and quantum operators precise.

Let us re-write the symplectic form (C.4) in terms of the familiar electric and magnetic fields. Recall that given a local observer with four velocity t^a ($t_a t^a = -1$), then the *electric field* with respect to this observer is given by $E_a := t^b F_{ba}$. It is naturally defined as a 1-form. Since we have a metric we can 'raise' the index and define the corresponding vector field. We can also define the *dual tensor* of the field F_{ab} by: ${}^*F^{ab} := \frac{1}{2}\epsilon^{abcd} F_{cd}$, where ϵ^{abcd} is the canonical volume form defined by the metric g_{ab} with all its indices raised with the metric. The magnetic field is defined by $B_a := t^b\, {}^*F_{ba}$. In the integrand of the symplectic form, one is contracting the tensor F^{ab} with the unit normal n_a to the surface Σ (that is the meaning of $\mathrm{d}\Sigma_a := \epsilon_{abcd}\, \mathrm{d}\Sigma^{bcd}$), so we get naturally the electric field E^a with respect to Σ. We can now express (C.4) as follows,

$$\omega(F, \tilde{F}); = \int_\Sigma (E^a \tilde{A}_a - \tilde{E}^a A_a)\, \sqrt{h}\, \mathrm{d}^3 x\,. \tag{C.7}$$

This expression can be rewritten in terms of objects defined purely on the hyper-surface Σ. We can write,

$$\begin{aligned} F^{ab} A_b\, \mathrm{d}\Sigma_a &= \frac{1}{2}\, \epsilon^{abcd}\, {}^*F_{cd}\, A_b\, \mathrm{d}\Sigma_a\,, \\ &= \frac{1}{2}\, {}^*F_{cd}\, A_b\, \epsilon^{abcd} \epsilon_{afgh}\, \mathrm{d}\Sigma^{fgh}\,, \\ &= -\frac{1}{2}\, {}^*F_{cd}\, A_b\, \mathrm{d}\Sigma^{cdb}\,. \end{aligned}$$

Therefore, one can take the 3-form ${}^*F \wedge A$ and integrate it on Σ,

$$\omega(F, \tilde{F}) = -\frac{1}{2} \int_\Sigma ({}^*F_{[ab} \tilde{A}_{c]} - {}^*\tilde{F}_{[ab} A_{c]})\ \mathrm{d}\Sigma^{abc}. \tag{C.8}$$

Note that the pullback to Σ of the dual tensor ${}^*F_{ab}$ is, in a 3-dimensional sense, the electric field two-form: $E_{ab} := {}^*F_{ab}$. This is naturally dual to a vector density of weight one $\tilde{E}^c := \tilde{\eta}^{cab} E_{ab}$, which is the electric field arising from the canonical approach.

Finally, one can ask what the Poisson Bracket of the observables defined by (C.5) is. Given h_a and $\tilde{h}_a$ in Γ the Poisson bracket of the observables they define is given by,

$$\{\mathcal{O}[h], \mathcal{O}[\tilde{h}]\} := \omega(T, \tilde{T}) = \int_\Sigma (T^{ab}\tilde{h}_b - \tilde{T}^{ab}h_b)\ \mathrm{d}\Sigma_a\,. \tag{C.9}$$

We shall now go to the canonical approach.

C.1.2 Canonical Phase Space

The action (C.1) can be written in a $3+1$ fashion. First we write the expression for the action as follows,

$$S = -\frac{1}{4}\int_M g^{ab}g^{cd}F_{ac}F_{bd}\,\sqrt{|g|}\,\mathrm{d}^4x \tag{C.10}$$

Next, we decompose $g^{ab} = h^{ab} - n^a n^b$ with n^a the unit normal to Σ, and introduce an everywhere time-like vector field t^a and a 'time' function t such that the hyper-surfaces t =constant are diffeomorphic to Σ and such that $t^a\nabla_a t = 1$. We can write $t^a = Nn^a + N^a$. The volume element is given by $\sqrt{|g|} = N\sqrt{h}$. Using this identities in Eq.(C.10) we get,

$$\begin{aligned} S &= -\frac{1}{4}\int_I \mathrm{d}t \int_\Sigma N\sqrt{h}\left\{h^{ac}h^{bd}F_{ab}F_{cd} - \right. \\ &\quad \left. \frac{2}{N^2}h^{ac}\left[(\pounds_t A_a - \nabla_a(t\cdot A) + N^bF_{ab})(\pounds_t A_c - \nabla_c(t\cdot A) + N^dF_{cd})\right]\right\}, \end{aligned} \tag{C.11}$$

where $(t\cdot A) := t^bA_b$, and $I = [t_0, t_1]$ is an interval in the real line. Note that since for all the terms in the previous equation, both the connection and the curvature are contracted with purely "spatial" objects ($n^aN_a = n^ah_{ab} = 0$), then both A_a and F_{ab} in (C.11) are the pull-backs to Σ of the space-time objects. For simplicity, we shall continue writing A_a for the 3-dimensional connection.

From the $3+1$ form of the action (C.11) we can find the momenta canonically conjugated to A_a:

$$\tilde{\Pi}^a := \frac{\delta S}{\delta(\pounds_t A_a)} = \tfrac{\sqrt{h}}{N}h^{ac}(\pounds_t A_c - \nabla_c(t\cdot A) + N^dF_{cd})\,. \tag{C.12}$$

It can be rewritten as,

$$\tilde{\Pi}^a = \tfrac{\sqrt{h}}{N}h^{ac}(t^b - N^b)F_{bc} = \tfrac{\sqrt{h}}{N}h^{ac}Nn^bF_{bc} = \sqrt{h}\,E^a\,, \tag{C.13}$$

thus, the canonically conjugated momenta is just the *densitized* electric field (w.r.t. Σ).

The Eq.(C.12) can be solved for the 'velocity', $\pounds_t A_a$,

$$\pounds_t A_a = \tfrac{N}{\sqrt{h}}\,h_{ac}\tilde{\Pi}^a + \nabla_c(t\cdot A) - N^dF_{cd} \tag{C.14}$$

We can perform a Legendre transform of the Lagrangian density in order to find the Hamiltonian:

$$\begin{aligned} H &:= \int_\Sigma \mathrm{d}^3x\left(\tilde{\Pi}^a\pounds_t A_a - \tilde{\mathcal{L}}\right) \\ &= \int_\Sigma \mathrm{d}^3x\left(-(t\cdot A)\nabla_a\tilde{\Pi}^a - N^dB_{ad}\tilde{\Pi}^a + \tfrac{N}{2\sqrt{h}}h_{ac}\tilde{\Pi}^a\tilde{\Pi}^c + \tfrac{N\sqrt{h}}{4}h^{ac}h^{bd}B_{ab}B_{cd}\right). \end{aligned} \tag{C.15}$$

We have denoted by B_{ab} the curvature of the 3-dimensional connection A_a. It is related to the magnetic field in the following way: $B^a := \frac{1}{\sqrt{h}}\tilde{\eta}^{abc}B_{bc}$. The last term in (C.15) can be rewritten: $h^{ac}h^{bd}B_{ab}B_{cd} = B^eB^f\epsilon^{cd}{}_e\epsilon_{cdf} = 2h_{ab}B^aB^b$. In the 'Dirac analysis' of the action (C.10) the first step is to identify the *configuration variables.* In this case, these are pairs $(\phi := (t \cdot A), A_a)$. In the action there is no term corresponding to time derivative of ϕ so we have a primary constraint $\chi_1 = \Pi_\phi \approx 0$. The basic Poisson brackets are,

$$\{A_a(x), \tilde{\Pi}^b(y)\} = \delta_a^b\delta^3(y,x) \quad ; \quad \{\phi(x), \Pi_\phi(y)\} = \delta^3(y,x)\,. \tag{C.16}$$

Asking that the constraint be preserved in time with respect to the Hamiltonian (C.15) leads to the secondary constraint $\nabla_a\tilde{\Pi}^a \approx 0$. There are no extra constraints. They form a *First Class* system. One can eliminate the first one by giving the gauge condition $\chi_2 := \phi - \lambda(\bar{x}) \approx 0$, with λ an arbitrary function on Σ. We can reduce the constraints (χ_1, χ_2) since they form a second class pair. We are then left with the Gauss constraint $\nabla_a\tilde{\Pi}^a \approx 0$. Now, ϕ has the role of a *Lagrange multiplier.* Therefore, the phase space Γ' is coordinatized by the pairs $(A_a, \tilde{\Pi}^b)$. The constraint surface $\hat{\Gamma}$ are the point in Γ' where the Gauss constraint is satisfied. The reduced phase space Γ_c is the space of orbits generated by the gauss constraint in $\hat{\Gamma}$. The canonical transformation generated by the (smeared) Gauss constraint, $G[\lambda] = \int_\Sigma \lambda\nabla_b\tilde{\Pi}^b \mathrm{d}^3x$, is given by,

$$A_a \longrightarrow A_a - \nabla_a\lambda\,. \tag{C.17}$$

Therefore, the (reduced) phase space is given by pairs $([A], \tilde{\Pi})$ of gauge equivalence class of connections and vector densities satisfying Gauss' law. One convenient gauge choice is to ask that $\nabla^aA_a = 0$. This is a good gauge condition so one can coordinatize Γ_c by (A_a, E^a), a pair of divergence-less (transverse) vector fields on Σ. We have used the fact that we have a metric on Σ to de-densitize the momenta $\tilde{\Pi}$.

The Poisson brackets (C.16) induce a (weakly) non-degenerate symplectic form ω on pairs of tangent vectors $(\delta A, \delta E)$:

$$\omega\left((\delta A, \delta E); (\widetilde{\delta A}, \widetilde{\delta E})\right) = \int_\Sigma \sqrt{h}\,\mathrm{d}^3x\left(\widetilde{\delta A}_a\delta E^a - \delta A_a\widetilde{\delta E}^a\right)\,. \tag{C.18}$$

We can now relate the two approaches and see that the phase space Γ from last section is precisely the space Γ_c constructed via the canonical approach. The key observation is that there is a one to one correspondence between initial data of compact support on Σ and solutions to the Maxwell equations on M. Therefore, to each element F_{ab} in Γ there is a pair (A_a, E^a) on Γ_c ($2\nabla_{[a}A_{b]} = F_{ab}{}^*$ and $E^a = h^{ab}n^cF_{cb}$ and more importantly, for each pair, there is a solution to Maxwell's equations that induces the given initial data on Σ. From now on, we shall refer to elements of the vector space Γ in-distinctively either as F_{ab} or as (A_a, E^b).

Observables for the space Γ can be constructed directly by giving smearing functions on Σ (compare to the discussion of the previous section in which the observables were constructed from *space-time* smearing objects). Given a 1-form g_a on Σ we can define,

$$E[g] := \int_\Sigma \sqrt{h}\;\mathrm{d}^3x\,E^ag_a\,. \tag{C.19}$$

Similarly, given a vector field f^a we can construct,

$$A[f] := \int_{\Sigma} \sqrt{h}\, \mathrm{d}^3x\, A_a f^a\,, \tag{C.20}$$

Asking that $E[g]$ be gauge invariant does not impose any condition on g_a, since Gauss' law does not 'move' the electric field. Note however that $E[g]$ takes the same value for g_a and $g_a + \nabla_a\lambda$. It is convenient to restrict ourselves to g_a satisfying $\nabla^a g_a = 0$. The requirement that $A[f]$ be gauge invariant tells us that $\nabla_a f^a = 0$. Therefore, in order to get well defined operators, we need the pairs (g_a, f^b) to belong to the phase space Γ. These are the precise images of the observables (C.5) given by the identification of phase spaces. The relation is given by $g_a = h_a{}^*$ and $f^a = 2\nabla^{[a}h^{b]}n_b$.

Note that any pair of test fields $(g_a, f^a) \in \Gamma$ defines a linear observable, but they are 'mixed'. More precisely, a connection g_a in Σ, that is, a pair $(g_a, 0) \in \Gamma$ gives rise to a *electric field* observable $E[g]$ and, conversely, a vector field $(0, f^a) \in \Gamma$ defines a *connection* observable $A[f]$.

C.2 Quantization

In this section we shall construct the quantum theory. The first step is to identify the 1-particle Hilbert space $\mathcal{H}$. The strategy is the following: start with (Γ, ω) a symplectic vector space and define $J : \Gamma \to \Gamma$, linear operator such that $J^2 = -1$. The *complex structure* J has to be compatible with the symplectic structure. This means that the $\mu(\cdot,\cdot) := \omega(\cdot, J\cdot)$ is a positive definite metric on Γ. The Hermitian inner product is then given by,

$$\langle \cdot,\cdot \rangle = \tfrac{1}{2\hbar}\mu(\cdot,\cdot) + i\tfrac{1}{2\hbar}\omega(\cdot,\cdot)\,. \tag{C.21}$$

The complex structure J defines a a natural splitting of $\Gamma_{\mathbb{C}}$, the complexification of Γ, in the following way: Define the 'positive frequency' part to consist of vectors of the form $\Phi^+ := \frac{1}{2}(\Phi - iJ\Phi)$ and the 'negative frequency' part as $\Phi^- := \frac{1}{2}(\Phi + iJ\Phi)$. Note that $\Phi^- = \overline{\Phi}^+$ and $\Phi = \Phi^+ + \Phi^-$. Since $J^2 = -1$, the eigenvalues of J are $\pm i$, so one is decomposing the vector space Γ in eigenspaces of J: $J(\Phi^\pm) = \pm i\Phi^\pm$. We have used the term 'positive frequency' since in the case of M Minkowski space-time that is the standard decomposition. The Hilbert space $\mathcal{H}$ is the completion of Γ with respect to the inner product (C.21).

There are two alternative but completely equivalent description of the 1-particle Hilbert space $\mathcal{H}$:

1. $\mathcal{H}$ consists of *real* valued functions (solution to the Maxwell equation for instance), equipped with the complex structure J. The inner product is given by (C.21).

2. $\mathcal{H}$ is constructed by complexifying the vector space Γ (tensoring with the complex numbers) and then decomposing it using J as described above. In this construction, the inner product is given by,

$$\langle \Phi, \tilde{\Phi} \rangle = \tfrac{i}{\hbar}\omega(\Phi^-, \tilde{\Phi}^+) \tag{C.22}$$

Note that in this case, the 1-particle Hilbert space consists of 'positive frequency' solutions.

It is important to note that the only input we needed in order to construct $\mathcal{H}$ was the complex structure J. For a general space-time there is no preferred one. This in turn leads to the infinite ambiguity in the representation of the CCR. In the case of stationary space-times there is a preferred, canonical, complex structure given by the Killing field. This construction for the case of the Klein Gordon field is described in Chapter 4 (see Eq. (4.29) for the explicit form of the complex structure). For Minkowski space-time there are several ways of characterizing the usual quantization. The standard textbook treatment uses a (globally inertial) time coordinate t to perform the positive-frequency decomposition. Another way of selecting this decomposition is to ask that the vacuum on the resulting theory be Poincaré invariant. A third way is to ask that the coherent states in the quantum theory have the same energy as the classical solution on which they are peaked [83].

C.2.1 Fock Space

We shall describe the universal construction of the Fock space associated to the Hilbert space $\mathcal{H}$ and then give in detail the representation for the Maxwell field in Minkowski space-time.

The *symmetric Fock space* associated to $\mathcal{H}$ is defined to be the Hilbert space

$$\mathcal{F}_{\rm s}(\mathcal{H}) := \bigoplus_{n=0}^{\infty}\left(\bigotimes^{n}{}_{\rm s}\mathcal{H}\right), \tag{C.23}$$

where we define the *symmetrized tensor product* of $\mathcal{H}$, denoted by $\bigotimes^{n}{}_{\rm s}\mathcal{H}$, to be the subspace of the n-fold tensor product $(\bigotimes^{n}\mathcal{H})$, consisting of totally symmetric maps $\alpha : \overline{\mathcal{H}}_1 \times \cdots \times \overline{\mathcal{H}}_n \to \mathbb{C}$ satisfying

$$\sum |\alpha(\bar{e}_{i_1}, \ldots, \bar{e}_{i_n})|^2 < \infty\,. \tag{C.24}$$

The Hilbert space $\overline{\mathcal{H}}$ is the *complex conjugate* of $\mathcal{H}$ with $\{\bar{e}_1, \cdots, \bar{e}_j, \cdots\}$ an orthonormal basis. We are also defining $\bigotimes^{0}\mathcal{H} = \mathbb{C}$.

We shall introduce the abstract index notation for the Hilbert spaces since it is most convenient way of describing the Fock space. Given a space $\mathcal{H}$, we can construct the spaces $\overline{\mathcal{H}}$, the complex conjugate space; $\mathcal{H}^*$, the *dual space*; and $\overline{\mathcal{H}}^*$ the dual to the complex conjugate. In analogy with the notation used in spinors, let us denote elements of $\mathcal{H}$ by ϕ^A, elements of $\overline{\mathcal{H}}$ by $\phi^{A'}$. Similarly, elements of $\mathcal{H}^*$ are denoted by ϕ_A and elements of $\overline{\mathcal{H}}^*$ by $\phi_{A'}$. However, by using Riesz lemma, we may identify $\overline{\mathcal{H}}$ with $\mathcal{H}^*$ and $\mathcal{H}$ with $\overline{\mathcal{H}}^*$. Therefore we can eliminate the use of primed indices, so $\overline{\phi}_A$ will be used for an element in $\overline{\mathcal{H}}^*$ corresponding to the element $\phi^A \in \mathcal{H}$. An element $\phi \in \bigotimes^{n}{}_{\rm s}\mathcal{H}$ then consists of elements satisfying

$$\phi^{A_1\cdots A_n} = \phi^{(A_1\cdots A_n)} \tag{C.25}$$

An element $\psi \in \bigotimes^{n}\overline{\mathcal{H}}$ will be denoted as $\psi_{A_1\cdots A_n}$. In particular, the inner product of vectors $\psi, \phi \in \mathcal{H}$ is denoted by

$$\langle\psi, \phi\rangle =: \overline{\psi}_A\phi^A \tag{C.26}$$

A vector $\Psi \in \mathcal{F}_{\rm s}(\mathcal{H})$ can be represented, in the abstract index notation as

$$\Psi = (\psi, \psi^{A_1}, \psi^{A_1 A_2}, \dots, \psi^{A_1 \dots A_n}, \dots)\,, \tag{C.27}$$

where, for all n, we have $\psi^{A_1 \dots A_n} = \psi^{(A_1 \dots A_n)}$. The norm is given by

$$|\Psi|^2 := \overline{\psi}\psi + \overline{\psi}_A \psi^A + \overline{\psi}_{A_1 A_2}\psi^{A_1 A_2} + \dots < \infty\,. \tag{C.28}$$

Now, let $\xi^A \in \mathcal{H}$ and let $\overline{\xi}_A$ denote the corresponding element in $\overline{\mathcal{H}}$. The *annihilation operator* $\mathcal{A}(\bar{\xi}) : \mathcal{F}_{\rm s}(\mathcal{H}) \to \mathcal{F}_{\rm s}(\mathcal{H})$ associated to $\overline{\xi}_A$ is denoted by

$$\mathcal{A}(\bar{\xi}) \cdot \Psi := (\overline{\xi}_A \psi^A, \sqrt{2}\,\overline{\xi}_A \psi^{A A_1}, \sqrt{3}\,\overline{\xi}_A \psi^{A A_1 A_2}, \dots)\,. \tag{C.29}$$

Similarly, the *creation operator* $\mathcal{C}(\xi) : \mathcal{F}_{\rm s}(\mathcal{H}) \to \mathcal{F}_{\rm s}(\mathcal{H})$ associated with ξ^A is defined by

$$\mathcal{C}(\xi) \cdot \Psi := (0, \psi\xi^{A_1}, \sqrt{2}\,\xi^{(A_1}\psi^{A_2)}, \sqrt{3}\,\xi^{(A_1}\psi^{A_2 A_3)}, \dots)\,. \tag{C.30}$$

If the domains of the operators are defined to be the subspaces of $\mathcal{F}_{\rm s}(\mathcal{H})$ such that the norms of the right sides of eqs. (C.29) and (C.30) are finite then it can be proven that $\mathcal{C}(\xi) = (\mathcal{A}(\bar{\xi}))^\dagger$. It may also be verified that they satisfy the commutation relations,

$$\left[\mathcal{A}(\bar{\xi}), \mathcal{C}(\eta)\right] = \bar{\xi}_A \eta^A\, \mathrm{I}\,. \tag{C.31}$$

A more detailed treatment of Fock spaces can be found in [84, 85, 86].

C.2.2 Representation of the CCR

In the previous section we saw that we could construct linear observables in (Γ, ω), in either of the constructions. For the covariant picture the observables are given by (C.5) and in the canonical by (C.19) and (C.20). This is the set $\mathcal{S}$ of observables for which there will correspond a quantum operator. Thus, for $\mathcal{O}[h] \in \mathcal{S}$ there is an operator $\hat{\mathcal{O}}[h]$. We want the *Canonical Commutation Relations* to hold,

$$\left[\hat{\mathcal{O}}[h], \hat{\mathcal{O}}[\tilde{h}]\right] = i\hbar\{\mathcal{O}[h], \mathcal{O}[\tilde{h}]\} = i\hbar\,\omega(h, \tilde{h})\,. \tag{C.32}$$

Then we should find a Hilbert space and a representation thereon of our basic operators satisfying the above conditions. We have all the structure needed at our disposal. Let us take as the Hilbert space the symmetric Fock space $\mathcal{F}_{\rm s}(\mathcal{H})$ and let the operators be represented as

$$\hat{\mathcal{O}}[h] \cdot \Psi := \hbar\left(\mathcal{C}(h) + \mathcal{A}(\overline{h})\right) \cdot \Psi\,. \tag{C.33}$$

Let us denote by h^A the abstract index representation corresponding to h_a in $\mathcal{H}$. First, note that by construction the operator is self-adjoint. It is straightforward to check that the commutation relations are satisfied,

$$\begin{aligned}
\left[\hat{\mathcal{O}}[h], \hat{\mathcal{O}}[h']\right] &= \hbar^2\left[\mathcal{C}[h], \mathcal{A}[\overline{h}']\right] + \hbar^2\left[\mathcal{A}[\overline{h}], \mathcal{C}[h']\right] \\
&= \hbar^2\left(\overline{h}_A h'^A - \overline{h}'_A h^A\right) \\
&= \hbar^2\left(\langle h, h'\rangle - \langle h', h\rangle\right) \\
&= 2i\hbar^2\,\mathrm{Im}(\langle h, h'\rangle) = i\hbar\,\omega(h, h')\,,
\end{aligned} \tag{C.34}$$

where we have used (C.31) in the second line and (C.21) in the last line. Note that in this last calculation we just used general properties of the Hermitian inner product and therefore we would get a representation of the CCR for *any* inner product $\langle\cdot,\cdot\rangle$.

C.2.3 Examples

Let us restrict our attention to Minkowski space-time and inertial hyper-surfaces Σ. Therefore, the induced metric h_{ab} is the Euclidean flat metric. We will perform two different decompositions of Γ, for two different complex structures. First, we shall consider the ordinary 'positive frequency' decomposition. This leads to the standard quantum theory of the free Maxwell field found in textbooks. This decomposition is relevant for chapter 2, since we are using that representation of the CCR to analyze the flux observables. Next, we decompose Γ in self-dual and anti-self-dual fields. This decomposition is of relevance to chapter 3, since we are using such a decomposition for the quantization of the Maxwell field.

Positive Frequency Decomposition

Since it is completely equivalent to use the covariant or canonical notation, we shall denote elements of Γ as pairs $(A^{\rm T}_a, E^a_{\rm T})$, of transverse (i.e. divergence-free) vector fields. The first step in the quantization is the introduction of the complex structure $J : \Gamma \to \Gamma$. It is given by,

$$J \cdot \begin{pmatrix} A_a \\ E_a \end{pmatrix} := \begin{pmatrix} -\Delta^{1/2} E_a \\ \Delta^{-1/2} A_a \end{pmatrix} . \tag{C.35}$$

Next, we can construct the *projector operator* $K^+ : \Gamma \to \Gamma_{\mathbb{C}}$, such that $F^+_{ab} = K^+(F_{ab})$ is the *positive frequency* part of $F_{ab} \in \Gamma$. The projector is given by the following action in terms of the pairs of initial data,

$$K^+ \cdot \begin{pmatrix} A_a \\ E_a \end{pmatrix} := \frac{1}{2} \begin{pmatrix} A_a - i\Delta^{-1/2} E_a \\ E_a + i\Delta^{1/2} A_a \end{pmatrix} . \tag{C.36}$$

With this definitions, we can construct the inner product in $\mathcal{H}$. For $F, \tilde{F}$ in $\mathcal{H}$ we have,

$$\begin{aligned} \langle F, \tilde{F} \rangle &= \frac{i}{\hbar} \omega(\overline{F}^+, \tilde{F}^+) \\ &= \frac{i}{\hbar} \int_\Sigma \mathrm{d}^3x (\overline{E}^{+a} \tilde{A}^+_a - \tilde{E}^{+a} \overline{A^+}_a) \\ &= \frac{i}{4\hbar} \int_\Sigma \mathrm{d}^3x \, \big[(E^a \tilde{A}_a - \Delta^{1/2} A^a \Delta^{-1/2} \tilde{E}_a - \tilde{E}^a A_a + \Delta^{1/2} \tilde{A}^a \Delta^{-1/2} E_a) \\ &\quad -i(\tilde{A}_a \Delta^{1/2} A^a + E^a \Delta^{1/2} \tilde{E}_a + A_a \Delta^{1/2} \tilde{A}^a + \tilde{E}^a \Delta^{-1/2} E_a) \big] . \end{aligned} \tag{C.37}$$

The norm of $(g_a, f^a) \in \mathcal{H}$ is given by,

$$\langle (g,f),(g,f) \rangle = \frac{1}{2\hbar} \int_\Sigma \mathrm{d}^3x \, \big(g_a \Delta^{1/2} g^a + f^a \Delta^{-1/2} f_a) \big) . \tag{C.38}$$

One should keep in mind that all the objects (g_a, f^a) are transverse. The reason for this requirement is that the complex structure takes a very simple form (C.35) in terms of transverse vector fields, making also the expression for the norm look simple (C.38).

We are now in position of asking whether an observable generated by the pair (g_a, f^a) induces a well defined operator on $\mathcal{F}_s(\mathcal{H})$. Clearly, if the pair (g_a, f^a) belongs to the 1-particle Hilbert space $\mathcal{H}$ the answer is in the affirmative. We shall take this criteria also as necessary condition. The question is now whether the pair (g_a, f^a) defines an element of Γ, namely, whether they are 'well behaved' initial data for a solution of Maxwell equations with finite norm. This will be the case iff the norm of (g_a, f^a), given by Eq. (C.38), is finite. In Chapter 2 we have an example of such observables given by the fluxes of electric and magnetic field across surfaces bounded by closed loops.

Self-dual Decomposition

As we mentioned in the last section, one can define the dual tensor to the electro-magnetic field tensor F_{ab}, by ${}^*F_{ab} := \frac{1}{2}\epsilon_{abcd}F^{cd}$. Note that if we apply the duality $*-$operator again we get:

$$\begin{aligned} {}^*({}^*F_{ab}) &= \tfrac{1}{4}\epsilon_{abcd}\epsilon^{cdef}F_{ef} \\ &= -F_{ab}\,, \end{aligned} \tag{C.39}$$

since $\epsilon_{abcd}\epsilon^{cdef} = -4\delta^e_{[c}\delta^f_{d]}$. Therefore, the $*-$operator defines a complex structure J on Γ. Note that this structure is available for any 4-dimensional Lorentzian manifold (M, g_{ab}) without the need to introduce extra structure. As discussed above, the $*-$operation decomposed the complexification of Γ into eigenspaces with eigenvalues $\pm i$. The elements of $F^{\uparrow}_{ab}$ of $\Gamma_{\mathbb{C}}$ such that ${}^*F^{\uparrow}_{ab} = iF^{\uparrow}_{ab}$ are called *self-dual*; and those that satisfy ${}^*F^{\downarrow}_{ab} = -iF^{\downarrow}_{ab}$ are *anti-self-dual*. The corresponding projector is given by,

$$K^{\uparrow}{}_{ab}{}^{cd} = \tfrac{1}{2}(\delta^c_{[a}\delta^d_{b]} - i\epsilon_{ab}{}^{cd})\,. \tag{C.40}$$

Therefore, the self-dual electro-magnetic field is of the form: $F^{\uparrow}_{ab} = \frac{1}{2}(F_{ab} - i{}^*F_{ab})$. In terms of objects defined on the hyper-surface Σ, namely electric and magnetic fields, a self dual element is of the form $E_a - iB_a$. Let us now write the projector $K^{\uparrow}$ acting on the pairs (A_a, E^a),

$$K^{\uparrow}\cdot\begin{pmatrix} A_a \\ E^a \end{pmatrix} = \frac{1}{2}\begin{pmatrix} A_a + id_a \\ E^a - iB^a \end{pmatrix}\,, \tag{C.41}$$

where d_a is the *electric vector potential*, i.e., such that $E^a = \epsilon^{abd}\partial_b d_c$.

Finally, we could follow the same steps as in the previous case and write the 'norm' in the 1-particle Hilbert space constructed from the $*-$operator decomposition as follows,

$$\langle (A,E),(A,E)\rangle = -\frac{1}{2\hbar}\int_{\Sigma} \mathrm{d}^3x\,(E^a d_a + A^a B_a)\,. \tag{C.42}$$

Note that this norm, in contrast to the positive frequency decomposition case, is not *positive definite*, and is therefore, physically incorrect. In math jargon, one says that the complex structure defined by the $*$-operator is not compatible with the simplectic structure. If one were to quantize naively this "Hilbert space", one would get a Fock representation with negative norm states. A holomorphic quantization with a positive definite inner product was constructed in [41], and the corresponding loop representation is the subject of Chapter 3.

D

Segal-Bargmann-Hall Transform for U(1)

In this appendix we construct in (gory) detail the Hall transform for the group U(1). Recall that a Holomorphic quantum theory for the harmonic oscillator can be constructed using a transform from the Hilbert space of square integrable functions on the real line to the space of holomorphic functions on $\mathbb{C}$ with a Gaussian measure. This is the so called Segal-Bargmann transform. There exists a similar transform for compact gauge groups, defined recently by Hall [68]. It is a generalization of the Segal-Bargmann transform and maps the Hilbert space of normalized functions on a compact, connected, Lie-group K with respect to the (left and right invariant) Haar measure to the Hilbert space of Holomorphic functions on the complexification of the Lie-group. The transform gives also a prescription for the measure in the complexified group and provides an isometric isomorphism of Hilbert spaces.

Let us denote by K the compact Lie-group and $x \in K$ an element of it. An element of G, the complexified group will be denoted by g. Let us now define the mapping $C_t : L^2(K, \mathrm{d}x) \to \mathcal{H}(G)$ as follows,

$$C_t(\phi)(g) = \int_K \phi(x)\rho_t(x^{-1}g)\,\mathrm{d}x \qquad \phi \in L^2(K, \mathrm{d}x),\ g \in G\,. \tag{D.1}$$

We will now proceed to construct the kernel $\rho_t(x)$ of the transform. The first step involves the solution of the 'heat equation' on S^1. The heat equation is defined using the Laplace-Beltrami operator for a bi-invariant metric on the Lie algebra of U(1). In terms of a normalized left invariant vector field X^a on U(1) it is written as,

$$\Delta = X^2\,. \tag{D.2}$$

We can choose $X^a = \left(\frac{\partial}{\partial\theta}\right)^a$ and therefore, $\Delta = \frac{\mathrm{d}^2}{\mathrm{d}\theta^2}$. Now, let ρ_t denote the heat kernel for Δ, that is, the solution of

$$\frac{\mathrm{d}h(\theta,t)}{\mathrm{d}t} = \tfrac{1}{2}\,\Delta h(\theta,t)\,, \tag{D.3}$$

such that $h(\theta, t = 0) = \delta_H(\theta)$, the delta function with respect to Haar measure. The heat equation is then,

$$\frac{1}{2}\frac{\partial^2}{\partial\theta^2}h(\theta,t) = \frac{\partial}{\partial t}h(\theta,t)\,. \tag{D.4}$$

Expanding the function in "Fourier series" $h(\theta,t) = \sum_m a_m(t)\, e^{im\theta}$ we have,

$$\sum_m \left[-\frac{m^2}{2}a_m(t) - \dot{a}_m(t)\right]\, e^{im\theta} = 0\,, \tag{D.5}$$

where $\dot{a}_m := \frac{\mathrm{d}a_m}{\mathrm{d}t}$. This implies that $\frac{m^2}{2}a_m(t) = -\dot{a}_m(t)$ and therefore,

$$a_m(t) = A_m e^{-\frac{m^2}{2}t}\,, \tag{D.6}$$

with A_m constants to be determined. We now have to impose the initial value condition,

$$h(\theta, t=0) = \sum_m A_m e^{im\theta} \equiv \delta(\theta)\,. \tag{D.7}$$

This condition implies that $A_m = 1\ \ \forall\ m$. Hence, the heat kernel reads then

$$\rho_t(\theta) = \sum_m e^{-\frac{m^2}{2}t}\, e^{im\theta}\,. \tag{D.8}$$

If we denote, as before, an element of U(1) by $x = e^{i\theta}$ and the element of $\mathbb{C}^*$ by $z = e^{i\tilde{\theta}} = e^{p+i\theta}$ then we can define the analytic continuation of the heat kernel in the following way,

$$\rho_t(z) = \sum_m e^{-\frac{m^2}{2}t}\, e^{im\tilde{\theta}} = \sum_m e^{-\frac{m^2}{2}t}\, z^m\,. \tag{D.9}$$

We can now write the kernel of the transform as follows

$$\rho_t(x^{-1}z) = \sum_m e^{-\frac{m^2}{2}t}\, x^{-m}\, z^m\,, \tag{D.10}$$

and therefore for an arbitrary function $f(x) = \sum_n a_n x^n$ the transform will be,

$$\begin{aligned} C_t(f)(z) \;&:=\; \tfrac{1}{2\pi}\int_{S^1}\left(\sum_m e^{-\frac{m^2}{2}t}\, x^{-m}\, z^m\right)\left(\sum_n a_n x^n\right)\mathrm{d}\theta \\ &=\; \tfrac{1}{2\pi}\sum_m\sum_n a_n\, e^{-\frac{m^2}{2}t}\, z^m \int_{S^1} x^n x^{-m}\,\mathrm{d}\theta \\ &=\; \sum_m\sum_n a_n\, e^{-\frac{m^2}{2}t}\, z^m \delta_{mn} \\ &=\; \sum_m a_m\, e^{-\frac{m^2}{2}t}\, z^m\,. \end{aligned} \tag{D.11}$$

We have completed the construction of the Segal-Bargmann transform for U(1). The next step is to find the averaged heat kernel measure ν_t^c. But first we have to define the heat kernel in the complexified group $\mathbb{C}^*$. We define the Laplace-Beltrami operator for a inner-product defined in the complexification of the Lie algebra $\mathbb{R}$ seen as a (two-dimensional) real vector space. The operator is then,

$$\Delta_G = \frac{\partial^2}{\partial \theta^2} + \frac{\partial^2}{\partial p^2} \,. \tag{D.12}$$

Let μ_t denote the fundamental solution at the identity of the following equation on G:

$$\frac{\mathrm{d}h}{\mathrm{d}t} = \tfrac{1}{4}\Delta_G h \,. \tag{D.13}$$

The (real) heat kernel μ_t on G is different from the analytic continuation of the heat kernel on K. The solution is given by,

$$\mu_t(z) = \tfrac{1}{\sqrt{t\pi}} e^{-\frac{p^2}{t}} \left(\sum_m e^{-\frac{m^2 t}{4}} e^{im\theta} \right) \,. \tag{D.14}$$

Denoting $x = e^{i\theta'}$ we have,

$$\mu_t(xz) = \tfrac{1}{\sqrt{t\pi}} e^{-\frac{p^2}{t}} \left(\sum_m e^{-\frac{m^2 t}{4}} e^{im(\theta+\theta')} \right) \,. \tag{D.15}$$

The averaged heat kernel measure $\nu_t(z)$ is defined as

$$\nu_t := \tfrac{1}{2\pi} \int_{S^1} \mathrm{d}\theta \mu_t(xz) \,. \tag{D.16}$$

Therefore, it takes the form,

$$\begin{aligned} \nu_t &= \tfrac{1}{2\pi} \int_{S^1} \mathrm{d}\theta' \tfrac{1}{\sqrt{t\pi}} e^{-\frac{p^2}{t}} \left(\sum_m e^{-\frac{m^2 t}{4}} e^{im(\theta+\theta')} \right) \,, \\ &= \tfrac{1}{\sqrt{t\pi}} e^{-\frac{p^2}{t}} \,. \end{aligned} \tag{D.17}$$

References

[1] See for instance: T.W.B. Kibble *Commun. Math. Phys.* **65** 189 (1979);
A. Heslot, *Phys. Rev. D***31** 1341 (1985);
A. Ashtekar and T.A. Schilling, "Geometry of Quantum Mechanics" CAM-94 Physics Meeting, in AIP Conf. Proc **342**, 471, ed. Zepeda A (AIP Press, Woodbury, New York, 1995);
T.A. Schilling, "Geometry of Quantum Mechanics", (Ph.D. diss., The Pennsylvania State University, 1996);
L.P. Hughston, "Geometric Aspects of Quantum Mechanics" Chapter 6 *Twistor Theory* ed. S.A. Hugget pp. 59-79 (Marcel Dekker, New York, 1995)

[2] A. Sen, *J. Math. Phys.* **22**, 1781 (1981); A. Sen, *Phys. Lett* **11B**, 89 (1982).

[3] A. Ashtekar, *Phys. Rev. Lett.* **57**, 2244 (1986); A. Ashtekar, *Phys. Rev.* **D36**, 1587 (1987).

[4] There are now more than 600 papers related to this program. For a complete list see: C. Beetle and A. Corichi, 'Bibliography of publications related to classical and quantum gravity in terms of connections and loop variables'. preprint gr-qc/9703044.

[5] W.H. Thomson, *Trans. R. Soc. Edin.* **25**, 217 (1869).

[6] J.H. Przytycki, *Aportaciones Matemáticas, Comunicaciones* **11**, 173 (1992).

[7] M. Atiyah, *Publ. Math. Inst. Hautes Etudes Sci. Paris* **68**, 175 (1989); M. Atiyah, *The Geometry and Physics of Knots* (Cambridge University Press, Cambridge, 1990).

[8] E. Witten, *Commun. Math. Phys.* **117**, 353 (1988).

[9] E. Witten, *Commun. Math. Phys.* **121**, 351 (1989).

[10] S.K. Donaldson and P.B. Kronheimer, *Geometry of 4-Manifolds* (Clarendon Press, Oxford, 1990).

[11] L. Kauffman, *Knots and Physics* (World Scientific, Singapore, 1994).

[12] J.C. Baez and J. Munaian, *Gauge Fields, Knots and Gravity* (World Scientific, Singapore, 1994).

[13] For early references on the application of loop variables to gravity see: S. Mandelstam, *Ann. of Phys.* **19**, 1 (1962);

A. Polyakov, *Phys. Lett.* **82B**, 247 (1979); *Nucl. Phys.* **B164**, 171 (1979);
Y.M. Makenko and A.A. Migdal, *Phys. Lett* **88B**, 135 (1979);
L. Smolin, *Nucl. Phys.* **B148**, 333 (1979).

[14] P. Renteln and L. Smolin, *Class. Quantum Grav.* **6**, 275 (1989);

[15] T. Jacobson and L. Smolin, *Nucl. Phys.* **B299** 295, (1988).

[16] C. Rovelli and L. Smolin *Phys. Rev. Lett.* **61** (1988) 1155; C. Rovelli and L. Smolin *Nuc. Phys.* **B331** (1990) 80.

[17] See for instance: A. Polyakov, *Mod. Phys. Lett.* **A3**, 325 (1988);
L. Smolin, *Mod. Phys. Lett.* **A4**, 1091 (1989);
S. Majid, *J. Math Phys.* **31**, 924 (1990);
C.H. Tze, *Int. J. Mod. Phys.* **A3**, 1959 (1988).

[18] Y. Aharonov and J. Anandan, *Phys. Rev. Lett.* **58**, 1593 (1987).

[19] M.V. Berry, Proc. R. Soc. Lond. **A392**, 45 (1984).

[20] A. Achúcarro and P.K. Townsend, *Phys. Lett.* **B180** (1986) 89;

[21] E. Witten, *Nucl. Phys.* **B311** (1988) 46.

[22] A. Ashtekar and J. Romano, *Phys. Lett.* **B229** (1989) 56;
A. Ashtekar, V. Husain, C. Rovelli, J. Samuel and L. Smolin, *Class. Quantum Grav.* **6** (1989) L185;
L. Smolin, in the *Proceedings of the John's Hopkins Conference on Knots, Topology and Quantum Field Theory*, ed. L. Lusanna (World Scientific, Singapore, 1989).

[23] C. F. Gauss, *Werke* vol V (Götingen, Königliche Gesellschaft der Wissenschaften, 1833) 605, Note of January 22.

[24] J. C. Maxwell, *A Treatise on Electricity and Magnetism* (Oxford, Clarendon, England, 1873), p 40-42.

[25] R. Gambini and J. Pullin, *Loops, Knots, Gauge Theories and Quantum Gravity* (Cambridge University Press, Cambridge, England, 1996).

[26] I. Bialynicki-Birula and Z. Bialynicka-Birula, *Quantum Electrodynamics* (Pergamon, Oxford, England, 1975).

[27] A. Ashtekar, Mathematical problems of non-perturbative quantum general relativity, in *Proceedings of the 1992 Les Houches School on Gravitation and Quantizations* ed B. Julia and J. Zinn-Justin (Elsevier, Amsterdam, 1995).

[28] A. Ashtekar, J. Lewandowski, D. Marolf, J. Mourão, T. Thiemann. *J. Math. Phys.* **36**, 6456 (1995).

[29] A. Widom and T.D. Clark, *Phys. Lett.* **90A**, 280 (1982).

[30] A. Ashtekar and A. Corichi, "Gauss linking number and electro-magnetic uncertainty principle", *Phys. Rev.* **D56**, 2073 (1997).

[31] A. Ashtekar and A. Magnon, *Proc. R. Soc.* (London) **A46**, 375 (1975).

[32] B. S. Kay, *Commun. Math. Phys.* **62**, 55 (1978)

[33] R. Gambini and A. Trias, *Phys. Rev.* **D22**, 1380 (1980); C. Di Bartolo, F. Nori, R. Gambini and A. Trias, *Lett. Nuov. Cim.* **38**, 497 (1983).

[34] A. Ashtekar and C. Rovelli, *Class. Quantum Grav.* **9** 1121, (1992).

[35] A. Ashtekar, *Lectures on Non-perturbative Canonical Gravity (Notes prepared in collaboration with R S Tate)* (World Scientific, Singapore, 1991)

[36] A. Ashtekar and A. Corichi, *Class. Quantum Grav.* **14**, A43 (1997).

[37] A. Ashtekar and J. Lewandowski, Representation theory of analytic holonomy C^* algebras *Knots and Quantum Gravity* ed J Baez (Oxford University Press, Oxford, 1994)

[38] A. Ashtekar, C. Rovelli and L. Smolin, *Phys. Rev.* **D44**, 1740 (1991).

[39] A. Ashtekar and R.S. Tate, *J. Math. Phys.* **35**, 6434 (1994).

[40] A. Rendall, *Class Quantum Grav.* **10**, 2261 (1993).

[41] A. Ashtekar, C. Rovelli and L. Smolin, *J. Geom. Phys.* **8**, 7 (1992).

[42] J.S. Dowker, *Nuovo Cimento* **B52**, 129 (1967).

[43] C.J.C. Burges, *Phys. Rev.* **D32**, 504 (1985);
V.B. Bezerra, *Phys. Rev.* **D35**, 2031 (1987); *J. Math. Phys.* **30**. 2895 (1989); *Ann. of Phys.* **203**, 392 (1990).

[44] S. Deser, *Phys. Rev. Lett.* **64**, 611 (1990);
S. Deser and J.G. McCarthy, *Nucl. Phys.* **B344**, 747 (1990);
Y.M. Cho, D.H. Park and C.G. Han, *Phys. Rev* **D43**, 1421 (1991);
M. Reuter, *Phys. Rev* **D44**, 1132 (1991);
M.E. Ortiz, *Nucl. Phys.* **B363**, 185 (1991).

[45] A. Ashtekar and A. Magnon, *J. Math. Phys.* **16**, 341 (1975);
J. Anandan, *Phys. Rev.* **D15**, 1448 (1977); **D24**, 338 (1981);
D. Harari and A.P. Polychronakos, *Phys. Rev.* **D38**, 3320 (1988).

[46] P.S. Gerbert and R. Jackiw, *Commun. Math. Phys.* **124**, 229 (1989).

[47] B.S. DeWitt, *Phys. Rev. Lett.* **16**, 1092 (1966).

[48] G. Papini, *Nuovo Cimento* **B52**, 136 (1967).

[49] J. Anandan, *Gen. Rel. Grav.* **26**, 125 (1994).

[50] A. Corichi and M. Pierri, *Phys. Rev.* **D51**, 5870 (1995).

[51] A. Mostafazadeh, "Adiabatic geometical phase for scalar fields in a curved space-time", preprint hep-th/9608051.

[52] P.O. Mazur, *Phys. Rev. Lett.* **57**, 929 (1986).

[53] S. Deser, R. Jackiw and G. t'Hooft, *Ann. of Phys.* **152**, 220 (1984).

[54] M.A. Peshkin and A. Tonomura, *The Aharonov-Bohm effect*, (Springer Verlag, New York, 1989), LNP v. 340.

[55] Y. Aharonov and D. Bohm, *Phys. Rev.* **115**, 485 (1959).

[56] J. Samuel and R. Bandhari, *Phys. Rev. Lett.* **60**, 2339 (1988).

[57] J. Anandan and P.O. Mazur, *Phys. Lett.* **A173**, 116 (1993).

[58] R. Loll, *J. Math. Phys.* **36** (1995) 6494.

[59] H-J. Matschull, *Class. Quantum Grav.* **12**, 2621 (1995).

[60] S. Carlip and J. E. Nelson, *Phys. Rev* **D51**, 5643 (1995).

[61] S. Carlip, *Journal of the Korean Physical Society (Proc. Suppl.)* **28**, S447 (1995).

[62] S. Carlip, *Class. Quantum Grav.* **12**, 2853 (1995).

[63] A. Ashtekar, *Phys. Rev* **D53**, 2865 (1996).

[64] T. Thiemann, *Class. Quantum Grav.* **13**, 1383 (1996).

[65] T. Thiemann, *Phys. Lett.* **380B**, 257 (1996).

[66] J.D. Romano, *Gen. Rel. Grav.* **25**, 759 (1993).

[67] J. Louko and D. M. Marolf, *Class. Quantum Grav.* **11**, 311 (1994).

[68] B. C. Hall, *J. Functional Analysis* **122**, 103 (1994).

[69] R. Loll, *Class. Quantum. Grav.* **12**, 1655 (1994).

[70] J. E. Nelson and T. Regge, *Nucl. Phys.* **B328**, 190 (1989); *Class. Quantum Grav.* **9**, 187 (1992); *Phys. Lett.* **B272**, 213 (1991).

[71] J. E. Nelson, T. Regge and F. Zertuche, *Nucl. Phys.* **B339**, 516 (1990).

[72] J. E. Nelson and T. Regge, *Commun. Math. Phys.* **141**, 211 (1991); *Commun. Math. Phys.* **155**, 561 (1993); *Phys. Rev.* **D50**, 5125 (1994).

[73] K. Ezawa, *Phys. Rev.* **D49**, 5211 (1994); addendum in *Phys. Rev.* **D50**, 2935 (1994).

[74] G. Mess, "Lorentz Spacetimes of Constant Curvature," Institut des Hautes Etudes Scientifiques preprint IHES/M/90/28 (1990).

[75] E. Witten, *Commun. Math. Phys.* **137**, 29 (1991).

[76] A. Ashtekar and A. Corichi, in preparation.

[77] C.J. Isham, *Modern Differential Geometry for Physicists* (World Scientific, Singapore, 1989).

[78] Y. Choquet-Bruhat, C. DeWitt-Morette and M. Dillard-Bleick, *Analysis, Manifolds and Physics* (North Holland, Amsterdam, 1982).

[79] R. Penrose, in *Battele Recontres*, ed C. DeWitt and J.A. Wheeler, (Benjamin, New York, 1968).

[80] R. Penrose and W. Rindler, *Spinors and Space-time* (Cambridge University Press, Cambridge, 1984).

[81] R.M. Wald, *General Relativity* (Chicago University Press, Chicago, 1984).

[82] A. Ashtekar, G.T. Horowitz and A. Magnon-Ashtekar, *Gen. Rel. Grav.* **14**, 411 (1982).

[83] A. Ashtekar and A. Magnon-Ashtekar, *Pramana* **15**, 107 (1980).

[84] R. Geroch, "Special Problems in Particle Physics", (unpublished).

[85] R.M. Wald, *Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics* (Chicago University Press, Chicago, 1994).

[86] M. Reed and B. Simon, *Functional Analysis* (Academic Press, London, 1980).

Vita

Alejandro Corichi Rodriguez Gil was born in Mexico City in 1967. He attended private schools for his elementary education. Later, he joined the Universidad Nacional Autónoma de México (UNAM), in the Escuela Nacional Preparatoria for the last three years of high-school. He enrolled in the Faculty of Science of UNAM in October 1985. He completed the physics courses at UNAM in 1990, and became then associated to the Instituto de Ciencias Nucleares (ICN, UNAM), where he wrote his B.S. thesis. He defended it and was conferred the degree of "Fisico", with high class honors, in october 1991. He attended Syracuse University from 1992 to 1993. In August 1993 he enrolled in the physics Ph.D. program at The Pennsylvania State University.

Since 1988 he has been awarded fellowships. From 1988 to 1991 he received support from the DGAPA, UNAM for completing undergraduate studies. From 1992 to 1997 he was given a Ph.D. fellowship also from DGAPA, UNAM. In 1991 he received the "Gabino Barrera Medal" for the best graduating student in physics. In 1991 he was awarded the "Bruno Gonzalez Fellowship" by the Mexican Physical Society (SMF) and the European Laboratory for Particle Physics (CERN). In 1994 The United States Achievement Academy conferred on him the honor of All-American Scholar. Since 1995 he has been part of the Sistema Nacional de Investigadores (SNI), México, as "Candidato a Investigador Nacional".

He has written the following articles: A. Corichi and D. Núñez, "Introduction to the ADM Formalism", *Rev. Mex. Fis.*, **37**, 720, 1991; A. Corichi, "Comment on Hamiltonian Formulation of a Theory with Constraints", *J. Math. Phys.*, **33**, 4066, 1992; A. Corichi and M. Pierri, "Gravity and Geometric Phases", *Phys. Rev.* **D51**, 5870, 1995; A. Corichi and M.P. Ryan, Jr., "Quantization of Nonstandard Hamiltonian Systems", *J. Phys. A: Math. and Gen.*, 1997 (in press); A. Ashtekar and A. Corichi "Gauss Linking Number and the Photon Norm", *Class. Quantum Grav.* **14**, A45, 1997; A. Corichi and J.A. Zapata "On Diffeomorphism Invariance for Lattice Theories", *Nucl. Phys.* **B**, 1997 (in press); A. Ashtekar and A. Corichi "Gauss Linking Number and Electro-magnetic Uncertainty Principle", Preprint hep-th/9701136; A. Corichi and K. Krasnov, "Loop Quantization of Maxwell Field and Electric Field Quantization" Preprint hep-th/9703177; A. Ashtekar and A. Corichi "3-dimensional Gravity with a Cosmological Constant", in preparation; A. Ashtekar, A. Corichi, J. Lewandowski and J.A. Zapata "Quantum Theory of Geometry II: Non-commutativity of Riemannian Structures", in preparation; A. Ashtekar, J. Baez, A. Corichi and K. Krasnov "Quantum Mechanical Origin of Black Hole Entropy", in preparation.