The polygon representation of three dimensional gravitation and its global properties

The polygon representation of three dimensional gravitation and its global properties

De polygoonrepresentatie van de zwaartekracht-theorie in drie dimensies en haar globale eigenschappen

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht, op gezag van de Rector Magnificus, Prof. dr. W. H. Gispen, ingevolge het besluit van het College voor Promoties in het openbaar te verdedigen op maandag 18 april 2005 des middags te 14.30 uur

door

Zoltán Kádár

geboren op 26 oktober 1976 te Boedapest, Hongarijë

Promotor: Prof. dr. G. 't Hooft
Copromotor: dr. R. Loll

Instituut voor Theoretische Fysica en
Spinoza Instituut
Universiteit Utrecht

ISBN 90-393-2698-3

Contents

Preface

There are four fundamental interactions in nature. The electromagnetic force describes not only electric and magnetic fields, but also optical phenomena and other forms of radiation. The strong force, active in atomic nuclei, keeps the quarks together in protons and neutrons. Finally, the weak force is responsible for phenomena such as the beta decay of neutron into a proton while emitting an electron and a neutrino in some radioactive nuclei.

These first three forces dominate small scales from millimeters to 10^{-16} millimeters. Together they describe the dynamical features of the so-called Standard Model. This theory requires a full use of *quantum mechanics*, which is very different from classical physics, like Newtonian mechanics, Maxwell's electromagnetism or general relativity. For example, it is an essential feature of quantum mechanics, that simultaneous measurement on all observables of a system cannot be made with arbitrary precision, even in principle. A classical theory is deterministic: knowing all initial conditions of a system enables us to predict its state in a later time exactly. A quantum theory is non-deterministic, it predicts probabilities for the state of a system at a later time.

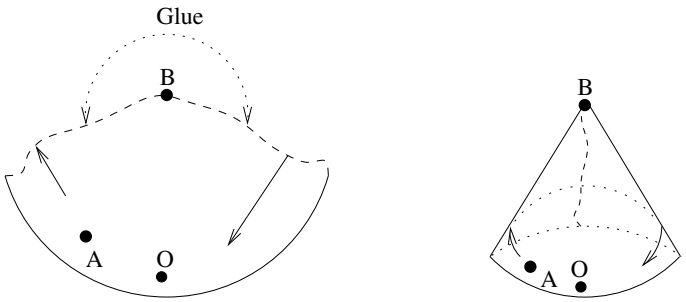
For the fourth force, that of gravitation, we only have a classical theory called General Relativity. General relativity is an extremely successful model of gravitation in four dimensions agreeing with all experimental data from cosmological distances to millimeter scales. So there is a classical theory describing gravitation on macroscopic scales and a quantum theory describing matter on microscopic scales. No deviations from their predictions have ever been observed, however, this situation cannot be logically consistent. There is one scale, called the Planck length, where effects coming from all theories are expected to be important. At this scale all forces should act together and therefore be in harmony with one another. It has been a long-standing open issue to unify the standard model and general relativity or to find the quantum theory of gravitation, which would predict new physics at that scale. The value of the Planck length is 10^{-32} millime-

ter, which is 16 orders of magnitude smaller than the best resolution we can achieve in modern particle physics, where large underground particle accelerators are being used to probe small scales using particles with the highest available energy. The Planck regime is therefore inaccessible using today's technology. It may seem that we do not need a quantum gravity theory and there is no experimental evidence for its existence. However, we do have theoretical evidence that such a theory should exist.

In quantum mechanics the position x and the momentum p of a particle cannot be measured with arbitrary precision. The more accurately its position is determined, the less precise our knowledge is about its momentum and vice versa. The uncertainty around a sharp, exact value of the location of the particle is Δx and the uncertainty of its momentum is Δp. According to quantum mechanics the product $\Delta x \cdot \Delta p$ cannot be smaller then $\hbar$, a fundamental constant in quantum mechanics. On macroscopic scales it is a tiny number: for an object of weight 1 kg moving with 1 meter/s, the uncertainty is around 10^{-34} meter. Classical general relativity nevertheless can be used to measure both x and p precisely which would be a contradiction. This indicates that there is an inconsistency of the combined theory of quantum mechanics of matter and classical general relativity as a fundamental theory of nature. There *has to* be a quantum gravity theory, which reduces to general relativity on macroscopic scales. Among others it should explain the formation and evaporation of black holes and the physics of the early universe.

Physicists have been trying to construct quantum gravity since the 60's. The situation is a bit different from the birth of fundamental theories of physics in the last couple of centuries. At that time, modifications about the current view about nature were forced upon us by numerous experimental observations. Now, for quantum gravity, the only guiding principle in the construction of the theory is that it must reduce to general relativity in the macroscopic regime. We can guess a great deal about features of quantum gravity, but the full theory itself has not yet been not found. To summarize the fundamental difficulty: the conventional quantized particle theories are formulated in a given spacetime, whereas in general relativity spacetime is an outcome of the dynamics of the theory. In general relativity, spacetime is determined or created by the type and the distribution of matter. How to reconcile these two fundamentally different approaches is a deep and fundamental problem.

In modern science, if a problem turns out to be too difficult to solve, one can try to solve a simpler model of similar nature, with the intention to generalize it to the original later. In this thesis, we have chosen to reduce the dimensionality of spacetime from four to three. That is to say,



the number of space dimensions went from three to two. We studied gravitating point-like particles moving in two space dimensions. For example, we have a sheet of paper, or a surface, like the surface of the Earth or a donut. The point-like particles have no size, but they have masses so they interact gravitationally. There are no such objects in the four dimensional world, since, if we try to "squeeze" matter in a too small region, it becomes a black hole. A black hole, however, is an extended object with a horizon separating its "inside" from the rest of the world. In other words, black holes have a finite size. This, however, cannot happen in three dimensions, where particles can be truly point-like.

Furthermore, what is peculiar about particles living in three dimensions, is that they do not really feel each other in a way planets do in nature, they follow straight lines. Both space and time surrounding a point particle in three dimensions are locally flat. However, a particle creates a "cone" around itself. Consequently, the locally straight trajectory of a particle may seem to be bent when regarded from a distance. This phenomenon is illustrated in the figure on the top of this page. The right side of the figure can be reproduced from the left with scissors and glue as indicated. We are at point O in the "two dimensional sea" and look to the left towards a small ship A. There is a heavy tanker at point B. We see that the ship is going straight and we see it from behind, but looking to the right we see the front of the ship far away coming towards us! This was an illustration of the effect caused by a point-like massive particle B when Einstein's general relativity is applied in three dimensions.

An important motivation for studying three dimensional quantum gravity is that this theory is classically soluble. Quantization of classical systems has been studied and a general experience is that if the classical system is exactly solvable, then so is the quantized system. One can say, that if the space of classical solutions is known, then quantization becomes easier. This is a motivation to study the space of the solutions of three dimensional gravity. Apart from conventional physical motivations, it is also highly in-

teresting by itself as a mathematical problem.

Given some initial configuration, the Einstein equations of general relativity determine all of spacetime. There is a great arbitrariness in choosing this initial configuration, because there is no absolute notion of time in relativity. Suppose that the initial system were a disc with a circle at its boundary. Then spacetime would be similar to a solid cylinder as we view how the disc evolves in time. In other words, every slice of the cylinder which is parallel to the initial disc, shows the disc at different moments of time. However, one could slice this cylinder not only parallel to an initial disc but also slightly askew. This would correspond to a different choice of the time coordinate. Then the same spacetime would be represented as a collection of ellipses. One feature remains universal. The notion of an event A being *before* B means that A can send a message to B and this signal can travel at most with the speed of light. We then say that A and B are causally connected. Whenever we give an initial slice, it is not allowed to have any pair of points in it that are causally connected. Apart from this restriction, we are free to choose the initial slice. It does not have to be an ellipse, it can be a complicated surface sliced from that cylinder.

In the polygon representation of three dimensional gravity, the time slice is a collection of polygons in such a way that every edge of a polygon is identified with another edge of another (or the same) polygon. In other words, if an observer leaves a polygon, she reappears in another. The particle we described in the beginning easily fits in this picture: a corner of a polygon is a point particle if the two edges it separates are identified. Then, there would be a cone around that corner if we glued the identified edges together as it was shown in the beginning. This simple model is the main topic of this thesis. It can be simulated on a computer. It contains striking features: big bang, chaos and mathematical richness. Furthermore, constructing and studying this simple system made physicists come to important qualitative conclusions concerning real four dimensional quantum gravity. For the brave and the expert, a notable reference is [1].

Chapter 1

Introduction to 2+1 gravity

The quest for a theory of quantum gravity has been pursued in many approaches in mathematical physics. Apart from direct proposals, which may be tested in the future by experiments, numerous branches of research focus on certain problems in formulating the theory in simplified settings. Results and experience coming from toy models often generalize to the physically more interesting cases and help us to identify methods and general features of the system that we wish to understand.

2+1 gravity is one important example of such a toy model. It is much simpler than the realistic four dimensional case. It teaches us that discreteness can result from quantization without breaking Lorentz invariance. It gives us an indication how, at least in principle, a quantization procedure for the gravitational force might be set up.[1] It teaches us that timelike and spacelike quanta of geometry can have different spectra, see section 3.3.

The reason why gravity in three dimensions is much simpler than in four is the absence of local degrees of freedom. We can see this by means of a counting argument. The number of phase space degrees of freedom is given by the number of independent components of the dynamical variables, the spatial metric and its conjugate momentum, minus the sum of the number of independent constraints and symmetries of the theory. For gravity in d dimensions, both the spatial metric and its conjugate momentum have $d(d-1)/2$ independent components. d components of the (Einstein) equations of motion contain no time derivatives, they are constraints among the independent variables, and there are d reparametrization symmetries of the coordinates also known as diffeomorphism invariance. The number of local degrees of freedom is thus

$$d(d-1) - 2d = d(d-3), \tag{1.1}$$

[1] There has been substantial progress for the simplest case of the torus topology.

which is zero for the case of $d = 3$. One can also argue in the following way. In three dimensions the Ricci tensor $R_{\mu\nu}$ determines the curvature tensor $R_{\mu\nu\sigma\rho}$ completely. The consequence of this and the vacuum Einstein equations $2R_{\mu\nu} = -\Lambda g_{\mu\nu}$, where $g_{\mu\nu}$ is the *spacetime* metric, is that the spacetime has constant curvature (proportional to the cosmological constant Λ). Phrasing it in physical terms: there are no gravitational waves in three dimensions.

Nevertheless, the theory is nontrivial due to global degrees of freedom and is worth studying. Let us give a short list of arguments why.

- Physical degrees of freedom may come from nontrivial topology of space, but also from objects like point particles, which can be considered as a special limit of matter fields. A point particle is a point-like mass, a naked singularity in three dimensions, which creates a conical geometry around itself. This will be explained in more detail in the next chapter. Note that besides the point particle, the Einstein equations also have black hole solutions: geometries, where the singularity is behind a horizon. These solutions, however, only exist if the cosmological constant $\Lambda < 0$.

- The model is classically soluble. Due to developments in the field of three dimensional geometry and topology [2], we can enumerate all possible manifolds which are solutions of the Einstein equations. Furthermore, we can explicitly reduce the infinite dimensional phase space to a *reduced phase space* of finite dimensions. Quantization of a finite dimensional phase space is quantum mechanics rather than quantum field theory, hence, considerably simpler.

- Viewed as a field theory defined by the Einstein-Hilbert action

$$I[g_{\mu\nu}] = \frac{1}{16\pi G}\int d^3x\sqrt{-g}(R - 2\Lambda), \tag{1.2}$$

  where G is Newton's constant, 2+1 gravity is non-renormalizable, since Newton's constant has the dimension of a length (in units $c = \hbar = 1$).[2]

- Many conceptual problems are inherited from four dimensional gravity: The observables are invariants under diffeomorphisms, in particular time translations. They are thus constants of motion. Where

[2] If the first order formalism is used, this problem seems to disappear, but we have not been able to apply it in the presence of matter [3].

to find the dynamics and what is the appropriate *time* coordinate with respect to which one should formulate the dynamics, are difficult questions. One hopes to be able to answer them first in the simplified setting of 2+1 gravity.

- In the canonical formalism of three dimensional gravity, there are no second class constraints. These are present in a general constrained Hamiltonian system, if only a subset (called first class) of the Poisson algebra of constraints closes. If this is the case, before quantization, the Poisson bracket has to be replaced with the Dirac bracket [4] or the second class constraints have to be solved by introducing a new, smaller set of dynamical variables to the system. In four (and higher) dimensional gravity this causes major complications.

- Due to the absence of local degrees of freedom, the theory is topological, which is manifest in the first order formalism as we will see in the next section. Among others, this means that one can discretize spacetime, e.g. by means of triangulation, while still keeping the exact set of dynamical variables.

In the rest of this chapter we shall review the classical approaches concentrating mostly on the simplest sector of the theory, the case when the cosmological constant vanishes. Since classical general relativity in $2+1$ dimensions is well understood[3], the different approaches to the theory are developed in order to serve as starting points for the quantum theory. Apart from a few remarks, we will concentrate on the classical formulations and postpone the aspects of quantization to the last chapter.

Two influential publications appeared at the end of the 80's. In [5] Deser, Jackiw and 't Hooft wrote down the generic one-particle solution to the Einstein equations for $\Lambda = 0$. They started from a static ansatz for the metric and the expression for the stress energy tensor of a static point-like source given by $T_{00} = m\,\delta^{(2)}(\vec{r} - \vec{r}_{(0)})$, and solved the Einstein equations. They also computed the solution for a stationary, axially symmetric ansatz for the metric and derived the angular momentum. As a combination of these results the line element of a spinning, massive particle reads:

$$ds^2 = -(dt + 4Gsd\phi)^2 + dr^2 + (1 - 4Gm)^2 r^2 d\phi^2 \ . \qquad (1.3)$$

After the transformations $t' = t + 4Gs\phi$ and $\phi' = (1 - 4Gm)\phi$, the form of the above metric is Minkowskian everywhere, except at the origin. Furthermore, there is a cusp stretching from the origin with an unusual matching

[3]There are many unsolved questions though; we will see some examples in the next chapter.

condition

$$(t + 8\pi G s,\, r,\, \phi + 8\pi G m) = (t,\, r,\, \phi)\ . \tag{1.4}$$

This geometry is of a "helical cone", its deficit angle at the tip is $8\pi G m$ and the time shift is $4Gs$. We will discuss this geometry in more detail in the beginning of the next chapter.

The second important result is that of references [6, 3], where it was proven that the pure gravity sector is equivalent to a Chern-Simons theory with a Poincaré ($ISO(2,1)$), de Sitter ($SO(3,1)$) or anti de Sitter ($SO(2,2)$) gauge group depending on the sign of the cosmological constant Λ. In order to explain this result and for later convenience we shall now present a short introduction to the first order formulation of gravity.

1.1 First order formalism

This formulation of gravity in arbitrary dimensions is long known and widely used in the quantum gravity community, see [7] for an early paper on the topic. In this approach to gravity there are two sets of dynamical variables. One is given by the vielbein, which is locally an $so(d-1,1)$ Lie algebra valued one-form. The second field is an $SO(d-1,1)$ connection.[4] Hence, the theory resembles a gauge theory and in three dimensions it is in fact a gauge theory, as we will see.

The first order formalism is usually considered as classically equivalent to the second order formulation of gravity in terms of the metric. Whether this equivalence continues to hold for the quantized theory depends on how the quantization is done. There may well exist different, inequivalent quantum versions. The issue underlying this opinion is that the phase space of the first order formalism is larger than that of the second order formalism: it contains configurations, which correspond to degenerate metrics. The Einstein equations can only be recovered from the first order formalism, when the metric is non-degenerate. Note also that the symmetry algebra of the metric formulation, *i.e.* the diffeomorphism algebra, is contained in the gauge symmetries of the first order formalism only on shell. Furthermore, finite diffeomorphisms, which are not connected continuously to the identity map, are not symmetries of the first order approach, if we think of it as a gauge theory [8].

Let us now explicitly define the fundamental variables and explain how to interpret them physically. The vielbein is denoted by e^a_μ, where μ is a spacetime index and a is a flat Minkowski index. The vielbein is interpreted

[4] In the case of Riemannian gravity, $SO(d-1,1)$ is replaced by $SO(d)$.

as a local observer: in his Lorentz frame at the point x the spacetime vector $v^\mu(x)$ is measured as $v^a(x) = e^a_\mu(x)v^\mu(x)$. The metric in terms of the vielbein reads

$$g_{\mu\nu} = e^a_\mu \, e^b_\nu \, \eta_{ab} \, , \tag{1.5}$$

with η_{ab} being the Minkowski metric. We adopt the convention of using the signature $(-++\ldots)$ for the Minkowski metric η_{ab}. It is used to lower and raise flat indices labeled by latin letters $a, b = 0, 1, 2\ldots$ from the beginning of the alphabet. Greek indices stand for spacetime indices and latin indices i, j, k... for space indices.

There is an extra local Lorentz symmetry acting on the flat indices of the vielbein, which leaves the form of the metric (1.5) invariant. We need to introduce a Lorentz connection for parallel transporting those indices, just like the affine connection is used for parallel transporting spacetime indices. The Lorentz connection is called the *spin connection* and denoted by ω^{ab}_μ. It is antisymmetric in the flat indices, since it lives in the adjoint representation of the $so(d-1,1)$ Lie algebra. In three dimensions, the adjoint representation of the Lorentz Lie algebra $so(2,1)$ is equivalent to the fundamental one and one may use the quantity with one flat index for the spin connection defined by

$$\omega^a_\mu = \frac{1}{2}\epsilon^{abc}\omega_{\mu\, bc} \, , \quad \omega_{\mu\, ab} = \epsilon_{abc}\, \omega^c_\mu \, . \tag{1.6}$$

In order to indicate that the formalism is not specific to three dimensions, however, we will use the quantity with two indices. We impose the requirement that the vielbein should be covariantly constant, which can be regarded as the definition of the connection ω:

$$\partial_\mu \, e^a_\nu - \Gamma^\alpha_{\mu\nu} \, e^a_\alpha - \omega^a_{\mu\, b} \, e^b_\nu = 0 \, , \tag{1.7}$$

where $\Gamma^\alpha_{\mu\nu}$ is the affine connection. Eq. (1.7) also implies that the affine connection is compatible with the metric (1.5). Furthermore, the curvature of ω:

$$F^a_{\mu\nu\, b}(\omega) = \partial_\mu \, \omega^a_{\nu\, b} - \partial_\nu \, \omega^a_{\mu\, b} - \omega^a_{\mu\, c} \, \omega^c_{\nu\, b} + \omega^a_{\nu\, c} \, \omega^c_{\mu\, b} \tag{1.8}$$

coincides with the usual definition of the curvature in the following sense:

$$R^\delta_{\alpha\beta\gamma} = F^a_{\gamma\beta\, b} \, e^\delta_a \, e^b_\alpha = \partial_\alpha \, \Gamma^\delta_{\beta\gamma} - \partial_\beta \, \Gamma^\delta_{\alpha\gamma} + \Gamma^\delta_{\alpha\nu} \, \Gamma^\nu_{\beta\gamma} - \Gamma^\delta_{\beta\nu} \, \Gamma^\nu_{\alpha\gamma} \, . \tag{1.9}$$

The Hilbert-Palatini action of the first order formulation for $d = 3$ is

$$S[e, \omega] = \int_M \epsilon_{abc} \, \epsilon^{\mu\nu\rho} \left(e^a_\mu \, F^{bc}_{\nu\rho}(\omega) - \frac{\Lambda}{6} e^a_\mu \, e^b_\nu \, e^c_\rho \right) \, . \tag{1.10}$$

Note that we introduced the units $16\pi G = 1$, which will be used from this point on. The equations of motion read

$$2F^a_{\mu\nu\,b}(\omega) = \frac{\Lambda}{2}(e^a_\mu\, e_{\nu\,b} - e^a_\nu\, e_{\mu\,b}) \,, \tag{1.11}$$

$$D_\mu\, e^a_\nu - D_\nu\, e^a_\mu = 0 \,, \tag{1.12}$$

where $D_\mu e^a_\nu = \partial_\mu\, e^a_\nu - \omega^a_{\mu\,b}\, e^b_\nu$ is the gauge covariant derivative. The second equation means that the spin connection is torsion free. This torsion given by the lhs. of (1.12) is proportional to the usual definition $\Gamma^\rho_{\mu\nu} - \Gamma^\rho_{\nu\mu}$. Thus, (1.7) and (1.12) assure that the affine connection is the unique Levi-Civita connection associated to the metric (1.5).

The first set of equations of motion (1.11) implies for the Ricci tensor:

$$R_{\mu\nu} = F^a_{\mu\rho\,b}\, e^b_\nu\, e^\rho_a = -\frac{\Lambda}{2} g_{\mu\nu} \,, \tag{1.13}$$

where e^ρ_a is the inverse of the vielbein satisfying $e^\rho_a\, e^a_\sigma = \delta^\rho_\sigma$ and $e^\rho_a\, e^b_\rho = \delta^b_a$. We can now see, that Einstein's equations can only be recovered if the vielbein is non-degenerate.[5]

In three dimensions we can write down the following two sets of symmetries of the action (1.10). One is the local Lorentz transformations

$$\delta e^a_\mu = \epsilon^a_{bc}\, e^b_\mu\, \tau^c \,, \qquad \delta\omega^a_{\mu\,b} = \epsilon^a_{bc}\, \partial_\mu\, \tau^c - \epsilon^a_{bc}\, \omega^c_{\mu\,d}\, \tau^d \,. \tag{1.14}$$

The other is the local translations

$$\delta e^a_\mu = \partial_\mu\, \rho^a - \omega^a_{\mu\,b}\, \rho^b \,, \quad \delta\omega^a_{\mu\,b} = \Lambda(e_{\mu\,b}\, \rho^a - e^a_\mu\, \rho_b) \,, \tag{1.15}$$

where $\tau^a(x)$ and $\rho^a(x)$ are the parameters of the local gauge transformations. The action is a scalar, so it is invariant under diffeomorphisms. The infinitesimal diffeomorphism symmetry generated by the vector field $\xi^\mu(x)$ is contained in the above gauge symmetries. If we write $\rho^a = \xi^\mu e^a_\mu$, $\tau^a = \xi^\mu \omega^a_\mu$, the combined transformations (1.14) and (1.15) coincide with the expressions of the Lie derivatives of the fields with respect to the vector field ξ^μ modulo terms proportional to the equations of motion. It is this subtle modification of the symmetry algebra, when carried along off-shell, that may lead to inequivalent quantum models. Note also that the first order

[5] Note that everything written so far goes almost identically in d dimensions. However, because the number of vielbein factors in the curvature term is $d-2$, the variation with respect to the vielbein yields $F \wedge \underbrace{e \wedge \ldots \wedge e}_{d-3}$ for the lhs. of eq. (1.11), which, for $d > 3$, is proportional to the Einstein tensor rather than the curvature tensor.

phase space contains configurations that correspond to degenerate metrics and this even allows for spatial topology change [9, 10, 11].

It has been pointed out in [3] that one can construct a connection $A = A^X_\mu T_X \, dx^\mu$, where T_X are the 6 generators of a Lie algebra and A^X_μ are linear combinations of ω^a_μ and e^a_μ with the following properties. The symmetries given by eq. (1.14) and eq. (1.15) give a genuine gauge transformation for the connection A. Furthermore, the action in terms of A acquires the form of a Chern-Simons gauge theory

$$S[A] = \frac{k}{4\pi} \int_M \mathrm{Tr} \left(A \wedge dA + \frac{2}{3} A \wedge A \wedge A \right) \ . \tag{1.16}$$

The symbol Tr stands for an invariant, non-degenerate, bilinear form of the Lie algebra $iso(2,1)$, $so(2,2)$ and $so(3,1)$ for zero, negative and positive cosmological constant, respectively. Its existence is a nontrivial fact. The coupling constant is given by $k = -1/\sqrt{|\Lambda|}$ for $\Lambda \neq 0$ and by $k = -1$ for $\Lambda = 0$. The equation of motion for this action is

$$F(A) = 0 \ , \tag{1.17}$$

where $F(A)$, as usual, denotes the curvature or field strength of the connection A. This is another manifestation of the absence of local degrees of freedom.

We now restrict ourselves to $\Lambda = 0$ when specifying the details of how e and ω are encoded in A. The formulae for other values of the cosmological constant can be found in [12] or in the original paper [3]. The Poincaré connection for $\Lambda = 0$ reads

$$A = (e^a_\mu \, P_a + \omega^a_\mu \, J_a) dx^\mu \ , \tag{1.18}$$

where P_a and J_a are the translation and Lorentz generators of the Poincaré algebra $iso(2,1)$, respectively. A bilinear form is:

$$\mathrm{Tr}(J^a P^b) = \eta^{ab}, \quad \mathrm{Tr}(J^a J^b) = \mathrm{Tr}(P^a P^b) = 0 \ . \tag{1.19}$$

In other words, the triad and spin connection together constitute a genuine Poincaré gauge field. The action (1.16), defined on closed manifolds, is invariant under infinitesimal gauge transformations (this statement is non-trivial, because of the unusual dependence of the action on the gauge fields). A recipe for its quantization is given in [3], but explicit calculations have been done only for the case that the topology of space is the torus [13, 14].

We finish this section by mentioning the most important results for nonzero cosmological constant. A black hole solution was found in [15] and later it turned out that this three dimensional black hole can be created by point particles [16]. All possible black hole solutions were classified recently [17]. Let us note that a topical paper [18] indicates that the different sectors of 3D gravity corresponding to the sign of the cosmological constant and the signature of the metric (Riemannnian $(+++)$ or Lorentzian $(-++)$) are related. In this thesis however, we restrict ourselves to the flat case, that is, when the cosmological constant is zero.

1.2 Zero cosmological constant

This is the simplest sector of the theory. As will be explained below, the physical phase space has the usual global, linear structure parametrized by positions and momenta with the canonical Poisson bracket. Note that for non-vanishing cosmological constant this is not the case. It is straightforward to quantize such a phase space. In the Schrödinger representation, wave functions are given by square-integrable functions on the configuration space. The operators representing the coordinates act by multiplication and the momenta act as derivatives with respect to the coordinate. However, since the Hamiltonian is identically zero in the reduced phase space, it is difficult to extract dynamical information from the spectrum of the above operators. In the following we will show how the physical phase space can be extracted. Our starting point is the Hamiltonian framework, which we will describe in both the second and the first order formalism. The reduction of the phase space to the physical one will be done in both formalisms in the next section.

The starting point of the canonical formalism is the space-time decomposition of the metric à la Arnowitt Deser and Misner [19]:

$$ds^2 = -N^2 dt^2 + g_{ij}(dx^i + N^i dt)(dx^j + N^j dt) \ . \tag{1.20}$$

$g_{ij}(x^i, t)$ is the Riemannian two-metric on a spacelike hypersurface Σ_t labeled by the coordinates x^i. The functions $N(x,t)$ and $N^i(x,t)$ are called the lapse and the shift, respectively. The two-metric g_{ij} and its inverse g^{ij} are used to lower and raise the spatial indices i, j. Let us define the quantity

$$\pi^{ij} = \sqrt{{}^{(2)}g}\,(K^{ij} - g^{ij}K) \ , \tag{1.21}$$

where $K_{ij} = \frac{1}{N}(\partial_0 g_{ij} - {}^{(2)}\nabla_i N_j - {}^{(2)}\nabla_j N_i)$ is the extrinsic curvature for the ADM metric above and ${}^{(2)}\nabla_i$ denotes the Levi-Civita covariant derivative with respect to the metric g_{ij}. It characterizes the embedding of the

equal time surface Σ in the spacetime manifold.[6] We can now write the decomposed Einstein-Hilbert action (1.2) (with $16\pi G = 1$) in the following compact form

$$I = \int dt \int_\Sigma d^2x \left(\pi^{ij} \partial_0 g_{ij} - N\mathscr{H} - N_i \mathscr{H}^i \right) \tag{1.22}$$

with

$$\mathscr{H}(g_{ij}, \pi^{ij}) = \frac{1}{\sqrt{^{(2)}g}} (\pi^{ij}\pi_{ij} - \pi^2) - \sqrt{^{(2)}g}(\,^{(2)}R - 2\Lambda) \ , \tag{1.23}$$

called the Hamiltonian constraint and

$$\mathscr{H}^i(g_{ij}, \pi^{ij}) = -2\,^{(2)}\nabla_j \pi^{ij} \ , \tag{1.24}$$

called the momentum constraint. The Cauchy data $(g_{ij}(x, t_0), \pi^{ij}(x, t_0))$ at a given parameter t_0 determines the whole spacetime if Σ_{t_0} is a Cauchy surface.[7] The Poisson brackets can be read off from (1.22)

$$\{g_{ij}(x), \pi^{kl}(x')\} = \frac{1}{2}(\delta^k_i \delta^l_j + \delta^l_i \delta^k_j)\delta(x, x') \tag{1.25}$$

The equations of motion coming from the variations with respect to the Lagrange multipliers N and N^i are the constraints $\mathscr{H} = 0$ and $\mathscr{H}^i = 0$, respectively. They do not contain time derivatives of the fields. Furthermore, they are first class, the momentum constraints generate spatial, the Hamiltonian constraint generates timelike diffeomorphisms. To see this, we consider their Poisson brackets with the canonical variables. The rhs. of

$$\left\{ \int d^2x\, \xi_i(x)\, \mathscr{H}^i(x), g_{kl}(x') \right\} = -\left({}^{(2)}\nabla_k \xi_l + {}^{(2)}\nabla_l \xi_k \right)(x') \tag{1.26}$$

is the variation of the two-metric under a diffeomorphism of Σ generated by ξ^i. Turning to the momentum, we find

$$\begin{aligned} &\{ \textstyle\int d^2x\, \xi_i(x)\, \mathscr{H}^i(x), \pi^{kl}(x') \} = \\ &\left(\xi^i \partial_i \pi^{kl} + \pi^{ik} \partial_i \xi^l + \pi^{il} \partial_i \xi^k + \pi^{kl} \partial_i \xi^i \right)(x') \ , \end{aligned} \tag{1.27}$$

[6] The definition for the extrinsic curvature or second fundamental form is $K_{\mu\nu} = -\nabla_\mu n_\nu + n_\mu n^\rho \nabla_\rho n_\nu$, where n^ν is the unit normal to Σ. For the ADM metric given by (1.20) this vector reads $n_\mu = N\delta^0_\mu$.

[7] A *Cauchy surface* is a spacelike hypersurface which is intersected by every inextendible causal curve once and only once. An *inextendible* curve can end only at infinity or at an initial/final singularity.

which is the correct transformation law of the momentum under spatial diffeomorphisms, since π^{ij} is a tensor density: only the quantity $\pi^{ij}/\sqrt{^{(2)}g}$ transforms as a true tensor on Σ.[8] The action of the Hamiltonian constraint is more subtle. It can be shown that it generates diffeomorphisms in the timelike directions, but only on shell. We do not present the derivation of this fact, because we will not continue along these lines, but it can be found e.g. in [12]. Nevertheless, we write down the algebra of constraints. If the generators are written as

$$G[\xi,\xi^i] = \int_\Sigma d^2x\,(\xi\,\mathscr{H} + \xi_i\,\mathscr{H}^i)\ , \tag{1.28}$$

then a lengthy calculation shows [20] that their Poisson bracket can be written as

$$\{G[\xi_1,\xi_1^i], G[\xi_2,\xi_2^i]\} = G[\xi_3,\xi_3^i]\ , \tag{1.29}$$

with

$$\xi_3 = \xi_1^i\,\partial_i\xi_2 - \xi_2^i\,\partial_i\xi_1\ , \tag{1.30}$$

$$\xi_3^k = \xi_1^i\,\partial_i\xi_2^k - \xi_2^i\,\partial_i\xi_1^k + g^{ki}(\xi_1\,\partial_i\xi_2 - \xi_2\,\partial_i\xi_1)\ . \tag{1.31}$$

We see, that the algebra closes, but it is not a genuine Lie algebra: the canonical variable g^{ki} appears in the right-hand side.

The total Hamiltonian of the system is given by

$$H = \int_\Sigma d^2x(N\mathscr{H} + N_i\,\mathscr{H}^i)\ . \tag{1.32}$$

It is a linear combination of the constraints, thus vanishes on shell for closed universes. This fact is a generic feature of theories that are invariant under spacetime diffeomorphisms. If one solves all constraints, there is no dynamics in the reduced space of configurations unless an explicit time slicing is introduced. The reason is that diffeomorphisms in the timelike directions are also symmetries of the theory. For our variables g_{ij} and π^{ij} the Hamiltonian H generates the time evolution in the usual way:

$$\begin{aligned} \dot{g}_{ij} &= \frac{\delta H}{\delta\pi^{ij}} = \{g_{ij},H\}\ , \\ \dot{\pi}^{ij} &= -\frac{\delta H}{\delta g_{ij}} = \{\pi^{ij},H\}\ , \end{aligned} \tag{1.33}$$

[8] One can see this e.g. from the absence of the determinant of the metric in the d^2x integral of formula (1.22).

but the right-hand side depends on the arbitrary lapse and shift functions. They will acquire a definite form when an explicit time slicing is introduced.

Note that the Hamiltonian constraint (1.23) contains the inverse metric and the square root of the determinant of the metric potentially giving rise to problems when one tries to quantize the theory. As we will see below, in three dimensions one can solve the constraints classically and quantize the reduced phase space. However, the problem is more acute in four dimensions, where the phase space reduction technique is unavailable, the constraints have to be promoted to operators, whose joint kernel singles out the subspace of physical states in a larger Hilbert space. For a particular choice of first order variables, the Hamiltonian can be brought into polynomial form [21]. This observation sparked off an extensive program of canonical quantization called loop quantum gravity.

We now turn to canonical three dimensional gravity in the first order formalism. The space-time (ADM) decomposition of the action (1.10) yields the following formula[9]

$$S = 2 \int dt \int_\Sigma d^2x \eta_{ab}\, \epsilon^{0jk}\, (e^a_j\, \partial_0\, \omega^b_k + \omega^a_0\, D_j\, e^b_k + e^a_0\, F^b_{jk}) \ . \tag{1.34}$$

Note that for the sake of simplicity, we used of the spin connection with one index defined by (1.6). The curvature then reads

$$F^a_{\mu\nu} = \partial_\mu\, \omega^a_\nu - \partial_\nu\, \omega^a_\nu + \epsilon^a_{bc}\, \omega^b_\mu\, \omega^c_\nu \ . \tag{1.35}$$

The dynamical fields are the spatial components of the triad e^a_i and the connection ω^b_i. One reads off from (1.34) that they are canonically conjugate,

$$\{e^a_j(x), \omega^b_k(y)\} = \frac{1}{2} \epsilon_{0jk}\, \eta^{ab}\, \delta^{(2)}(x, y) \ , \tag{1.36}$$

where x and y denote coordinates on the surface Σ. The space of classical solutions is spanned by the solutions of the curvature and the Gauss constraint

$$F^a_{12}(\omega) = 0 \ , \quad D_1\, e^a_2 - D_2\, e^a_1 = 0 \ , \tag{1.37}$$

respectively. They are equations of motion coming from the variations with respect to the fields e^a_0 and ω^a_0, respectively. These fields are Lagrange multipliers similarly to the lapse and the shift in the second order formalism. Their time derivatives do not appear in the action. If the triad is invertible, the curvature constraint can be decomposed into a vector and a scalar

[9] A partial integration has been performed, but no boundary term arises, since we assume that Σ is closed.

constraint, implying invariance under space and time diffeomorphisms, respectively, see [22] for details.

Computing the Poisson brackets of the fields with the dynamical variables, one easily finds that the second equation of (1.37) generates the local Lorentz transformations given by (1.14), and the first generates local translations given by (1.15). Hence the constraint algebra is a closed Lie algebra, no further constraints appear. The total Hamiltonian

$$H = \int_\Sigma \eta_{ab}\, \epsilon^{0jk} \left(\omega_0^a\, D_j\, e_k^b + e_0^a\, F_{jk}^b \right) \tag{1.38}$$

is again a linear combination of the constraints.

The above discussion can be repeated in four dimensions. In the second order formalism, we nowhere used the dimensionality of spacetime. Also the space-time decomposition of the first order Hilbert-Palatini action looks similar to its three dimensional counterpart, eq. (1.34): it consists of a kinetic term linear in the vielbein and the time derivative of the spin connection, a Gauss constraint multiplied by the time component of the connection and a curvature constraint multiplied by the time component of the vielbein. However, the most important differences are the following:

(i) the curvature constraint does not imply that ω is a flat connection,

(ii) one finds secondary constraints when determining the Poisson algebra of constraints,

(iii) this algebra does not close, there are second class constraints in the theory.

In four dimensions, the complications with the constraint algebra also lead to problems in the quantum theory, since it is not possible to solve the constraints classically. Remarkable, in three dimensions the constraints can be solved at the classical level: one can determine the reduced phase space of the physical degrees of freedom explicitly. We present two different formulations of this reduction in the next section: one from the second, the other one from the first order formalism.

1.3 Phase space reduction: second order formalism

Let us first return to the canonical metric formulation defined by the action (1.22). Recall that the dynamical variables are the spatial metric g_{ij} and

its conjugate momentum π^{ij}. First, let us see, how the finite number of physical degrees of freedom appear in the space of metrics g_{ij}.

The uniformization theorem [23] asserts that every Riemann surface is conformally equivalent to another so-called *uniformizing* surface with constant curvature. That surface can be

- $\mathbb{C} \cup \{\infty\}$ with its standard metric of constant curvature $R = 1$

$$ds^2 = \frac{4dzd\bar{z}}{(1+|z|^2)^2}\,, \tag{1.39}$$

- the complex plane with the usual flat $R = 0$ metric

$$ds^2 = |dz|^2\,, \tag{1.40}$$

- or the Poincaré disc $D^2 = \{z : |z| < 1\}$ with its metric of curvature $R = -1$

$$ds^2 = \frac{4dzd\bar{z}}{(1-|z|^2)^2}\,, \tag{1.41}$$

modulo a discrete group G of isometries. This group action should behave sufficiently nicely.[10] For example, it cannot have fixed points, which would cause a singularity in the quotient space.

The first case above is the sphere. It is compact and simply connected, so it is topologically distinct from spaces in the second and the third class. We conclude that all metrics with constant positive curvature are diffeomorphic: the reduced configuration space for the spherical topology is zero dimensional (furthermore, there are no more Riemann surfaces in the first class, since every isometry of the sphere has fixed points).

Let us turn to the second class. The fixed point free isometries of (1.40) are the translations $z \to z + c$ with $c \in \mathbb{C}$. If we choose a discrete subgroup G of them with two generators $z \to z + a$ and $z \to z + b$, where $a, b \in \mathbb{C}$, then we find the tori ($a \neq c\,b$, with $c \in \mathbb{R}$). Note that we had to choose two generators in the isometry group for the two generators of the *fundamental group*[11] $\pi_1(\Sigma)$ of the torus. We can choose $a = 1$, $b = \tau$, $\tau \in \mathbb{C}$ without

[10] If the group action G on the manifold M is *properly discontinuous*, then the quotient M/G is a smooth manifold. The latter notion means that each $x \in M$ has a neighbourhood U_x such that $gU_x \cap U_x = \emptyset$ for every non-trivial $g \in G$.

[11] The elements of the fundamental group are equivalence classes (called *homotopy classes*) of oriented closed curves in Σ starting and ending at a common basepoint in Σ. Two curves are equivalent if one can be continuously deformed to the other. The multiplication is the composition of curves: $a \circ b$ is the curve which is obtained by drawing a and then b without raising our pencil from the paper. The unit element is the contractible curve, and the inverse is the same curve class, but with opposite orientation.

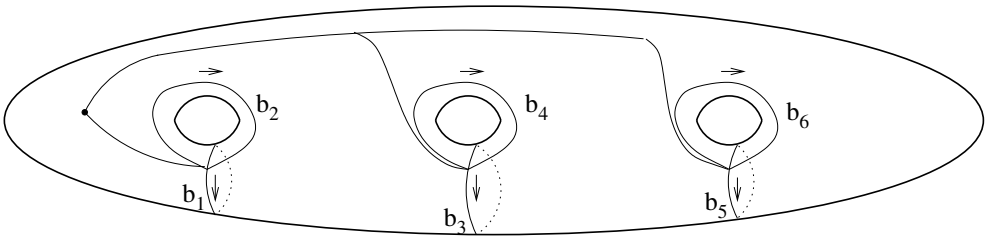


Figure 1.1: A Riemann surface of genus $g = 3$ with the standard generators of its fundamental group. These curves generate any homotopy class of curves by means linking the path of the generators one after the other in appropriate order. They satisfy the relation (1.42).

loss of generality. Hence, the reduced configuration space of metrics of the torus topology is two dimensional, and is parametrized by the complex modulus τ.

The most important for the following chapter is the third class. The isometry group of (1.41) is the three dimensional group $SU(1,1)$, see also appendix A.1. If g_{ij} is the two by two matrix in the defining representation of that group, then its action on D^2 is given by $z \to (g_{11}z+g_{12})/(g_{21}z+g_{22})$. By appropriately choosing the subgroups G of $SU(1,1)$, one can construct all compact Riemann surfaces with *genus* $g > 1$ (the genus is the number of holes or handles of the surface). The fundamental group of a surface with $g > 0$ has $2g$ generators $b_1, b_2, \ldots, b_{2g}$ and one relation:

$$b_1\, b_2\, b_1^{-1}\, b_2^{-1}\, b_3\, b_4\, b_3^{-1}\, b_4^{-1} \ldots b_{2g-1}\, b_{2g}\, b_{2g-1}^{-1}\, b_{2g}^{-1} = e\ , \tag{1.42}$$

see also fig. 1.1 for illustration. We will call them *standard generators* below. We can count the dimension of the reduced configuration space for $g > 1$. Specifying G amounts to assigning elements of $SU(1,1)$ to the generators of $\pi_1(\Sigma)$. The group $SU(1,1)$ is three dimensional, so we have $3\cdot 2g$ parameters. However, due to (1.42), three of them are not independent. Note also that G and gGg^{-1}, with $g \in SU(1,1)$, yields the same surface, which means that another three parameters are redundant. This way we arrive at the number $6g-6$ for the dimension of the reduced configuration space for $g > 1$. It is called Teichmüller space and denoted by $\mathcal{T}_g$. It turns out to be convenient to write the two-metric in the form

$$g_{ij} = e^{2\Phi} f^* \tilde{g}_{ij}\ , \tag{1.43}$$

where $\tilde{g}_{ij} \in \mathcal{T}_g$ and f denotes a diffeomorphism.

Now, let us turn to the momenta. We need the following definition of a differential operator P, which acts on vector fields:

$$(P\,Y)_{ij} = \tilde{\nabla}_i\, Y_j + \tilde{\nabla}_j\, Y_i - \tilde{g}_{ij}\, \tilde{g}^{kl}\, \tilde{\nabla}_k\, Y_l\ . \tag{1.44}$$

We can now write the following decomposition of π^{ij}:

$$\pi^{ij} = e^{-2\Phi}\left(\, p^{ij} + \frac{1}{2}\tilde{g}^{ij}\,\pi + \sqrt{\tilde{g}}\,\tilde{g}^{ik}\,\tilde{g}^{jl}(P\,Y)_{kl}\right)\ , \tag{1.45}$$

where $\pi = g_{ij}\,\pi^{ij}$, $\tilde{\nabla}$ stands for the Levi-Civita covariant derivative with respect to $\tilde{g}_{ij}$, and the first term in (1.45) is transverse and traceless with respect to the same metric. That is,

$$\tilde{\nabla}_i\, p^{ij} = 0\ , \quad \tilde{g}_{ij}\, p^{ij} = 0\ , \tag{1.46}$$

respectively. The field p^{ij} is called quadratic differential or moduli deformation, it parametrizes infinitesimal deformations of conformal classes of metrics. In other words, p^{ij} parametrize the velocities that are conjugate to the positions parametrized by $\tilde{g}_{ij}$.

The decomposition given by (1.45) is unique up to vector fields k_i, which satisfy $(Pk)_{ij} = 0$. These are conformal Killing vectors, they induce only conformal transformations of the metric. The second term of eq. (1.45) is the trace, which can always be eliminated by a conformal transformation (representing one local gauge degree of freedom). The third term is a traceless piece of metric deformation coming from reparametrization (two local gauge degrees of freedom) by the vector field Y^i. So locally, the symmetric p^{ij} containing three degrees of freedom can always be chosen to vanish, but there may be topological obstructions to doing that globally.

Without entering into the details, we mention that a version of the Riemann-Roch theorem [23] states that the dimension of conformal Killing vector fields d_k minus the dimensions d of the space $\{p^{ij}\}$ of transverse, traceless symmetric two-tensors is equal to 3χ, where $\chi = 2-2g$, the Euler characteristic of Σ. In other words

$$d = 6g - 6 + d_k\ . \tag{1.47}$$

For $g = 0$, the conformal Killing transformations are given by $SL(2,\mathbb{C})$. Elements of $SL(2,\mathbb{C})$ act on $\mathbb{C}\cup\{\infty\}$ as was shown for $SU(1,1)$ above, and leave $|dz|^2$ invariant. $SL(2,\mathbb{C})$ is six dimensional, so from the above formula we get $d = 0$ for the sphere. The torus has a two parameter

family of translation isometries, thus $d = 2$. Higher genus surfaces have no conformal Killing vectors, since acting with $\tilde{\nabla}_i$ on $(PY)_{ij} = 0$, one finds

$$\tilde{g}^{ik}\,\tilde{\nabla}_k\tilde{\nabla}_i Y_j = R\,Y_j = -Y_j = 0\,, \tag{1.48}$$

hence $d = 6g - 6$.

Now, the next step is to rewrite the constraints (1.23) and (1.24) using the decomposition (1.43) and (1.45). Then, an explicit time slicing should be introduced. A convenient choice, which is used in [24, 25] is the York time $T = -K_{ij}g^{ij}$. This formula should be interpreted as the definition for the equal time hypersurface Σ_T. For the topology $M = \Sigma \times \mathbb{R}$, such a foliation always exists with the necessary property that the function T is monotone increasing [26]. Note that it was proven in [27] that a three dimensional $\Lambda = 0$ universe, which contains a spacelike Caucy hypersurface, is always of the form $\Sigma \times \mathbb{R}$, see also appendix A.2. We will therefore restrict ourselves to this case throughout this thesis, and use the symbol Σ exclusively for a spacelike hypersurface of the spacetime manifold.

Using the York gauge defined above, it is possible to solve the momentum constraint, which implies that $Y_i = 0$. The Hamiltonian constraint then yields an elliptic differential equation for the conformal factor Φ. We will not need the details of the calculation, they can be found in [24, 25]. The conclusion from this analysis is that, since the constraints fix Y_i and the conformal factor Φ, the reduced phase space is the space $\{\tilde{g}_{ij}(m_\alpha), p^{ij}(p^\beta\} = T_*\mathcal{T}_g$, the (co)tangent space of the Teichmüller space. It is 0, 4 or $12g - 12$ dimensional, for $g = 0, 1$ and > 1, respectively. The action (1.22) after the phase space reduction reads

$$I = \int dT \left\{ p^\alpha \frac{dm_\alpha}{dT} - H_{\text{red}}(m, p, T) \right\}\,, \tag{1.49}$$

where the (time dependent) reduced Hamiltonian

$$H_{\text{red}}(m, p, T) = \int_{\Sigma_T} d^2x \sqrt{\tilde{g}}\, e^{2\Phi(m,p,T)} \tag{1.50}$$

is the area of the equal time surface Σ_T, expressed in terms of the moduli parameters via the differential equation for Φ.

Note that the momentum and Hamiltonian constraints are solved identically and the reduced action (1.49) originates from the kinetic term of (1.22). Note also that the finite dimensional dynamical system can only be made explicit for the torus, where an expression is available for $\tilde{g}_{ij}$ and p^{ij} in terms of the moduli parameters and the differential equation for the

conformal factor can explicitly be solved. For $g > 1$, none of these statements hold: the metric g_{ij} can only be given implicitly by specifying the subgroup of $SU(1,1)$ as explained above or the lengths of certain geodesic curves in the surface. There are existence theorems for the solutions of the differential equation for the conformal factor [24], but no explicit solution is known.

There is one more symmetry we have not factored out. The group of large diffeomorphisms $\mathrm{Diff}(\Sigma)\,/\,\mathrm{Diff}_0(\Sigma)$ is a discrete group with a complicated action on the Teichmüller space. With the exception of the torus even its fundamental domain[12] is unknown.

In practice, the results above are sufficient to quantize the Hamiltonian system coming from the torus topology with four degrees of freedom. It has a large literature, (see [12] and references therein), but it will not be discussed here.

1.4 Phase space reduction: first order formalism

Let us start with the definition of the holonomy, since it will be important for the rest of the discussion and the subsequent chapters. The holonomy or Wilson line of the gauge field along a curve $\alpha : [0,1] \to M$ is the parallel transport operator

$$U_\alpha(A) = \mathcal{P} \exp\left(\int_0^1 ds \frac{d\alpha^\mu(s)}{ds} A_\mu(\alpha(s)) \right) , \tag{1.51}$$

where $\alpha(0) = a$, $\alpha(1) = b$ and $\mathcal{P}$ stands for path ordering. The holonomy obtained for the connection (1.18), with $\Lambda = 0$, is a Poincaré group element, which can be decomposed as $U_\alpha(A) = (U_\alpha(\omega), u_\alpha)$. The first term

$$U_\alpha(\omega) = \mathcal{P} \exp\left(\int_0^1 ds \frac{d\alpha^\mu(s)}{ds} \omega_\mu^a(\alpha(s)) J_a \right) , \tag{1.52}$$

is the Lorentz holonomy, and $u(\alpha)$ is a finite translation. The reason for this fact can be traced back to the structure of the Poincaré group. The multiplication rule in the group has the form of a semi-direct product

$$(\Lambda_1, a_1) \circ (\Lambda_2, a_2) = (\Lambda_1\Lambda_2,\ \Lambda_1 a_2 + a_1) \tag{1.53}$$

[12]The fundamental domain is a set of points in Teichmüller space none of which are connected by the group action and all points outside the set are connected with one inside by the group action.

with $\Lambda_i \in SO(2,1)$ and $a_i \in \mathbb{R}^{2,1}$, which means that the Lorentz part is not affected by the translation part.

The corresponding holonomies transform under general local Poincaré transformations as

$$U_\alpha(A) \to g(a)\, U_\alpha(A)\, g^{-1}(b) \tag{1.54}$$

with $g(x) \in ISO(2,1)$ and

$$U_\alpha(\omega) \to h(a)\, U_\alpha(\omega)\, h^{-1}(b)\ , \tag{1.55}$$

with $h(x) \in SO(2,1)$ being the Lorentz part of the transformation $g(x)$. For closed loops, we will sometimes use the notation $U_\alpha(\omega) \equiv U_\alpha(\omega, x)$, indicating the dependence of the holonomy on the basepoint x of the loop α.

Due to (1.17), the space of physically distinct solutions of the Chern-Simons theory is given by the space of flat Poincaré connections modulo Poincaré gauge transformations. Because of (1.37), we can also say that the reduced configuration space is the space of flat connections ω in Σ modulo Lorentz gauge transformations. A flat connection is locally a pure gauge and completely specified by its holonomies along non-contractible closed curves. We know already that there are $2g-1$ independent closed curves in a surface with genus $g > 1$, and taking into account that the Lorentz group is three dimensional as well as the conjugation symmetry (1.55), we recover the dimension $6g-6$ of the reduced configuration space.

One can also solve the constraints and determine the fully reduced phase space by using phase space functions, which are explicitly gauge invariant. This is the path of the loop representation [28, 29, 30]. In that approach, one set of observables is given by the traces of the Wilson loops (1.52), denoted by $T^0_\alpha(\omega)$. They are gauge invariant as opposed to the untraced holonomies, which have a residual conjugation symmetry as discussed above. An additional set of variables which contain information about the momenta is defined by

$$T^1_\alpha[\omega, e] = \int_\alpha \mathrm{Tr}\left(J_a\, e^a_\mu \frac{d\alpha^\mu(s)}{ds}\, U_\alpha(\omega, x(s)) \right), \tag{1.56}$$

where J_a are the $so(2,1)$ generators. The loop variables T^0_α and T^1_β are not canonically conjugate, but form a closed Poisson algebra, whose right-hand side involve composition of the original loops α and β. The price one has to pay for working directly on the reduced phase space is that the loop variables are overcomplete. The difficulty amounts to finding an independent set of them, which also takes into account the algebraic identities between

traces of the holonomies in the representation that is used. For $g > 1$, for the T^0 variables, this problem was solved in [31]. Quantization in the loop representation amounts to representing the Poisson algebra of a suitable subset of these observables on a Hilbert space. This approach above is an example of the so-called frozen-time formalism. The parameters of the phase space are observables. It implies, due to invariance under timelike diffeomorphisms that they are constants of motion. To recover dynamics from the time independent observables is a difficult task in gravity.

One can follow a different path and choose the untraced $SO(2,1)$ holonomies and the edge vectors of a non-planar polygonal surface as basic variables [32, 33, 34, 35]. These Approaches are all covariant and the Hamiltonian is always a linear combination of the global constraints of the theory. One can fix the Lagrange multipliers by assigning timelike vectors to the corners of the polygonal surface and find the corresponding unique dynamics. The details of these models, in particular their direct relation to the original smooth fields ω and e will be described in chapter 3 as well as the procedure to reduce them to the 't Hooft polygon model. A characteristic of the latter formulation is that it allows for incorporating point particles representing matter in the universe. One of the most attractive features of three dimensional gravity is the fact that one can include point-like particles in it with mass and spin without much difficulty and without losing the simple finite dimensionality of the phase space of the theory.

1.5 Point particles

Despite the simplicity of 2+1 gravity with $\Lambda = 0$, explicit results concerning the quantum theory are only available when the space Σ is the torus. One can however explore a different sector of the theory. Instead of handles, one can include point-like sources. They will follow causal geodesics, while generating conical deficit angles in the spacetime surrounding them. The classification of the arising classical spacetimes is a much more difficult task than in the case of pure gravity without such particles. In the simplest cases of one- or two-particle universes, many results are available in the literature. These systems have been quantized explicitly [36, 37, 1], and scattering amplitudes of particles have also been calculated [38, 39, 40].

The action in the second order variables of a particle reads

$$S_{particle} = \int dt \left(P_i \dot{q}^i + N^i(q)\, P_i - N(q) \sqrt{P_i\, P_j\, g^{ij} + m^2} \right), \qquad (1.57)$$

where q^i is the position of the particle, P_i is its momentum and m is its

mass. This should be added to the gravitational action (1.22) to describe the coupled system. In case of an open universe there are additional boundary terms which also need to be added. If the space Σ is the plane $\mathbb{R}^2$ with punctures at the location of the particles, a convenient gauge is the so-called *instantaneous York gauge*. In this gauge Σ is characterized by the vanishing of the extrinsic curvature $K_{ij} \equiv 0$. Then, using conformal coordinates and the following form of the metric

$$ds^2 = -N^2 dt^2 + e^{2\Phi}(dz + N^z dt)(d\bar{z} + N^{\bar{z}} dt) \tag{1.58}$$

the Hamiltonian constraint takes the form of an inhomogeneous Liouville equation for the conformal factor [41]

$$2\Delta\Phi + e^{2\Phi} = -4\pi \sum_{i=1}^{N} (m_i - 1)\delta(z - z_i) - 4\pi \sum_{A=1}^{N-2} \delta(z - z_A) \ , \tag{1.59}$$

where A label the so-called apparent singularities, and i labels the particles. This equation has a long history in mathematics and it is solved up to $N = 2$. For more particles it is more involved and there is no explicit form of the solution. Nevertheless, the proof that the reduced dynamical system is Hamiltonian, has been recently completed [42] (see also [41] for an overview of this approach). This fact resembles the treatment of empty universes in the York gauge, where the dynamics of the reduced system was shown to be governed by the Hamiltonian (1.50).

A similar approach using the same gauge in the first order formalism was followed by Welling [43], who reduced the solution to the problem to that of solving the Riemann-Hilbert problem, which is as follows. Find an $SU(1,1)$ complex spinor $\xi^\alpha(z)$ on the plane with $N+1$ punctures with given monodromy properties around the punctures. He gave the analytic solution for the two-particle problem and reduced the solution of the general case to finding fermion correlation functions in a conformal field theory. The instantaneous York gauge, which he used is suited for open universes or the static problem of the sphere with particles.

Our interest is somewhat complementary to these approaches, which were briefly discussed above. We want to study closed universes with point particles and also possible topological degrees of freedom (higher genus). The reason is that the polygon model can most conveniently be applied to this case. In the framework of this model a complete (although non-analytic) solution was given to the generic classical N-particle dynamics with arbitrary spatial topology [44].

The starting point of the polygon model is the representation of a globally hyperbolic universe by local charts in Minkowski space. The physical

moduli parameters are then encoded in the transition functions between overlapping coordinate charts. This structure on the manifold M is called *geometric* or (X, G) *structure* by mathematicians. X is the modeling space (which is Minkowski space in our case) and G is a group to which the transition functions belong (the Poincaré group in our case). The equal time slices are characterized by Euclidean polygons, which are glued together along their edges yielding a non-planar surface with conical singularities, and the moduli parameters are encoded in the geometry of these surfaces. The phase space is a reduced phase space of finite dimension. However, it is not completely reduced, the variables have to satisfy constraints, which generate some remnant symmetries among them. The resulting dynamical system has unusual features: chaotic behaviour near a final singularity, and the fluctuation of the number of (the non-physical) variables used in the model. Even though the time evolution in terms of the phase space variables is simple, finding a set which solves the constraints is difficult,[13] especially when the genus of space is $g > 1$.

In this thesis we will solve and explain some of these questions for the case of closed universes. In particular, we will give an algorithm for constructing a $12g-12$ parameter family of matter free universes with Σ an arbitrary $g > 1$ Riemann surface (the torus universes have been constructed in [46]). This is achieved with the help of the following discovery. The position variables, appearing as boost parameters in a set of holonomy matrices, determine a point in the reduced configuration space. This point is represented by a hyperbolic surface with curvature $R = -1$. This surface has a basepoint and the boost parameters are lengths of geodesic loops in certain homotopy classes based at that basepoint. The basepoint corresponds to a chosen Lorentz frame in the physics problem. The algorithm involves a recipe to carefully choose the Lorentz frame such that the constraints between the momenta can be solved. This is the topic of the next chapter and of [45]. We also try to generalize the construction for particles and show what can go wrong when trying to follow the same construction as in the matter free case.

The variables of the dynamical system are divided into positions and momenta. A candidate for the Hamiltonian is the total curvature of the polygonal equal time surface. The observation that this Hamiltonian generates the correct time evolution if the Poisson brackets among the variables is $\{q_i, p_j\} = \delta_{ij}$ was crucial for a qualitative analysis of the quantum the-

[13] The author wrote a computer code to simulate the time evolution of the system, but, apart from a trivial family corresponding to a totally symmetric situation, could not find a set of parameters which solve the constraints for $g > 1$, before the work [45] was completed.

ory [47]. It turned out that the model can be derived from the first order formalism [48], and the Poisson bracket is induced from the canonical one given by (1.36). This is achieved in two steps. The first step is to solve the constraints among the dynamical fields of the first order formalism almost everywhere by choosing the spin connection to be pure gauge. This reduces the number of degrees of freedom from infinity to a finite number, while the model remains covariant. The new phase space variables are the Lorentz holonomies along a set of non-contractible cycles and certain three-vectors conjugate to them. Variants of this reduction have been worked out in the literature as well. The second step involves a gauge fixing and a further reduction to the scalar variables of the polygon model. Then, an alternative derivation is given for the Poisson bracket of the polygon variables directly from (1.36) of the triad and the spin connection. This is the topic of chapter 3 and of [48]. In the appendices we have collected some relevant results from three dimensional geometry, two dimensional hyperbolic geometry and technical proofs omitted in the main text.

Chapter 2

Polygon model

The main part of this chapter contains the construction of any[1] three dimensional Lorentzian manifold M, which is

(i) everywhere locally flat,

(ii) globally hyperbolic,

(iii) maximal, and

(iv) spatially closed.

In other words we will construct the solutions to the Einstein equations in three dimensions, which

(i) correspond to pure gravity with zero cosmological constant and without matter degrees of freedom,

(ii) contain a spacelike Cauchy surface,

(iii) have the property, that geodesics only end at infinity or at the initial/final singularity.

A consequence of criterion (ii) is that the topology of spacetime is necessarily $M = \Sigma \times \mathbb{R}$ [27]. Therefore, criterion (iv) means that Σ is a closed surface. We saw in the previous chapter that there are no moduli for the spherical topology. The results for the torus has been obtained in [46]. It will be summarized briefly at the end of section 2.3. The main part of the chapter, however, deals with the case when $g > 1$.

[1] We have no rigorous proof that the family of spacetimes we consider contain all spacetimes with the specified properties. We will argue that the statement is very likely to be true.

The last part of the chapter contains an attempt to generalize the classification program for the physically more interesting case, when there are also point particles present. We will expose the difficulties when trying to apply here the methods which worked well for the matter-free case. The framework used to complete these tasks is the gauge fixed polygon model of 't Hooft. There is an extensive introduction in the original papers [49, 44, 50] and in the thesis of Max Welling [51]. However, a self-contained introduction will be presented here from a different angle, which is suited for the classification and the comparison to first order gravity, the subject of the next chapter.

The general solution of the three dimensional Einstein equations for a spinning massive particle has been derived in [5] and was treated in the introduction in detail. In this chapter we shall use units $4G = 1$ which means that the mass creates a cone with deficit angle $2\pi m$, so the mass parameter is restricted to be $m < 1$. Non-vanishing spin causes a timelike shift in the identifications of the multi-valued Cartesian coordinates after a full rotation around the puncture, and this implies the existence of closed timelike curves (CTC) in the vicinity of the particle. These cannot be avoided. However, for the case of an open universe with timelike center of mass momentum, there are no CTC's, if the sources have no angular momentum individually [52]. CTC's are also absent when the space is closed, as opposed to the claim in [53]. This result is a direct consequence of the construction [49]: in the polygon model spacetime arises as a foliation in terms of Cauchy surfaces, thus, CTC's are excluded by construction. Moreover, it is explicitly shown in [49], that the moment when the CTC appears in the solution of [53] necessarily happens "after" the final singularity, hence it is an unphysical analytic continuation of the solution.

In the next section, the starting points of the polygon model are presented in full generality, with the inclusion of N point particles (punctures). In section 2.2, we discuss the properties of the Poincaré group elements appearing as matching conditions across the punctures. In section 2.3, the time slicing of the polygon model will be introduced. Degrees of freedom will be counted next, then the time evolution and the Hamiltonian of the system are specified. Section 2.7 describes the concept of a one-polygon tessellation (OPT). It means that the equal time Cauchy surface is one Euclidean polygon. The Poincaré group elements appearing in the matching conditions at the edges are associated to closed paths in this case. Considering the case $N = 0$ first, section 2.8 contains the construction of a hyperbolic smooth surface S with curvature $R = -1$ parametrized by the

momenta.[2] This is the surface mentioned at the end of the previous chapter, and it is the central notion of section 2.8. We will explain in detail the relation of S to an OPT in subsection 2.8.1. S is triangulated by closed geodesic loops with one common basepoint as is shown in subsection 2.8.2. The lengths of these loops are the absolute values of the boost parameters. They are essentially the Zieschang-Vogt-Coldewey coordinates of the Teichmüller space $\mathcal{T}_g$ as demonstrated in subsection 2.8.3. In subsections 2.8.4 and 2.8.5 we explain that S remains invariant under time evolution and the remaining two parameter family of Lorentz transformations of the physical Cauchy surface Σ. In the following section of the chapter, we will try to generalize the whole construction for $N > 0$. Finally, sections 2.10 and 2.11 are devoted to a discussion of the results.

2.1 Geometric structures

In this chapter we are going to describe locally Lorentzian spacetimes of the form $M = \Sigma \times \mathbb{R}$, where Σ is a compact Riemann surface of genus g with finite number of punctures on it. The discussion is based on the references [45, 48].

Suppose, that the metric is Minkowski everywhere and at neighborhoods of the punctures $U \times \mathbb{R}$ it can be written as

$$ds^2 = -dt^2 + dl^2 \,, \quad dl^2 = r^{-2m}(dr^2 + r^2 d\theta^2) \,. \tag{2.1}$$

where m is a mass parameter. Then, (M, ds^2) is a solution of the three dimensional Einstein equations with spinless point-like sources with mass m. The two-metric dl^2 describes the conical singularity mentioned in the introduction. One can easily understand this by inspecting the transformation to the flat element [5]

$$\rho = \frac{r^{1-m}}{1-m} \,, \quad \theta' = (1-m)\,\theta \,, \tag{2.2}$$

which gives $dl^2 = d\rho^2 + \rho^2\, d\theta'^2$, but the range of the angle parameter is $\theta' \in (0, 2\pi\,(1-m)$. We see that there is an upper bound $m < 1$, it corresponds to the Euclidean cone with maximal deficit angle 2π. We are interested

[2]Note that exchanging the momenta and the positions can be achieved by a canonical transformation, so it is a matter of convenience, which terminology is used. We will see that within the polygon model the boost parameters have a natural momentum interpretation, whereas they parametrize the points of Teichmüller space, so they can also be called positions.

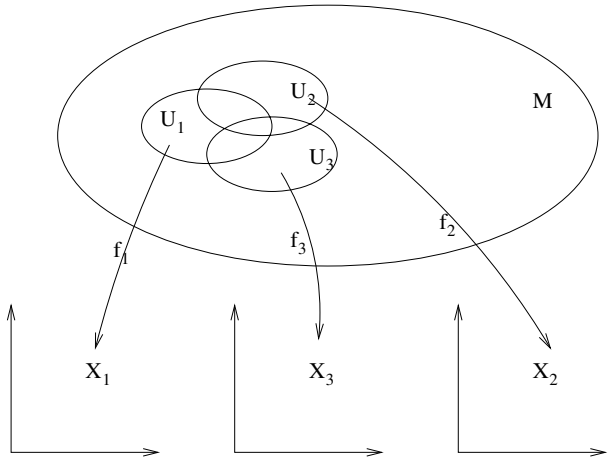


Figure 2.1: Three overlapping local neighborhoods U_i and their associated charts on M (diffeomorphisms $f_i : M \to \mathbb{R}^3$); $f_i \circ f_j^{-1}$ are the Poincaré transformations between neighboring charts.

in *globally hyperbolic* spacetimes: they contain a spacelike surface called *Cauchy surface*, which "determines" the whole spacetime[3]. We can describe a locally flat, punctured Lorentzian manifold in the language of geometric structures. M is covered with Minkowski charts and the matching condition between two neighboring charts $X = (t, x, y)$ and $X' = (t', x', y')$ is given by an element of the Poincaré group $ISO(2,1)$ in three dimensions:

$$X' = PX. \tag{2.3}$$

Consider now three adjacent patches $U_{1,2,3}$ covering matter free regions of M^3 with coordinate frames $X_{1,2,3}$ (fig. 2.1). If the matching conditions are

$$X_2 = P_1 X_1, \quad X_3 = P_2 X_2, \quad X_1 = P_3 X_3, \tag{2.4}$$

in the nonempty intersection of $U_1 \cap U_2 \cap U_3 \subset M$, it follows that

$$P_3 P_2 P_1 X_1 = X_1. \tag{2.5}$$

[3] The statement can be understood from the following equivalent definition of a Cauchy surface, alternative to the one given in the previous chapter. A Cauchy surface is a spacelike hypersurface, whose *domain of dependence* is the whole spacetime $D(S) = M$. The domain of dependence is defined as $D(S) = D^+(S) \cup D^-(S)$, where $D^+(S)$ is the set of points in M, through which all past directed causal curves intersect S, and $D^-(S)$ is defined analogously.

Writing $P_i X = \Lambda_i X + a_i$ with $\Lambda_i \in O(2,1)$ and a_i a Lorentz vector, we obtain

$$\begin{aligned} \Lambda_3 \Lambda_2 \Lambda_1 &= \mathbf{1}, & (2.6)\\ \Lambda_3 \Lambda_2 a_1 + \Lambda_3 a_2 + a_3 &= 0. & (2.7) \end{aligned}$$

We restrict ourselves to the proper ortochronous Lorentz group $SO_+(2,1)$, which is the identity component of $O(2,1)$, hence the manifold will be orientable and time orientable[4]. Every element in $SO_+(2,1)$ can be written as the product of two rotations and a boost,

$$\Lambda_i = R(\phi_i)\, B(\xi_i)\, R(\phi_i'), \tag{2.8}$$

where

$$B(\xi) = \begin{pmatrix} \cosh\xi & 0 & \sinh\xi \\ 0 & 1 & 0 \\ \sinh\xi & 0 & \cosh\xi \end{pmatrix}, \quad R(\phi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{pmatrix}. \tag{2.9}$$

Substituting these expressions into (2.6), one arrives at the vertex condition

$$B(2\eta_3)\, R(\alpha_1)\, B(2\eta_2)\, R(\alpha_3)\, B(2\eta_1)\, R(\alpha_2) = 1, \tag{2.10}$$

with the identifications

$$\begin{aligned} \xi_i &= 2\eta_i \\ \alpha_1 &= \phi_3' + \phi_2 \\ \alpha_2 &= \phi_1' + \phi_3 \\ \alpha_3 &= \phi_2' + \phi_1. \end{aligned}$$

We take the range of the angles to be $\alpha_i \in [0, 2\pi]$, and the factor of two in front of the η's is a convention which will turn out to be useful later. The above matrix equation has the following independent components:

$$\begin{aligned} \cosh(2\eta_k) &= \cosh(2\eta_i)\cosh(2\eta_j) + \sinh(2\eta_i)\sinh(2\eta_j)\cos\alpha_k, \\ \sinh(2\eta_i) &: \sinh(2\eta_j) : \sinh(2\eta_k) = \sin\alpha_i : \sin\alpha_j : \sin\alpha_k, \end{aligned} \tag{2.11}$$

and all permutations of the indices i, j, k.

[4] M is time orientable if there is a non-vanishing timelike vector field on it indicating the direction of time.

2.2 Particles

Now, we want to characterize a particle. If $m \in [0, 1)$, then it is sufficient to determine the Poincaré holonomy of it, that is the Poincaré transformation corresponding to going around the particle[5]. For this purpose the *axis* of a Lorentz transformation needs to be defined. A Lorentz group element is called *elliptic, (parabolic, hyperbolic)* if it has has a timelike (lightlike, spacelike, respectively) eigenvector corresponding to the eigenvalue 1. It always belongs to one of these categories and the corresponding eigenvector is called the *axis*. Note also, elliptic elements are conjugate to pure rotations R, hyperbolic elements are conjugate to pure boosts B, say, in the y direction. Using the decomposition given by equation (2.8) the axis is given by

$$p = c\,(\cosh\eta\,\sin\frac{\phi+\phi'}{2},\ -\sinh\eta\,\cos\frac{\phi-\phi'}{2},\ -\sinh\eta\,\sin\frac{\phi-\phi'}{2}). \tag{2.12}$$

In general we can rewrite the holonomy as

$$X' = P\,X = \Lambda\,X + a = \Lambda\,(X - b) + b + b' \tag{2.13}$$

We require, without loss of generality, that b is perpendicular to the axis while b' parallel to it.[6] If P is to describe the holonomy of a particle, then the world line of the particle should be described by $X' = X$, as we will see. If b' is non-vanishing, then the equation $X' = X$ has no solution, since we can write

$$(\Lambda - \mathbf{1})\,X + b' = (\Lambda - \mathbf{1})\,b \tag{2.14}$$

and the scalar product with p of both sides yields $b'p = 0$. If Λ is elliptic, then b' is a dilatation in a timelike direction. In case Λ is a pure rotation b' is proportional to the angular momentum leading to CTC's [5]. In a globally hyperbolic universe it cannot appear. The solution of equation (2.14) with $b' = 0$ can be written as

$$X = b + c'p, \quad c' \in \mathbb{R}\,, \tag{2.15}$$

where p is the axis of Λ defined above. It gives the world line of the corresponding particle and we can read off its causal properties. If Λ is hyperbolic then it is a tachyon (moving along the spacelike trajectory given by

[5]The holonomy is a homomorphism from the fundamental group to the Poincaré group, it will be properly defined in chapter 2.7. Note, that it coincides with the parallel transport operator of the Poincaré connection via the equivalence of the first and second order formalism explained in the first chapter.

[6]The translation vector is decomposed as $a = (\mathbf{1} - \Lambda)\,b + b'$. A possible component of b parallel to p is annihilated by $(\mathbf{1} - \Lambda)$ and b can be chosen such that b' is parallel to p.

p), if Λ is parabolic (elliptic) then it is a massless (massive) particle. Let us define the constant in front of the axis in eq. (2.12) as

$$c = -sgn\left(\sin\frac{\phi+\phi'}{2}\right) \tag{2.16}$$

whenever the rhs. is non-vanishing, and rewrite the Lorentz holonomy of a massive particle as $\Lambda_i = gR(2\pi - 2\pi m)g^{-1}$ with $g \in SO_+(2,1)$. Then the equality

$$tr\Lambda = 2\cos(2\pi\, m) + 1 \tag{2.17}$$

turns out to be equivalent to

$$p^2 + \sin^2(m\,\pi) = 0\ , \tag{2.18}$$

if the lhs. of (2.17) is parametrized according to eq. (2.8). After elementary manipulations it yields

$$\cos(m\,\pi) = \pm\cosh\eta\cos\frac{\alpha_{\rm def}}{2} \tag{2.19}$$

with $\alpha_{\rm def} = 2\pi - (\phi+\phi')$ is the angle of the wedge cut out of the frame causing the conical geometry. We can define the mass parameter to obey (2.19) with a plus sign. We conclude that p is the momentum of the particle and $2m\pi = \alpha_{\rm def}$ in its rest frame ($\eta = 0$), where the holonomy is a pure rotation. Note that the mass shell condition contains an unusual sine in front of the mass, which is a matter of convention: the current definition comes from writing $T^{00} = m\,\delta(x - x_{(0)})$ for the stress energy tensor leading to eq. (2.1) [5]. With this definition, the mass of particles at rest is additive. Note also, that the range of the mass parameter is $m \in (-\infty, 0) \cup (0, 1)$.[7] The upper bound corresponds to $G/4$, if we restore the usual units. Finally, we will also need the relation between the deficit angle and the boost parameter for a massless particle. It is given by writing down

$$p^2 = 0 \quad \text{or} \quad tr\Lambda = 3 \tag{2.20}$$

(both equations characterize a parabolic Lorentz transformation) yielding

$$\cosh\eta\cos\frac{\alpha_{\rm def}}{2} = 1\ . \tag{2.21}$$

[7]Note that $m = 0$ corresponding to the trivial holonomy is not a massless particle. However the relation (2.19) for a massless particle can be formally obtained by inserting $m = 0$. Negative deficit angles, still correspond to timelike momenta, thus unlike particles with hyperbolic holonomy, need not to be excluded, but we shall not use them.

2.3 Time slicing, phase space

We will introduce now the time slicing of M leading to a foliation of spacetime with Cauchy surfaces characterized by a fixed global time coordinate t. This leads to the polygon model. Fixing[8]

$$t_1 = t_2 \tag{2.22}$$

for two adjacent charts $U_1 \cap U_2 \neq \emptyset$ defines their common boundary in M. These hyperplanes defined for all intersecting regions define the boundaries of all the simplices in a chart. For a given time $t^{(0)}$, each region is a Euclidean polygon with a timelike normal vector. Setting

$$t_1 = t_2 = t_3 = ... = t^{(0)} \tag{2.23}$$

for all the regions defines a piecewise flat Cauchy surface as illustrated in fig. 2.2. When there are no particles we can embed the geometry in Minkowski space and the $t = t^{(0)}$ Cauchy surface is the non-planar tiling of polygons such that each has a timelike normal vector but these normal vectors are not parallel in general. We will assume that at most three polygons meet at each vertex, which presents no loss of generality since edges can have zero length at given instances of time.

The geometry of the Cauchy surface is completely fixed by the collection of straight edges of the polygons. The skeleton of them is a graph Γ and the geometry is specified by two real parameters associated to each of its edges. One is the boost parameter η of the Lorentz transformation in the matching condition $X' = R(\phi)\, B(2\eta)\, R(\phi')\, X + a$ between the two polygons sharing the edge. The other parameter is the length L of the edge. The angles α_i enclosed by edges incident at a vertex are functions of the boost parameters η_i given by eq. (2.11). The meaning of that relation is that a trivalent vertex of the Cauchy surface (to be called *3-vertex*) does *not* represent a particle because the holonomy around it is trivial. However, as a consequence of our choice of time slicing, the two dimensional geometry of the slice is singular and has a conical singularity at these vertices. We have a 2d geometry with conical singularities, where each singularity contributes with a deficit angle $2\pi - \sum_i \alpha_i$ to the total curvature, just like in 2d Regge calculus. By contrast, if there are no particles present, the three-geometry is *flat* everywhere, including the worldlines of the vertices, and the flatness condition is precisely eq. (2.10). Upon inspecting the independent components of the matrix equation (2.10) given by (2.11), one can notice that

[8] This condition could be relaxed into $t_1 = t_2 \cdot \text{constant}$.

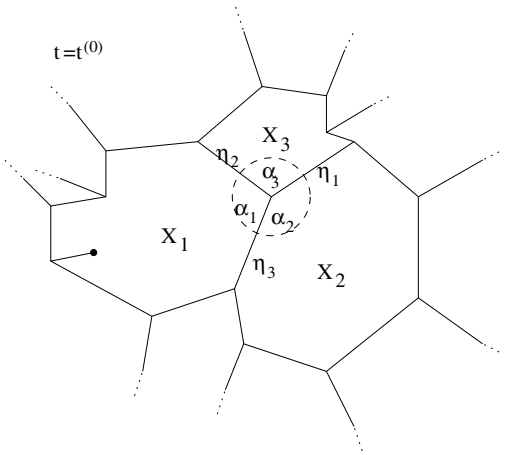


Figure 2.2: The geometry of an equal time slice is characterized by the intersection of the boundaries (2.22) with the lines $t = const$. The vertex in the middle is a trivalent vertex, while the open edge with a dot indicates a particle. The notation is schematic in the sense that both the trivalent and the particle vertex, in general, have conical singularities, which are not visible in this representation.

there is an ambivalence in the equations: if $(\alpha_i, \alpha_j, \alpha_k)$ is a solution, then so is $(2\pi - \alpha_i, 2\pi - \alpha_j, 2\pi - \alpha_k)$. This can be eliminated by the requirement that at most one angle may exceed the value of π at a 3-vertex. The opposite situation is forbidden as explained in the caption of fig. 2.3.

Now, consider a patch which contains a particle. In this case we identify the patch with itself after a rotation given by $\phi + \phi'$ of (2.8). In accordance with the last section, only those holonomies are allowed, where the Lorentz part is elliptic or parabolic, and the translation part can be written as $a = (\mathbf{1} - \Lambda)\, b$. In this case P has a fixed point, and the solution to $X' = X$ gives the world line of the particle. This line coincides with the intersection of the boundaries $t = (P\, X)^0$ and $t = (P^{-1} X)^0$ by construction. Hence at $t = t^{(0)}$ we have an edge identified with the adjacent edge, the angle between them is $\phi + \phi'$ and their common vertex is the location of the particle at the same instant of time. In figure 2.2 it is indicated as an open edge of the graph ending at a *1-vertex*. The figure does not show the deficit angle around the particle and depicts the two adjacent edges as glued together. The relation between the (deficit) angle and the boost parameter for a massive particle is given by (2.19). For a massless particle the corresponding relation is (2.21).

The phase space is characterized by a decorated graph Γ, being the skeleton of the polygons glued together yielding the piecewise flat Cauchy

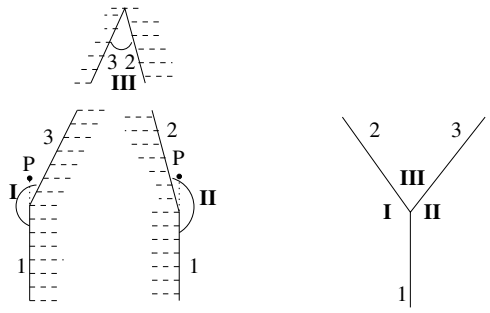


Figure 2.3: If there were two concave angles at a 3-vertex corresponding to region I and II in the figure, then due to the identifications the same spacetime point P would appear twice in the interior of two adjacent polygons. This is forbidden, and this requirement fixes the ambiguity in expressing the angles in terms of the boost parameters at the 3-vertices.

surface. We can also say that Γ is the collection of edges of the simplicial equal time surface. The edges carry two labels: the associated boost parameter η and the length L. This reduced phase space is thus $2E$ dimensional, where E denotes the number of edges of Γ. This is not the fully reduced phase space as the variables are subject to constraints, which are specified as follows. There are three constraints per face (polygon) which arise from the global condition that it must be closed. For a polygon with n sides we have

$$\sum_{i=1}^{n} \alpha_i = (n-2)\,\pi \tag{2.24}$$

and

$$\sum_{i=1}^{n} L_{I(i)} \exp(i\,\theta_i) = 0\ , \tag{2.25}$$

where $\theta_i = \sum_{j=1}^{i}(\pi - \alpha_i)$ and $I(i)$ is the label for the i-th edge starting from a chosen one in counterclockwise direction.

We can now count the independent degrees of freedom. Assume first that $g > 1$ and there are no particles present. Starting from $2E$ variables, there are $3F$ constraints, where F is the number of faces. The number of remaining symmetries is also $3F$, namely, one Lorentz transformation of the coordinate system at each face. However, this action is not free, because conjugating all Lorentz matrices with the same rotation affects neither the boost parameters nor the lengths; it simply amounts to an overall rotation of all coordinate systems. We therefore arrive (with V being the number vertices) at

$$2E - 3F - (3F - 1) = -6\,(F + V - E) + 1 = -6\,\chi + 1 = 12g - 11 \quad (2.26)$$

for the number of independent degrees of freedom. In the equalities above, the formula for the Euler characteristic $\chi = F + V - E = 2 - 2g$ and the trivalence of Γ ($2E = 3V$) have been used. Note that (2.26) is an odd number. Rescaling the lengths simultaneously with a constant corresponds to the reparametrization of time with a constant. Hence, this degree of freedom should also be subtracted and we arrive at the $12g - 12$ independent parameters: the dimension of the reduced phase space of the theory.

In the absence of particles ($N = 0$), there are still V punctures in the spatial slices. However, the correct physical phase space is not the cotangent bundle over the moduli space of Riemann surfaces with V punctures (which has dimension $12g - 12 + 4V$), but the cotangent space over the moduli space $\mathcal{M}_g$ (the space of smooth metrics of constant curvature on a genus-g surface modulo diffeomorphisms[9]), which has dimension $12g-12$. The reason is that we allow only the special class of geometries, where the $2E$ angles of the polygons depend on E parameters (η_i) via the relations (2.11). The conical singularities do not correspond to physical objects in the spacetime, but are merely a consequence of the gauge choice of the global time parameter. In the case $N > 0$, the four degrees of freedom per particle can be easily recovered by the following consideration. To add a 1-vertex to a trivalent graph we have to insert an extra vertex in the "middle" of an edge. This "breaks" that edge, which contains the insertion point, into two pieces (see fig. 2.2). Thus, we added two new edges, each carrying an (L, η) pair. Therefore, the dimension of the reduced phase space is

$$d = 12g - 12 + 4N \quad (2.27)$$

This was the generic situation of either $g > 1$, or $g = 1$ and $N > 1$, or $g = 0$ and $N > 2$. However the empty torus and the sphere with less then three particles needs separate treatment. Let us begin with the empty torus. We can assume that $F = 1$, since increasing the number of faces by one does not change the number of degrees of freedom, as can be seen from eq. (2.26). For $F = 1$ we have $E = 3$ and $V = 2$. As fig. 2.4 shows, the three edges are incident to both vertices and the closure condition of the corresponding hexagon yields $2\sum_i \alpha_i = 4\pi$, that is, no deficit angle for the vertices. However, in this case (2.11) is satisfied only in the special

[9]This moduli space differs from the Teichmüller space $\mathcal{T}_g$ by an additional quotient with respect to the discrete mapping class group action.

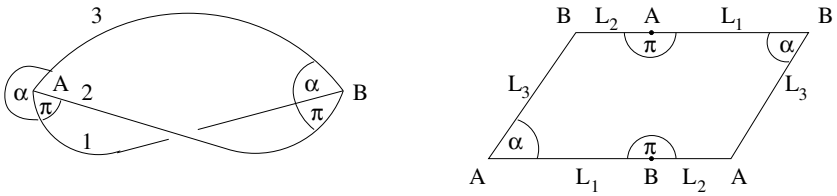


Figure 2.4: The non-static one-polygon torus in the diagrammatic notation and the corresponding hexagon, in which the angles are faithfully represented. The boost parameters are chosen as $\eta_3 = 0$ and $\eta_1 = \eta_2$.

configuration $(\eta, \eta, 0)$ corresponding to $(\alpha, \pi - \alpha, \pi)$ with α freely specifiable and independent of the value of η (hence we can might choose α_A for vertex A and α_B for vertex B). We obtain a Euclidean hexagon if we cut the torus open along the edges of Γ, see fig. 2.4 for illustration. This hexagon is special in the sense that it has three pairs of edges and the members of the pairs are parallel. Due to this, the constraint (2.25) does not restrict any of the lengths, but does require $\alpha_A = \alpha_B = \alpha$. Hence, we have two lengths, (the third one corresponds to time according to the discussion above) one boost parameter and one angle to specify freely[10]. The number of degrees of freedom for the torus is four, as is familiar from many studies in the literature [12]. The reason why the torus violates eq. (2.26) lies in the fact that it has a nontrivial group of isometries. For a detailed treatment of the empty torus in the polygon representation, see [46].

The sphere requires at least three particles, and in the case $N = 3$, it has a trivial phase space. One can understand this by applying the Gauss-Bonnet theorem for the equal time slice, which says, that the total curvature should be equal to 4π (that is 2π times the Euler characteristic of Σ). This cannot be produced by one or two particles, since we saw, that the deficit angle corresponding to a particle is at most 2π. In case $N = 3$, one may argue in several ways to show that there are no moduli. Within the context of the polygon model, the argument goes as follows. Suppose that $F = 1$. Then the polygon is shown in fig. 2.5. If we go to the rest frame of one of the particles, then the boost parameter of the corresponding edge is zero as we will see in the next section. If a boost parameter is zero at a vertex, then the other two must coincide and their value is fixed by eq. (2.24). Two of the lengths are constrained by eq. (2.25) and the last one is identified with time.

[10]There is a static sector, where all boost parameters are zero and two angles can be specified freely

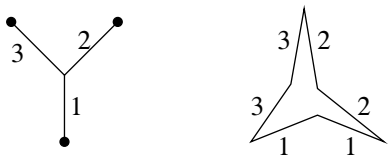


Figure 2.5: The spherical universe with three particles represented by one polygon.

A more general argument is the following. The sphere with three particles is conformal to $\mathbb{C} \cup \infty$ with three punctures and $ds^2 = |dz|^2$. Then, the three punctures can always be transformed to 0,1 and ∞ by an $SL(2,\mathbb{C})$ transformation (the action of such a matrix in the complex plane have been explained in section 1.3). Such a transformation is an isometry of $|dz|^2$, so we conclude that all metrics on the sphere with three punctures belong to the same conformal equivalence class. In other words, the Teichmüller space $\mathcal{T}_{0,3}$, where 0 stands for the genus and 3 is the number of punctures, consists of one point.

2.4 Dynamics

After having specified the time slicing and described the phase space we can study the arising dynamics. First, the boost parameters are constant in time by construction, they are associated to the holonomies independent of the time slicing. The evolution of the edges is fixed by the matching conditions between the neighboring polygons. Assume, for simplicity that the origins of X_1 and X_2 coincide so that the matching condition between them is given by

$$(R(\phi)\, B(2\eta)\, R(\phi')\, X_2 = X_1 \,. \tag{2.28}$$

Its first component is

$$t_2 \cosh 2\eta + (y_2 \cos\phi' + x_2 \sin\phi') \sinh 2\eta = t_1 \,. \tag{2.29}$$

After imposing the gauge condition $t_1 = t_2$ and after elementary manipulations becomes

$$y_2 \cos\phi' + x_2 \sin\phi' = -t_2 \tanh\eta \,, \tag{2.30}$$

or, in the other coordinate system

$$y_1 \cos\phi - x_1 \sin\phi = +t_1 \tanh\eta \,. \tag{2.31}$$

These are the equations of the lines containing the edge in coordinate system X_1 and coordinate system X_2, respectively. We see that the time

evolution of the edges is linear: they move with constant velocity $\tanh\eta$ perpendicular to themselves and either into or away from both of the polygons. In case of a particle, the magnitude of the velocity is a consequence of formulae (2.19) and (2.21). Changing the sign of all boost parameters simultaneously corresponds to time reversal.

The time evolution is well defined (and trivial) for an infinitesimal time interval, but the parametrization can break down whenever an edge shrinks to zero length or a concave angle ($\alpha > \pi$) hits an opposite edge, and part of the variables has to be reshuffled. There are nine types of such transitions, during which edges can disappear and/or be newly created. The figures on the next page show the topological transitions of the graph. That is, parts of Γ are depicted and the result of the possible discrete transitions. Solid lines represent the edges of Γ, while dotted lines are the edges of the *dual graph*. For the time being we can ignore those, they are drawn in the figure for later use. In the first example the arrows indicate the time evolution of the edges. It is clear that the length of edge 5 is decreasing, and the transition happens when it becomes zero. All of the transitions except the first occur only if there are particles or concave angles. The second group preserves, the third decreases, the fourth increases the number of polygons. Notice that the number of polygons, hence the number of variables can fluctuate. The values of the associated new boost parameters are given by the relations (2.11), (2.19), (2.21) and the restriction forbidding more than one angle being concave incident to a vertex.

The total energy is given by the Euler characteristics [5]

$$\frac{1}{4\pi}\int_\Sigma \sqrt{{}^{(2)}g}\,R = \frac{1}{4\pi}\sum_V (2\pi - \sum_{i\,\mathrm{at}\,V}\alpha_i) = 1 - g\ , \tag{2.32}$$

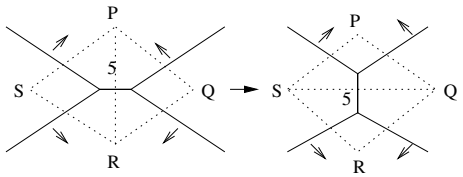
where the last equality is the Gauss-Bonnet theorem. Despite the fact the Hamiltonian is a constant, it generates the time evolution as a function of the boost parameters

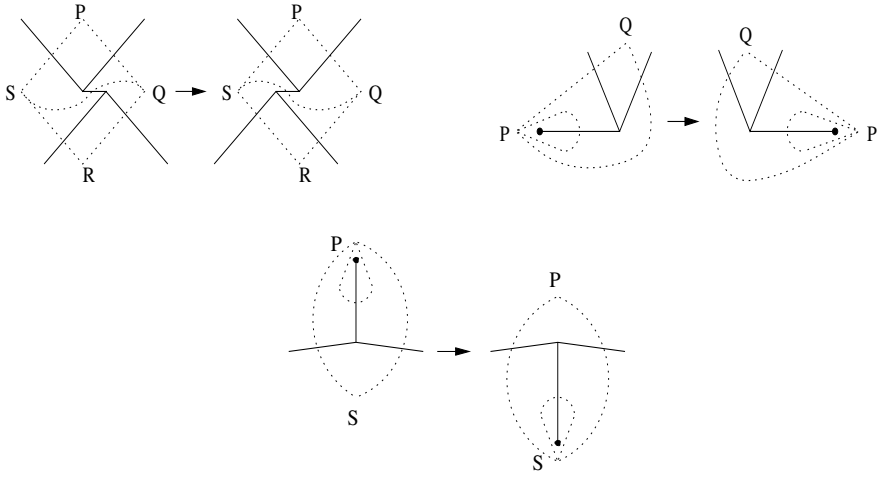
$$H(\eta_i) = \frac{1}{4\pi}\sum_V (2\pi - \sum_{i\,\mathrm{at}\,V}\alpha_i(\eta_j)) \tag{2.33}$$

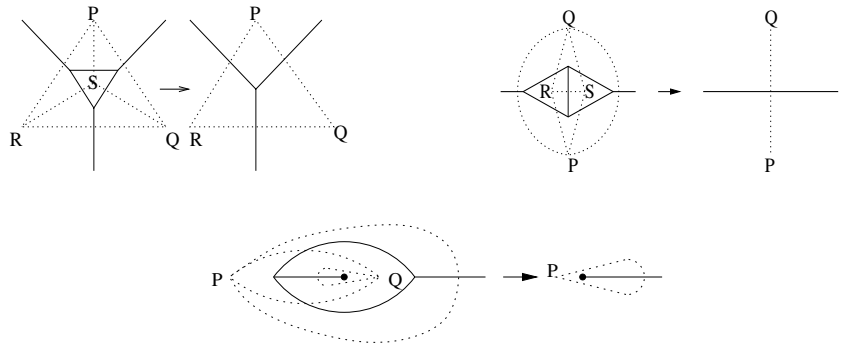
if the symplectic structure is

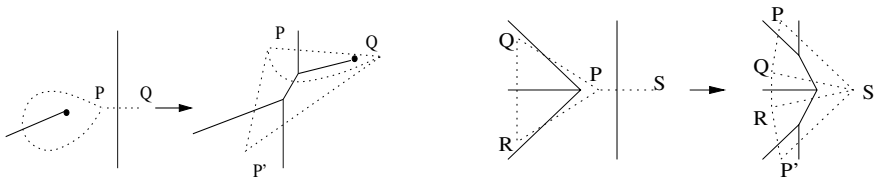
$$\{2\eta_i, L_j\} = 4\pi\,\delta_{ij}\ . \tag{2.34}$$

This was discovered in [47] and it had important consequences for the quantum theory.


P
5
S
Q
R
P
S
5
Q
R


P
S
Q
R
P
S
Q
R
Q
P
Q
P
P
S
P
S


P
S
R
Q
P
R
Q
Q
R
S
P
Q
P
P
Q
P


P
Q
P'
Q
P
S
R
P
Q
R
S
P'

We have introduced the polygon model as 't Hooft constructed it in references [5, 49]. Now we address one of the disadvantages of this otherwise neat way of describing 2+1 dimensional gravity. It is difficult to classify the set of all distinct universes, or to identify the truly reduced phase space. There are many ways to represent the $12g - 12 + 4N$ parameter family of universes in terms of polygons. The first major obstacle is the solution of the constraints (2.24), (2.25). Then, performing a Lorentz transformation inside one polygon causes the geometry of the polygons to change highly non-trivially.

We now address the problem to find solutions to the constraint equations. A particular symmetric solution (with all L's and all η's equal) is known [46]. One can perturb this solution and find a small open neighbourhood in the $12g - 12 + 4N$ dimensional space of independent parameters. Other solutions can be found with intelligent guesses. A straightforward procedure for the case $F = 1$, with given masses for the particles, is to guess all boost parameters except one, such that they satisfy the triangle inequalities $|\eta_i| + |\eta_j| > |\eta_k|$ whenever the edges i, j and k are incident to a common vertex. This is needed for the solution of the relations (2.11) to exist. Then, it should be possible to determine the unspecified boost parameter from the constraint, that the sum of angles of the Euclidean polygon with $2E$ sides is $(2E - 2)\,\pi$. The angles are now fixed. The next step is to guess $2E - 2$ edge lengths and determine the value of the last two from equation (2.25). This way of guessing solutions seems simple, but it does not always work. In the first step, it is difficult to satisfy all triangle inequalities, and eq. (2.25) often yields negative solutions for the undetermined lengths parameters for a fixed set of angles. Note that this problem is not specific to the polygon representation, but is present also in other formulations of 2+1 gravity, for example, in the canonical "frozen-time" loop formulation of 2+1 gravity [29]. The problem of finding an independent set of "loop variables" in this formulation was solved in [31].

We will address this problem in the polygon model and give an algorithmic solution for the constraints when $F = 1$, $g > 1$ and $N = 0$. To complete the task we go through a mathematical exercise: we give a new interpretation to the boost parameters in terms of geodesic lengths in hyperbolic space.

2.5 Vertex conditions

Each polygon carries a timelike unit normal vector, which is a point in hyperbolic space given by the unit hyperboloid $\mathbb{H}^2 = \{(t, x, y) |\ -t^2 + x^2 +$

$y^2 = -1\}$. This space has a canonical metric, which is induced from the Minkowski metric. It has constant curvature $R = -1$. The relation (2.6) says, that if n_1 is the unit normal vector of the polygon of X_1 then $n_2 = \Lambda_1 n_1$, $n_3 = \Lambda_2 n_2$ and $n_1 = \Lambda_3 n_3$. They determine a triangle in hyperbolic space. Indeed, it was pointed out in [54], that eq. (2.10) (or equivalently eq. (2.11)) is the relation between the lengths $|2\eta_i|$, $|2\eta_j|$, $|2\eta_k|$ and the angles $\pi - \alpha_i$, $\pi - \alpha_j$, $\pi - \alpha_k$ of a hyperbolic triangle, if the latter angles are positive. The triangle inequalities,

$$|\eta_i| + |\eta_j| \geq |\eta_k| \tag{2.35}$$

are constraints on the system as mentioned earlier. It should hold for all permutations of edges (i, j, k) at a 3-vertex, otherwise (2.11) has no real solutions for the angles. One has to be careful when interpreting the boost parameters as hyperbolic lengths, since the boost parameters can also be negative and the angles can also be concave. However, all cases can be associated with standard hyperbolic triangles of positive lengths $\tilde{\eta}_i$ and angles $0 < \tilde{\alpha}_i < \pi$. There are three different cases,

- homogeneous vertex: $\mathrm{sgn}(\eta_i) = const$, for which we set

$$\begin{aligned} \tilde{\eta}_i &= |2\eta_i| \,, \\ \tilde{\alpha}_i &= \pi - \alpha_i \,, \end{aligned}$$

  where $i \in \{1, 2, 3\}$ labels the edges (and opposite angles) incident at the given 3-vertex;

- mixed vertex: $\mathrm{sgn}(\eta_3) \neq \mathrm{sgn}(\eta_2) = \mathrm{sgn}(\eta_1)$, where the identification is made according to fig. 2.6, namely,

$$\begin{aligned} \tilde{\eta}_i &= |2\eta_i| \,, \\ \tilde{\alpha}_3 &= \alpha_3 - \pi \\ \tilde{\alpha}_i &= \alpha_i, \quad i = 1, 2 \,; \end{aligned}$$

- the degenerate case: $|\eta_i| + |\eta_j| = |\eta_k|$, characterized by $\tilde{\alpha_1} = \tilde{\alpha_2} = 0$ and $\tilde{\alpha_3} = \pi$.

Note that in the first case $2\pi \leqq \sum_{i=1}^{3} \alpha_i < 3\pi$ and in the second case $\pi < \sum_{i=1}^{3} \alpha_i \leqq 2\pi$. Thus, homogeneous vertices correspond to negative, mixed vertices to positive curvature.

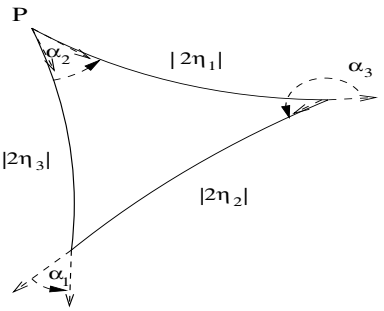


Figure 2.6: Also a "mixed" vertex can be associated with a true hyperbolic triangle with sides $|2\eta_i|$, but the identification of angles is different from the homogeneous case. Assuming that α_3 is concave, and proceeding clockwise from P, one reads off $\tilde{\alpha}_1 = \alpha_1$, $\tilde{\alpha}_2 = \alpha_2$ and $\tilde{\alpha}_3 = \alpha_3 - \pi$.

Let us now take the conical singularity due to a particle, and consider the formulae connecting its boost parameter and its deficit angle $\alpha_{\rm def}$ from the hyperbolic point of view. If we define $\tilde{\alpha}_p = \alpha_{\rm def} - \pi$ then (2.19) reads

$$\cos(1-m)\,\pi = \cosh\eta \sin\frac{\tilde{\alpha}_p}{2}\ . \tag{2.36}$$

The latter can be derived from the generic formula (2.11) for a hyperbolic triangle, with one angle being $\pi/2$. In other words, it is the relation for a hyperbolic triangle with a right angle. The angle $(1-m)\pi$ is opposite to the side with length $|\eta|$ adjacent to the right angle; $\tilde{\alpha}_p/2$ is the other angle. Formally, if $0 \leqq m \leqq 1/2$ then $-\pi \leqq \tilde{\alpha}_p \leqq 0$, and if $1/2 < m < 1$ then $0 < \tilde{\alpha}_p < \pi$. We will see in section 2.9, that the relation (2.36) has a hyperbolic triangle interpretation also for the case when $\tilde{\alpha}_p < 0$. Finally, eq. (2.21) corresponds to a right triangle with two right angles. It has two infinite and parallel sides meeting each other at the conformal boundary of hyperbolic space enclosing zero angle, opposite to the finite side of length $|\eta|$.

2.6 The dual graph

For the next step toward the solution of the polygon closure constraints we need the notion of the *dual graph*. The hyperbolic triangles (encoded by the vertex conditions) are associated to vertices, therefore, they are in one-to-one correspondence with the faces of a graph γ dual to Γ. The dual

of a graph is another graph with the following properties (uncapitalized v, e and f denotes vertices, edges and faces of the dual graph, respectively).

- It has a vertex in the middle of each face of the original graph thus $v = F$.
- Edges connect neighboring vertices. They cross one and only one edge of the graph Γ, so $e = E$.
- Each dual face encloses an original vertex, therefore $f = V$.

Before we turn to a specific example to illustrate the above structure, let us explain what a *tessellation* means. In the general discussion, we had many polygons glued together and we did not need to talk about tessellations. Nevertheless, in practice, one usually represents the Cauchy surface that has the topology of a Riemann surface, with as few polygons as possible. There are different ways to represent a graph Γ. One is to list the polygons and to provide the edges with labels, such that edges with coinciding labels are to be glued together. One can also say, that this picture emerges from cutting the surface Σ open along the edges of Γ. This is what we will call a tessellation.

Fig. 2.7 represents a two polygon universe without particles. Γ satisfies the general requirement that its complement $\Sigma\backslash\Gamma$ is simply connected. This property is a reformulation of the fact that it bounds a number of polygonal faces (that is discs, topologically). The dual graph of the figure can be constructed by placing a point P above edge 1 and another point Q below it. Then, one should connect each point to the middle of each edge of the polygon surrounding it, such that the curve does not intersect any other edge of Γ. At the end of this procedure, there will be three classes of curves: according to their source and target vertex: $Q \leftrightarrow P$, $P \leftrightarrow P$ and $Q \leftrightarrow Q$. The complete curve system forms a triangle graph and it satisfies the following properties:

(i) Each curve intersects one and only one edge of Γ.

(ii) The curves are pairwise non-homotopic with fixed endpoints.

(iii) None of the closed curves are contractible.

In fig. 2.8 both graphs are shown.

Finally, the inclusion of particles modifies the dual graph, such that its faces are not all triangles, but there are also *monogons*: faces entirely bounded by one (closed) curve corresponding to the particles. Considering

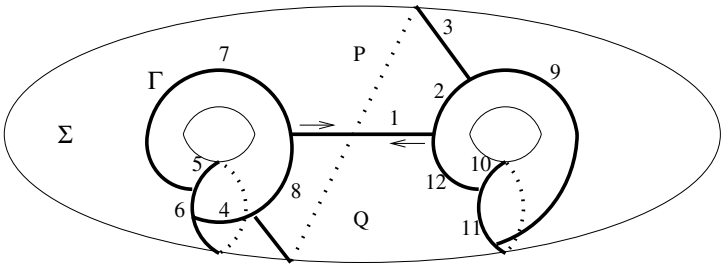


Figure 2.7: The spatial universe Σ is a piecewise flat genus-2 surface (thin lines), whose Cauchy data are associated with the graph Γ (thick lines). The spatial metric on Σ is flat everywhere, but has conical singularities at the vertices of Γ, and the edges of Γ are straight lines. The (thickened-out) graph has two boundary components, $(1, 2, 3, 4, 5, 6, 4, 8, 7, 5, 6, 7)$ and $(1, 8, 3, 9, 10, 11, 9, 2, 12, 10, 11, 12)$, and represents a two-polygon universe. Note also, that the arrows represent a circular ordering of going around the boundary graph Γ thought as a fat graph. Then, the orientation of an edge and its partner is always opposite.

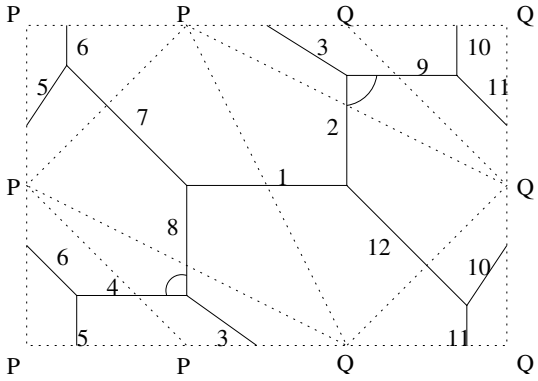


Figure 2.8: The genus two surface with the graph Γ (solid lines) of fig. 2.7, but it is cut open along those dual edges (dotted lines), which are on the boundary of the schematic rectangle. As a consequence we see many copies of the points P and Q. The identification of the (dual) edges at the boundary of this figure (by matching the numbers shown) goes always with opposite orientation.

the particles as punctures of the surface Σ, the dual graph satisfies the properties 1-3 above also in the general case when there are particles present. Fig. 2.9 shows a genus $g = 2$ universe containing two particles and a spherical four-particle universe, their corresponding tessellations and the dual graphs.

2.7 One polygon tessellation (OPT)

From this point on, we suppose that $g > 1$, or $g = 1$ and $N > 0$, or $g = 0$ and $N > 2$. The more polygons we use for describing the initial surface, the more variables we need to work with. We would like to be economic and work with only one polygon. Then, there are only three constraints to solve which will be done below. To justify studying only OPT's it will be necessary to see whether all universes admit such a slicing. We will return to this question after exploiting the advantages of working with OPT's. An example of a graph Γ on Σ and the associated dual curve system is shown in figures 2.10 and 2.11.

The *holonomy map* is naturally associated to a geometric structure. It is a homomorphism from the fundamental group of the manifold to the group, to which the transition functions belong. In our case it is a map from the fundamental group $\pi_1(M)$ to the Poincaré group $ISO(2,1)$. But due to the topology $M = \Sigma \times \mathbb{R}$ of the universes we study, $\pi_1(M) = \pi_1(\Sigma)$.

In case of an OPT, the edges of γ are all closed curves connecting the (only) vertex to itself. They can be identified with elements of the fundamental group. So we can say, that the holonomy map assigns Poincaré group elements to each edge of γ. These are the same group elements as the ones appearing in the matching conditions. Since the number e of the edges of γ is more than the number of independent generators of the fundamental group[11] ($E = 6g-3+2N > 2g+N-1$), the set of corresponding generators is over-complete. We can understand the $4g - 2 + N$ vertex relations of type (2.5) obeyed by them.

It has been proven in [27], that the holonomy of a globally hyperbolic, locally flat universe $\Sigma \times \mathbb{R}$ with $N = 0$ has a special property: its Lorentz part $f : \pi_1(\Sigma) \to SO(2,1)$ gives a discrete embedding, and the quotient $\mathbb{H}^2/G$ (with $G = f(\pi_1(\Sigma))$, a discrete subgroup of the Lorentz group) is a

[11] When there are isolated punctures on a Riemann surface, the fundamental group has N extra generators besides the $2g$ standard generators: those curves, which can be continuously contracted to the individual punctures. The relation (1.42) is modified in such a way, that the lhs. is multiplied from the right with these generators in the appropriate order.

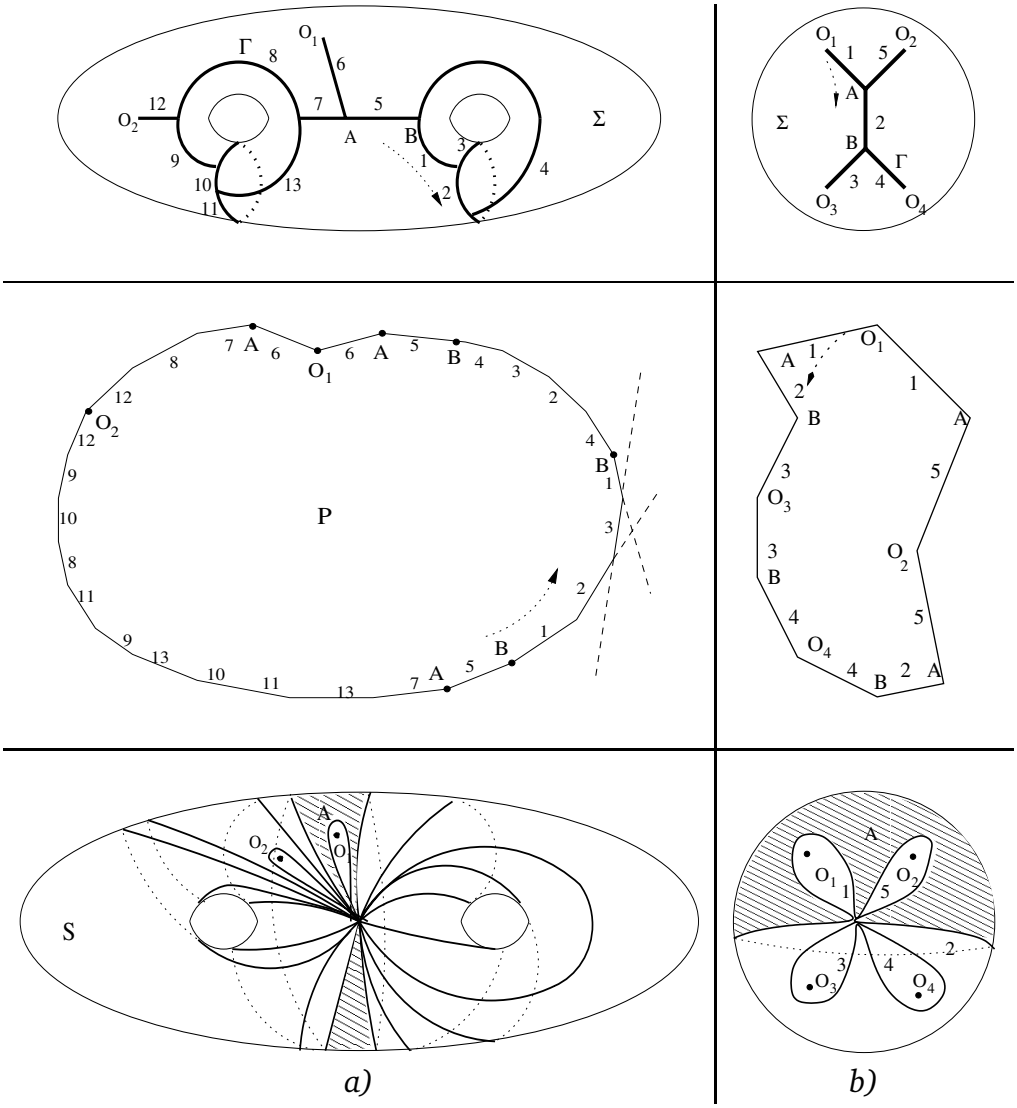


Figure 2.9: *a)*: a genus $g = 2$, $N = 2$ universe, *b)*: a genus $g = 0$, $N = 4$ universe. The graph Γ is shown in the top figures, the corresponding polygons $P = \Sigma \backslash \Gamma$ in the middle, and the dual graphs in the bottom (Note that the dual graph of the genus two surface on the bottom left is reflected with respect to the plane of the paper). The few drawn dashed lines in the middle left picture are the intersections of the hyperplanes given by $t = (P_i X)^0$ with the $t = 0$ spacelike planes. The edges of the polygon P are determined by the segments between the intersections of these lines. Those edges of the polygon bearing the same number have the same length and are to be glued with opposite orientation. The points A and B are examples of 3-vertices, O_i are 1-vertices. At vertices the sum of the angles are different from 2π. The dual graph cuts the surface into triangles and monogons, corresponding to 3-, and 1-vertices, respectively. In the bottom figures one triangle is shaded. It corresponds to those edges of Γ, which are incident to the 3-vertex A.

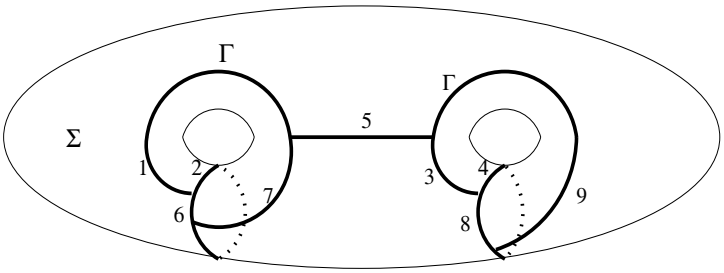


Figure 2.10: Example of a graph Γ corresponding to a OPT of genus 2.

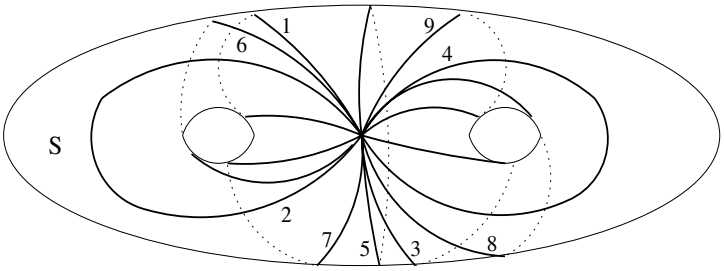


Figure 2.11: The curve system γ corresponding to the OPT in fig. 2.10 has been mapped to a set of homotopic geodesic loops on the smooth constant-curvature surface S of the same genus. All loops are oriented, and the labels $i = 1, \ldots, 9$ indicate the outgoing direction from the base point. Note that the picture is a mirror image (with respect to the paper plane) of the "correct" one.

smooth hyperbolic surface with the same genus as Σ. We will see in the next section that the boost parameters corresponding to an OPT are lengths of geodesic loops based at a common basepoint in this smooth surface; these geodesic loops cut the surface into triangles.

2.8 Uniformizing surface without particles

Let us assume for the this section 2.8, that there are no particles present ($N = 0$). Take the hyperbolic triangles associated to the vertices (with side lengths given by the corresponding boost parameters) and glue them together along the edges according to the structure of the dual graph. We obtain a surface S with the topology of Σ, which has a hyperbolic structure on it. The first statement is clear: the gluing identifications of the edges of γ fix the topology without any metric information. The metric information is encoded by hyperbolic triangles, the sides of which are in one-to-one correspondence with the edges of γ. These were curves connecting vertices of the dual graph in definite homotopy class with fixed endpoints, the corresponding sides of the hyperbolic triangles are geodesic arcs in S. In the following we will use the symbol γ for both the dual graph of Γ and also the corresponding graph on S, whose edges are geodesics.

To verify that S is a genuine hyperbolic surface without singularities, the only thing we have to check is whether the sum of angles is 2π at the vertices. If we have $\tilde{\alpha}_i = \pi - \alpha_i$ for all values i, then the statement is easily seen to be true. The angles $\tilde{\alpha}_i$ are then the outer angles of the Euclidean polygons, thus, they add up to 2π. Recall, that the above condition for a vertex means that all boost parameters at a 3-vertex have the same sign (such a vertex was called homogeneous). Since γ is connected, the above condition is fulfilled if all boost parameters have the same sign.

Homogeneous vertices have only convex angles, as we saw in section 2.5. If there is a concave angle, the identification of the (positive) angles of the corresponding hyperbolic triangle changes. If we want keep the interpretation that $\pi - \alpha$ is the angle appearing at a basepoint in the hyperbolic surface (a vertex of γ), then we should explain what it means, if it is negative. The solution is that triangles corresponding to mixed vertices have to be cut out instead of glued. Suppose that the angle α indicated in fig. 2.12 is concave. Then $\pi - \alpha$ is negative and this means that the angle $\tilde{\alpha}$ of the geodesic triangle PQR should count with a negative sign when summing the angles at basepoint P in S. The corresponding piece of the geodesic triangulation of S is the right-most figure rather than the middle. The angles of the hyperbolic triangle PQR is $\tilde{\alpha} = \alpha - \pi$, $\tilde{\beta} = \beta$ and $\tilde{\gamma} = \gamma$ according

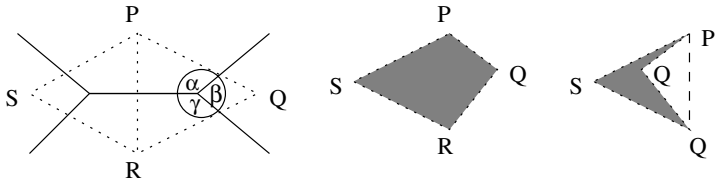


Figure 2.12: The piece of Γ represented by the solid lines prescribes the identification of the dual triangles PSR and PQR along the edge PR. If the triangle PQR corresponds to a homogeneous (mixed) vertex, then it should be glued to (cut from) the triangle PRS as shown in the middle (right) picture. The shaded region indicates the piece of the hyperbolic structure S corresponding to the two triangles.

to the analysis in sec. 2.5. When the triangle PQR is cut out rather than glued, then the angles count as $-\tilde{\alpha}$, $2\pi - \tilde{\beta}$ and $-\tilde{\gamma}$, as can be seen in the right figure. Suppose now, that P,Q and R are identified in S. Then the angles together count as $-\tilde{\alpha} + 2\pi - \tilde{\beta} - \tilde{\gamma} = (\pi - \alpha) + (\pi - \beta) + (\pi - \gamma)$, when summing all angles at the basepoint. Hence, summing the angles at the basepoint still correspond to summing the outer angles of the Euclidean polygon.

For a multi-polygon configuration when Q is not identified with R, the previous interpretation is not possible: $-\tilde{\alpha} = \pi - \alpha$ is alright, but $2\pi - \tilde{\beta} \neq \pi - \beta$ and $-\tilde{\gamma} \neq \pi - \gamma$, only the sum of the latter two relations is an equation. Thus, $\tilde{\beta}$ and $\tilde{\gamma}$ should be incident to the same basepoint for the proposed interpretation to hold. We could not resolve this problem. Note that it is not present in an OPT.

In the remainder of this section, we will prove various properties of OPT's and their associated smooth surfaces by proceeding in a number of steps:

(i) If all η's have the same sign, we can identify their (absolute) values directly with the lengths of the corresponding geodesic loops in the triangulation of S. Special properties of such configurations will be elucidated in sec. 2.8.1.

(ii) The image in S of a triangulation γ coming from an OPT can be obtained directly as follows. After choosing a basepoint (corresponding to the one and only vertex of γ in S)), take the $2g$ standard generators b_j, $j = 1, \ldots, 2g$, of the fundamental group $\pi_1(S)$. They have the property that their complement in S is a geodesic polygon with $4g$ sides in the sequence

$$b_1\, b_2\, b_1^{-1}\, b_2^{-1}\, b_3\, b_4\, b_3^{-1}\, b_4^{-1} \ldots b_{2g-1}\, b_{2g}\, b_{2g-1}^{-1}\, b_{2g}^{-1}\ .$$

The notation b_k^{-1} indicates that the side is to be glued with opposite orientation to its partner b_k (with the same k) to obtain S from the geodesic polygon. (An example is given by the four loops labeled 1 to 4 in fig. 2.11.) One then triangulates the $4g$-sided polygon by drawing geodesic diagonal arcs until the polygon is triangulated. This is the subject of section 2.8.2.

(iii) When two out of the resulting $6g-3$ closed loops in S are taken to be *smooth* geodesics (ie. without any kinks), this uniquely fixes the common basepoint for all loops to be the intersection point of the two smooth loops. In this case only $6g-6$ of the length parameters are independent, and can be identified with the so-called Zieschang-Vogt-Coldewey coordinates of $\mathcal{T}_g$, as we discuss in more detail in section 2.8.3.

(iv) The transitions occurring during the time evolution correspond to changing the triangulation by deleting one arc and drawing another one, but not altering the surface S. This will be explained in section 2.8.4.

(v) The polygon picture exhibits a Lorentz symmetry: one can change the basepoint of the set of loops in S. The length variables η_i will transform in a well-defined way, but the abstract geometry encoded by S does not change. These properties will be explained in section 2.8.5.

(vi) Finally, the usefulness of the previous parts of the construction will become apparent in section 2.8.6 where we also solve the constraints for the length variables L_i, and give an algorithm to construct a large family of universes. The space of these universes is $\mathbb{R}^{12g-12}$ modulo an infinite, discrete group with a complicated action. We conjecture, that the family we have constructed is in fact the complete physical phase space.

2.8.1 Properties of an OPT

If the tessellation consists of a single polygon, all edges of the dual graph γ as well as their images in S are closed loops (geodesics on S) which begin and end at the chosen basepoint. The number of triangles and edges are

$$V = 4g-2\ , \quad E = 6g-3\ , \tag{2.37}$$

as follows immediately from the Euler characteristic and the trivalence of Γ. The closure constraint

$$\sum_{i=1}^{n} \alpha_i = (12g - 8)\,\pi \ , \tag{2.38}$$

which is equivalent to the Gauss-Bonnet theorem (2.32), can be rewritten for the angles of the hyperbolic triangles as

$$\sum_{i=1}^{n} \tilde{\alpha}_i = 2\pi \ , \tag{2.39}$$

as explained in the beginning of this section.

After recalling the main properties of an OPT, let is discuss the special case when all boost parameters have the same sign. In that case, there are only convex angles, so the polygon can never split. Second, there is only one transition which can occur, since the other eight require either the presence of particles and/or more polygons. Third, since the velocity of an edge is given by $\tanh\eta$, the polygon is either expanding or shrinking at all times. In the direction of time corresponding to a shrinking, the universe always runs into a singularity, because of the lower bound for the value of the boost parameters η_i (and therefore, for the velocities $\tanh\eta_i$ of the edges of Γ): the length of the shortest (non-contractible) smooth geodesic loop on S. This result is in accordance with [27], where it has been proven that all globally hyperbolic matter-free universes with $g > 1$ contain an initial or final singularity, while it is geodesically complete in the other direction of time.

2.8.2 Geodesic polygon

Consider a triangulation of S consisting of geodesic loops based at one point. This corresponds to an OPT. The action of deleting a loop from the triangulation merges two triangles into a quadrilateral. We can draw the other diagonal of this quadrilateral, which changes the structure of the triangulation. The combined action of deleting and drawing is called *elementary move*. We can use the algorithm given in [55] to show, that any triangulation of a given genus-g surface S that arises in our construction (called an "ideal triangulation" in [55]), can be reached from any other one by a finite sequence of elementary moves. The numbers of moves required is bounded by a function of g.

For example, given any graph Γ, its dual γ and a set of boost parameters $\{\eta_i\}$ for genus two, we can calculate the values of $\{\eta_i'\}$ corresponding

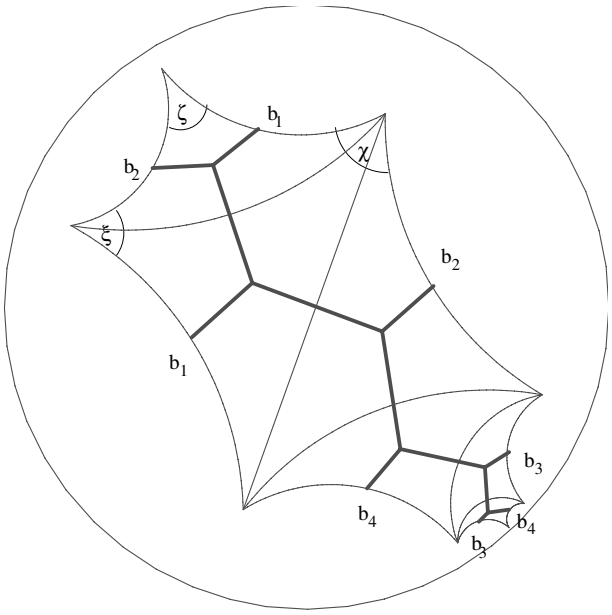


Figure 2.13: A "triangulated" geodesic polygon (thin geodesic arcs) faithfully representing the smooth surface S on the unit disk D^2 (see appendix A.1 for the metric, geodesics and isometry group of this model of hyperbolic space). Edges are to be glued pairwise and with opposite orientation as indicated by the numbers in accordance with relation (1.42) for the generators of the fundamental group. We have also included the original graph Γ (thick lines), the edges of which has no geometric meaning in S.

to the curve system in fig. 2.11. The latter has four geodesic loops labeled 1, 2, 3 and 4, and by cutting the surface along them, one obtains the geodesic polygon with consecutive geodesic arcs $b_1\, b_2\, b_1^{-1}\, b_2^{-1}\, b_3\, b_4\, b_3^{-1}\, b_4^{-1}$, as explained above. Fig. 2.13 shows the same geodesic polygon drawn on the unit disk with the standard hyperbolic metric. A similar construction can be performed for any genus g.

2.8.3 The ZVC coordinates

In general the geodesic polygon is described by $6g$ parameters, to wit, $4g$ angles and $2g$ side lengths. Since all of the angles appear at the same basepoint P, they must sum up to $\sum_i \tilde{\alpha}_i = 2\pi$. Assume now, that P is the intersection point of the smooth geodesics labeled 1 and 2, say, and that these two form part of the triangulation. In terms of the geodesic polygon, which is called *normal canonical* polygon in this case, this means that the two angles at b_1 (χ and ζ in fig. 2.13), as well as the two angles at b_2 (ζ

and ξ) should add up to π. Together with the constraint $\sum_i \tilde{\alpha}_i = 2\pi$ we therefore have three equations for the angles. The closure condition for the polygon makes two more sides and one angle redundant, and we arrive at $6g - 6$ degrees of freedom. These ($4g - 2$ independent sides and $2g - 4$ independent angles) are the Zieschang-Vogt-Coldewey (ZVC) coordinates of the Teichmüller space $\mathcal{T}_g$ [56, 57].

Going to the normal polygon amounts to a gauge fixing of the boost parameters η_i, which are the lengths of the arcs of the geodesic polygon, and can be calculated from the ZVC coordinates by using the triangle relations (2.11). In Appendix A.3 we sketch the algorithm of how to construct a set of boost parameters corresponding to an element of $\mathcal{T}_g$ in practice.

2.8.4 The exchange transition

We have seen above how the boost parameters of an OPT can be used to characterize a surface S uniquely. However, as we have pointed out already, within a finite amount of time Σ may undergo a transition which changes both Γ and its dual γ. In the absence of particles and concave angles, only one transition can take place, the so-called exchange transition. This transition is illustrated in the figure on top of page 37. The edge (5) drawn in the middle shrinks to zero size and is subsequently replaced by another edge (5') as indicated. The graph Γ is represented by thick lines. The associated change in the dual diagram γ (dotted lines) corresponds to the substitution of a diagonal 5 of a quadrilateral 1432 by its other diagonal $5'$, whose length can be obtained by elementary trigonometry. This is an elementary move of the triangulation of the surface S in the terminology of sec. 2.8.2. The surface S itself remains unchanged, and is therefore invariant under the time evolution. (Note that this does not imply the absence of time evolution from the original picture, but only reflects the constancy of the edge momenta or velocities.) The fact that S is also left invariant by the residual Lorentz gauge transformations of the polygon model will be demonstrated in the next section.

2.8.5 Lorentz transformation

An important issue that we have not addressed so far is the role played by the choice of the basepoint in S. As we will show in the following, the action of a Lorentz transformation on the boost parameters (which is a symmetry transformation of the piecewise flat formulation) precisely induces a change in the location of the basepoint.

Recall that the matching condition between neighboring coordinate systems X_1 and X_2, in the polygon picture, is

$$\Lambda_1 X_1 + a_1 = X_2 \ . \tag{2.40}$$

If the edge in question separates two distinct polygons, the corresponding coordinate systems X_1 and X_2 can be Lorentz-transformed with independent group elements. In the special case of an OPT we have only one coordinate system, and X_2 is just a "copy" of X_1. Under the action of a Lorentz transformation Λ we have $X_j \to \tilde{X}_j = \Lambda X_j$, $j = 1, 2$, and the matching condition gets modified to

$$\tilde{\Lambda}_1 \tilde{X}_1 + \tilde{a}_1 = \tilde{X}_2 \tag{2.41}$$

with $\tilde{\Lambda}_1 = \Lambda \Lambda_1 \Lambda^{-1}$ and $\tilde{a}_1 = \Lambda a_1$.

We saw that the smooth surface is presented as $S = \mathbb{H}^2/G$, where G is the Lorentz part of the holonomy group (see the end of sec. (2.7)). It is generated by the Lorentz group elements Λ_i associated to the edges. The space $\mathbb{H}^2$ is the universal cover of S. One of the inverse images (called lifts) of the basepoint (that is a point in $\mathbb{H}^2$, which is mapped to the basepoint on S under the covering) is $n = (1, 0, 0)$, the normal vector of the polygon. The universal cover consists of infinite number of copies of the geodesic polygon, and the complete set of lifts of the basepoint is the set $\{\Lambda n \,|\, \Lambda \in G\}$. If one considers the geodesic arcs emanating from n to $\Lambda_i n$ where i labels the $6g - 3$ edges, then one recovers the angles $\tilde{\alpha}_i$ at n enclosed by these arcs, which are lifts of the corresponding geodesic loops on the surface.

What is the effect of a Lorentz transformation? If F was a fundamental domain of S in $\mathbb{H}^2$, then ΛF is a new one after the Lorentz transformation given by Λ, since $\Lambda G \Lambda^{-1} \Lambda F = \Lambda G F$. But ΛF is isometric to F, since $\Lambda \in SO(2, 1)$ is an isometry of $\mathbb{H}^2$. So the surface does not change. However, the triangulation does, since the lift of the original basepoint $n = (1, 0, 0)$ was also mapped to Λn (note that Λ is not an element of G). In other words, the basepoint on the surface has been moved, the point n has a different image on S after Lorentz transformation. The effect of this in the universal cover is that the geodesic arcs $n \leftrightarrow \Lambda_i n$ have been replaced by $n \leftrightarrow \Lambda \Lambda_i \Lambda^{-1} n$. They are lifts of the unique geodesic loops on the surface in the same homotopy classes as their ancestors before the Lorentz transformation, but based at a different basepoint. Note that the boost parameters and the angles, as defined in equations (2.8) and (2.9), all change after a Lorentz transformation.

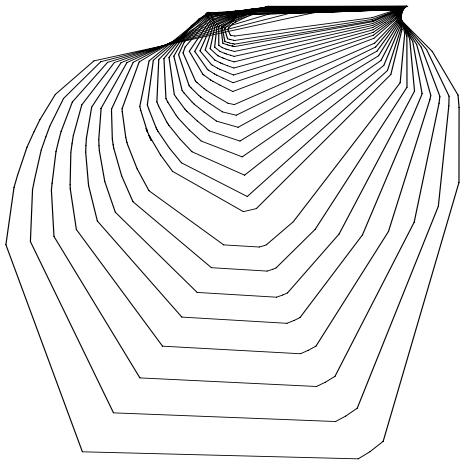


Figure 2.14: An example of how an OPT changes along a one-parameter family of Lorentz transformations. As opposed to time evolution generated by the Hamiltonian (2.33), angles change in a highly nontrivial manner (encoded by the changing of the ideal triangulation on S when moving the basepoint). Within finite value of the parameter transitions appear much like during time evolution.

2.8.6 The complex constraint

While we have found an abstract geometric re-interpretation for the boost parameters of the polygon representation, which has enabled us to identify the independent physical and the redundant gauge degrees of freedom, nothing has been said so far about the other half of the canonical variables, the edge length variables $\{L_i\}$. We will show in the following that the complex constraint (2.25) admits a solution for any triangulation of any surface $S \in \mathcal{T}_g$, provided that the basepoint $P \in S$ is chosen carefully. Furthermore, we will show that from a particular solution (of eq. (2.46) for all relevant triplets z_i) one can construct a $(6g-6)$-parameter family of solutions, thus spanning an entire sector $\mathcal{P}$ of the full phase space where $\mathcal{P} = \mathbb{R}_+^{6g-6} \times \mathcal{T}_g \cong \mathbb{R}^{12g-12}$.

For an OPT, we can rewrite the complex constraint (2.25) as

$$\tilde{C} \equiv \sum_{I=1}^{6g-3} L_I \, z_I = 0 \tag{2.42}$$

with $z_I = \exp(i\,\theta_i) + \exp(i\,\theta_j)$, since each label I will appear exactly twice, namely, at positions i and j when counting the edges of the polygon in counterclockwise direction. The angles θ_i are expressible as sums of angles α_i of the polygon, but there is a more straightforward way of writing them. If the matching condition was $X_2 = \Lambda_1 X_1 + a_1$ with

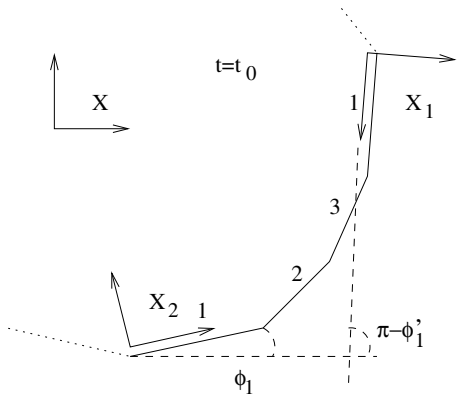


Figure 2.15: Starting from the edge labeled 1 and proceeding in counterclockwise direction one has $I(1) = 1$, $I(2) = 2$, $I(3) = 3$, $I(4) = 1$, ... and therefore $z_1 = \exp(i\theta_1) + \exp(i\theta_4)$. Instead of adding the outer angles in terms of α_i to obtain θ_i, we can determine them directly from the angle parameters appearing in the Lorentz transformation part of the matching conditions corresponding to the edge 1.

$$\Lambda_1 = R(\phi_1)\, B(2\eta_1)\, R(\phi_1') \ , \tag{2.43}$$

one can rewrite it as $R(-\phi_1)\, X_2 = B(2\eta_1)\, R(\phi_1')\, X_1 + a_1'$, and fig. 2.15 then shows that

$$\theta_1 = \phi_1 \ , \quad \theta_4 = \pi - \phi_1' \ . \tag{2.44}$$

In order to derive this relation, one has to take into account that (after an appropriate translation) X_2 has to be rotated by an angle $-\phi_1$ and X_1 by an angle ϕ_1' to align their spatial axes with those of X, and that the two occurrences of an edge always have opposite orientation. After rotating the spatial axes of X_1 and X_2 to those of X, the matching condition between the new coordinate systems is a pure boost $B(2\eta_1)$. The coefficient z_1 of the corresponding edge is therefore given by

$$z_1 = \exp(i\,\phi_1) + \exp(i\,(\pi - \phi_1')) \ . \tag{2.45}$$

The constraint (2.42) is a complex linear equation, and has a solution if and only if the complex coefficients z_i, thought of as vectors based at the origin of the complex plane, are not all contained in a half plane. If they are, there is no non-trivial linear combination with positive coefficients L_i that vanishes. In appendix A.4 we prove the non-trivial fact that after an appropriate conjugation of the generators Λ_i of G, which correspond to the loops of the triangulation, the coefficients z_i will be transformed into a generic position, not contained in any half plane.

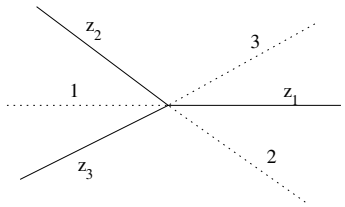


Figure 2.16: For any point z_i ($i > 3$) in one of the convex sections of the complex plane bounded by two half lines $j, k \in \{1, 2, 3\}$, the triangle with corner points (z_i, z_j, z_k) contains the origin. In other words, the corresponding eq. (2.46) has a unique solution with $\lambda_j, \lambda_k, \lambda_j + \lambda_k \in (0, 1)$.

Suppose now that this has been achieved, ie. there are three points z_1, z_2, z_3 in the desired generic position. We can divide the complex plane as depicted in fig. 2.16, and z_i, $i > 3$, is some other point lying in the convex section of the plane bounded by lines 1 and 3, say. Then the equation

$$\lambda_1 z_1 + \lambda_2 z_3 + (1 - \lambda_1 - \lambda_2) z_i = 0 \tag{2.46}$$

admits a unique solution with $\lambda_1, \lambda_2, \lambda_1 + \lambda_2 \in (0, 1)$. A similar statement holds for any point z_i contained in one of the other two convex sections of the plane. There are $6g-5$ independent triangles ($(1, 2, 3)$, and every other index $i \in [4, 6g-3]$ matched with two of $(1, 2, 3)$ according to the location of z_i as explained above). Adding up the resulting $6g - 5$ equations of the form (2.46), each one multiplied by an arbitrary number $\rho_i > 0$, we get a solution to the constraint (2.42). Each index is represented, and each z_i appears with a positive coefficient L_i, namely, a positive linear combination of the ρ_j. All ρ_i's are independent, but we can fix ρ_1 to be 1, which fixes the global time parameter or, equivalently, an overall length scale. Thus we have completed the explicit construction of a $(12g - 12)$-parameter set $\mathcal{P}$ of independent and unconstrained initial conditions for the polygon model, each corresponding to a one-polygon universe. We have obtained this space in the explicit form $\mathcal{P} = \mathbb{R}_+^{6g-6} \times \mathcal{T}_g$. We conjecture that $\mathcal{P}$ is identical with the full phase space of the model, and not just an open subset of it. To prove this conjecture one should prove two statements.

(i) All OPT's (if any) with a concave angle can be Lorentz transformed to one with only convex angles.

(ii) All universes can be represented with an OPT.

We will discuss both statements in the following subsections. Note that

there is ongoing work establishing this result in the mathematics literature [58].

2.8.7 OPT with a concave angle

Let us try to construct an OPT with a concave angle. We start with giving a bound to the maximal number of concave angles. Recall that $\pi < \sum_i \alpha_i \leqq 2\pi$ holds for mixed vertices of Γ, and $2\pi \leqq \sum_i \alpha_i < 3\pi$ for the homogeneous ones. Since the graph Γ has $4g-2$ vertices, in the case of two mixed vertices the sum of all angles in the polygon is given by (cf. (2.24))

$$(12g-8)\pi = \sum_{i=1}^{12g-6} \alpha_i < 4\pi + (4g-4)\,3\pi = (12g-8)\,\pi\,. \qquad (2.47)$$

This is a contradiction, so we can exclude the appearance of more than one mixed vertex. Note that if the graph Γ was one-particle irreducible,[12] it would be impossible to have only one single mixed vertex. This bound is $2F-1$ for an F polygon tessellation without particles. So we can have at most one mixed vertex for $F=1$. The graph can be e.g. that of fig. 2.10, with η_1, η_2, η_5, η_6, $\eta_7 < 0$ and η_3, η_4 , η_8, $\eta_9 > 0$. Then the vertex with edges 359 is mixed, and the left handle is shrinking, whereas the right handle is expanding (with the convention that $\eta > 0$, whenever the corresponding edge is moving out of the polygon).

Suppose that the bound 2π for the sum of angles at the mixed vertex is saturated. Then, an angle at any other vertex should be equal, on average, to $\pi(\,12g-11)\,/\,(12g-9)$. It means that the boost parameters have to be large to give rise to angles close to π or, in other words, hyperbolic triangles with very small angles. Nevertheless, one can find a consistent set of boost parameters, which is consistent with the triangle inequalities and determines a set of angles satisfying the closure constraint of the Euclidean polygon. Equivalently, one can construct a hyperbolic geodesic polygon of sec. 2.8.2, which contains a concave angle. Recall that the vertex, which has a concave angle gives rise to a cut hyperbolic triangle, as illustrated in the right-most picture of fig. 2.12. Hence, there is a concave angle (at point Q in the figure) in the hyperbolic polygon.

The next step is the solution of the complex constraint. It turns out that there are configurations discussed in the previous paragraph, for which we can find solutions to that constraint. They can also be made to avoid self-intersection of the polygon. However, the only solutions we found appear

[12] That is, a graph, which does not fall into disconnected pieces, if we cut any of its edges.

to be stronly Lorentz contracted in one direction, suggesting that another Lorentz frame exists, where also these universes are described by a convex OPT. This however we could not prove.

2.8.8 Multi-polygon tessellation

What we have described up to now is the sector of the theory corresponding to a single polygon. For this case, we have identified a complete set of initial data (the phase space $\not{D}$), and shown that it is mapped into itself under time evolution. However, as we have mentioned in the introduction, a generic universe in the 't Hooft representation is a whole collection of flat polygons glued together at their boundaries. In the following we will analyze multi-polygon configurations, and discuss their relation to the one-polygon sector of the theory. Because of technical problems in connection with the complicated action of the Lorentz gauge transformations on these configurations, we have so far been unable to establish explicit solutions to the analogues of the constraint equation (2.42) and to construct a complete set of initial data. These questions may turn out to be irrelevant due to the closing remark of sec. 2.8.6.

However, in the generic case with particles, the question is not only whether any universe admits an OPT, but also whether one can avoid polygon splitting transitions (bottom pictures on page 37) due to concave angles. A universe consisting of F polygons has a graph Γ with $6g+3(F-2)+2N$ edges. Since every edge comes with a canonical variable pair (η_i, L_i), it is clear that $3F$ of these pairs must correspond to unphysical or redundant information. For $F > 1$, one would expect that at the level of the η's alone, three more boosts can be gauge-fixed for every additional polygon in the tessellation.

We will illustrate now by a specific example how a multi-polygon universe can be effectively reduced to a universe with fewer polygons by a suitable gauge-fixing. Fig. 2.8 represents (a fundamental domain of) the surface S triangulated by means of the geodesic arcs as explained above with $F = 2$, corresponding to the piecewise flat universe of fig. 2.7. Fig. 2.8 is analogous to fig. 2.13 with the difference that the dual graph γ (dotted lines) now has two distinct basepoints P and Q as indicated (note, that fig. 2.8 is only schematic as opposed to fig. 2.13, which renders the angles of the hyperbolic polygon faithfully). The solid lines represent edges of Γ. Any such edge which reaches the boundary of the big quadrilateral must be

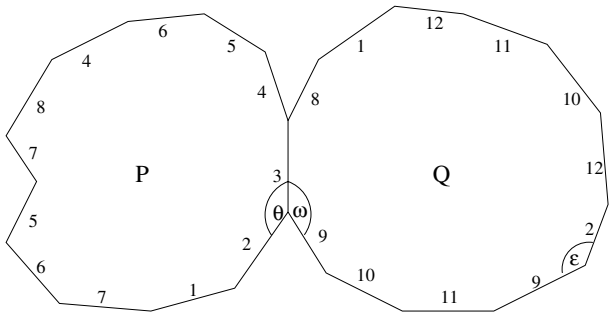


Figure 2.17: The polygons corresponding to the two-polygon universe of figs.2.7 and 2.8. If one Lorentz-transforms the two coordinate frames such that η_3 becomes zero, the angles $\theta + \omega$ and ϵ between edges 2 and 9 become π. Similar statement holds for edges 4 and 8. Switching to the one-polygon representation of the same universe amounts to deleting edge 3 and the vertices on it (which no longer carry any deficit angles) and considering 2 and 9 as well as 4 and 8 as single edges.

glued to the outgoing edge labeled by the same number. The two flat polygons corresponding to this example are shown schematically in fig. 2.17. Closed dual loops based at P (Q) correspond to edges which appear twice in the polygon with center P (Q), and dual curves connecting P to Q correspond to edges appearing once in both polygons. After the gluing the two polygons form a connected piecewise flat surface of genus 2. Consider now the situation when one of the dual edges in S, say, edge 3, has zero length. The basepoints P and Q then fall on top of each other, and the triangles 348 and 392 are degenerate, since the two edges 4 and 8 coincide, as do 2 and 9. The angles enclosed by these pairs of edges are zero, and consequently the angles between the edges of Γ (indicated in fig. 2.8) on the piecewise flat surface are π. Furthermore, since the boost parameter is zero, the matching condition in the most general case is given by

$$X_P = R(\phi)X_Q - a \ . \tag{2.48}$$

This implies that we can redefine X_Q to be exactly X_P without changing the shape of the polygon (since the transformation is a pure rotation). It also means that edge 3 is redundant, and edges 4 and 8 (as well as 2 and 9) can be represented by just single edges. We have therefore re-derived the situation of fig. 2.10.

The conclusion from this analysis is that if the boost parameter of an edge bounding two distinct polygons can be made to vanish by an appropriate gauge transformation, then we have an OPT. The general matching

condition of edge 3 (omitting the translational part) reads

$$X_P = \Lambda\, X_Q \ , \tag{2.49}$$

where X_P and X_Q are now distinct coordinate systems. After performing two independent Lorentz transformations on both frames ($X_P \to \tilde{X}_P = \Lambda_P X_P$ and $X_Q \to \tilde{X}_Q = \Lambda_Q X_Q$) eq. (2.49) becomes

$$X_P = \Lambda_P^{-1} \Lambda\, \Lambda_Q X_Q \ . \tag{2.50}$$

There are now many choices for Λ_P and Λ_Q which reduce the matching condition to $X_P = X_Q$ (for example, $\Lambda_P = \Lambda$ and $\Lambda_Q = I$ will do), which seems to mean that we can effectively get back to a one-polygon tessellation by performing a symmetry transformation. Unfortunately, there is no guarantee that when performing a finite Lorentz transformation (just like upon finite time translations) the gauge does not crash: we do not need to perform a transition. If the transition happens to be a polygon splitting one, then we did not reduce the number of polygons. This prevents us from proving that any multi-polygon universe admits an OPT.

2.9 Uniformizing surface with particles

In sec. 2.5 we have seen that the relation between the deficit angle and the boost parameters also allows for interpretation in terms of a hyperbolic triangle with a right angle. This is a promising sign toward generalizing the constructions of the smooth surface S for $N > 0$. In this case, we expect to have conical singularities also on S (the dimension of the moduli space of these N times punctured surfaces is greater accordingly); by the invariance property we also expect that the corresponding deficit angles are given by the masses.

In the following we will show how this construction works for a large class of universes. It is a two step procedure. The first step is the construction of a bordered hyperbolic surface with N boundary components. This is the result of cutting and pasting the triangles corresponding to the 3-vertices of Γ. Then we glue (or cut) certain isosceles on the boundaries along their base and identify their two equal sides. This way we obtain a hyperbolic cone surface with the above property: the deficit angles are given by 2π times the masses of the particles. This procedure will be explained below in detail.

The invariance property of the surface S under a Lorentz transformation is difficult to establish, since the presentation of S as a quotient is not available any more. The invariance of S under time evolution is also not proven;

the appearance of all transitions makes this task difficult. We believe, that under certain circumstances, and infinitesimally, these properties hold. We will also see, that there are counter examples, that is, universes where the recipe does not work. In such a universe, we cannot construct the hyperbolic cone surface S with N punctures. Unfortunately, we were unable to identify the class of universes, for which the construction does not work.

Let us study the specific example of a genus two $N = 2$, OPT universe of fig. 2.9 a. Consider the hyperbolic structure arising from gluing triangles together according to the dual graph of a generic polygonal Cauchy surface as shown in fig. 2.18 (the solid lines represent the hyperbolic geodesics with lengths $|2\eta_i|$). It has the following properties:

- It has $F = 1$ vertices.
- Those (dual) edges, which form the boundary of a monogon in γ (a 1-vertex in Γ), see the labels 6 and 12 in the figure, are always closed geodesic boundaries (that is, for $F > 1$ as well): their source and target vertices always coincide (the particle's holonomy is associated to the *closed* curve around the particle).

First, assume that the masses of the particles are in the range $m \in (1/2, 1)$. A crucial observation is that the hyperbolic isosceles (triangle with two equal sides) with base length 2η, (the boost parameter corresponding to a particle) and angle $2\pi\,(1-m)$ opposite to the base has angles

$$\tilde{\alpha}_p/2 = (\alpha_{\text{def}} - \pi)/2 \tag{2.51}$$

adjacent to the base. For the chosen range of m the latter is always a positive acute angle. It is a direct consequence of the relation (2.19) and the fact discussed in the end of sec. 2.5, that α_p, $(1-m)\,\pi$ and η_i determine a hyperbolic triangle with a right angle, cf. fig. 2.19. If we glue this isosceles to the boundary geodesic with length 2η on the hyperbolic structure and identify the equal sides of the isosceles, then the boundary disappears and we obtain a conical singularity with deficit angle $2\pi m$.

The constructed hyperbolic surface is thus a candidate for the generalization of S in the matter-free case. The sum of angles at the basepoint indeed add up to 2π, because the hyperbolic angles are always the outer angles of the Euclidean polygons. This statement is true for the particles as well, since we have twice $\tilde{\alpha}_p/2$ at one basepoint and $\tilde{\alpha}_p$ is really the corresponding outer angle of the polygon.

What happens if the particle is light $m \in (0, 1/2)$? We can repeat the previous discussion, but the angles α_p are negative and $(1-m)\pi$ is an

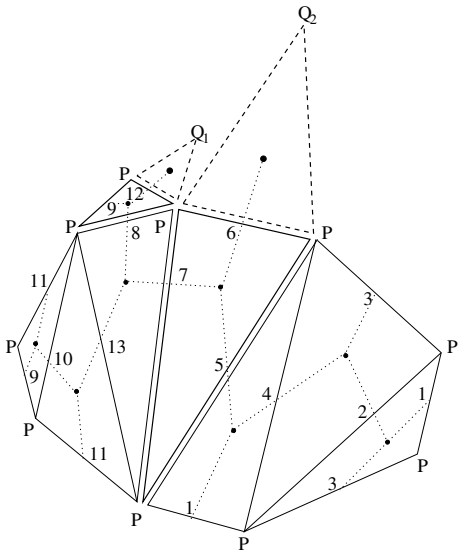


Figure 2.18: An illustration of the dual graph (solid lines) of a genus two OPT with two particles corresponding to the left side of fig. 2.9. Edges labeled by 6 and 12 are boundary geodesics of the bordered surface in the intermediate stage of constructing the hyperbolic surface S with $N = 2$ conical singularities. The two dashed isosceles are to be glued on these edges and their equal sides are to be identified. Then we obtain the hyperbolic surface with conical singularities at Q_1 and Q_2, the corresponding deficit angles are given by the particle masses and the vertex P turns out to be a regular point as explained in the text. The double lines are to be used in figure 2.21.

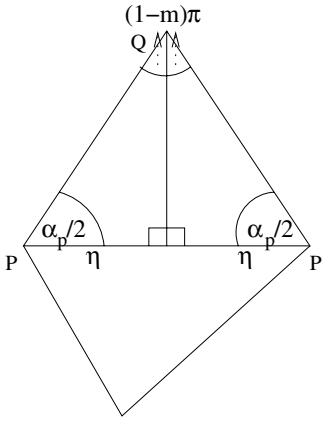


Figure 2.19: Two isometric hyperbolic triangles with a right angle glued together yield the isosceles with the correct base length (to glue to the corresponding boundary geodesic) and angles on the base, the sum of which correspond to the outer angle of the Euclidean polygon at the particle's vertex. The isosceles, when its equal sides are glued together, produces the conical singularity $2\pi m$ at the vertex opposite to the base.

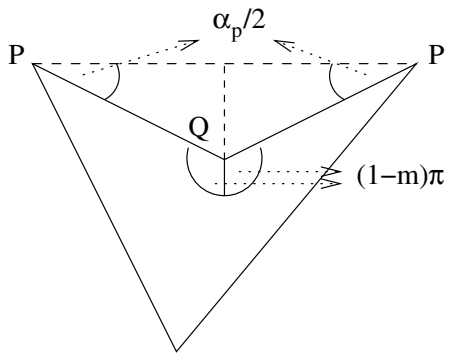


Figure 2.20: In case of light mass $0 < m < 1/2$ the isosceles is to be cut out of the bordered surface, the angles at the base count with a negative sign in the sum of angles at the vertex P, in accordance with the fact that the corresponding outer angle of the Euclidean polygon is negative.

obtuse angle. It means that the triangle is to be cut rather than glued to the corresponding boundary geodesic as shown in fig. 2.20. Hence, the conical singularity at the vertex opposite to the base is again $2\pi m$, the angle $\tilde{\alpha}_p$ is the outer angle (with sign) of the Euclidean polygon. Consult figures 2.21 and 2.22 to gain insight to these procedures on the surface itself.

2.10 Discussion

We have completed the construction on the formal level of the invariant hyperbolic surface for the general case $N > 0$. However the result is far from complete. Let us enumerate the as yet unsolved problems.

(i) Range of the mass
Let us start with discussing the limit cases. The case $m = 1$ corresponds to the maximal deficit angle, it describes a lightlike particle (so m no longer comes from the angle of the unique pure rotation in the conjugacy class of an elliptic holonomy. The equal sides of isosceles becomes infinite, its corner opposite to the base is now the boundary of hyperbolic space and the angle at that corner is zero. The case $m = 0$ is similar to $m = 1$. However, in this case the isosceles should be cut out instead of glued, which is difficult to understand, since the infinite triangle does not fit in a bounded geodesic polygon. If $m = 1/2$, then the corresponding boundary geodesic is smooth (the angle at its vertex is π. The identification amounts to folding the geodesic to half; in the middle point there will be a conical singularity

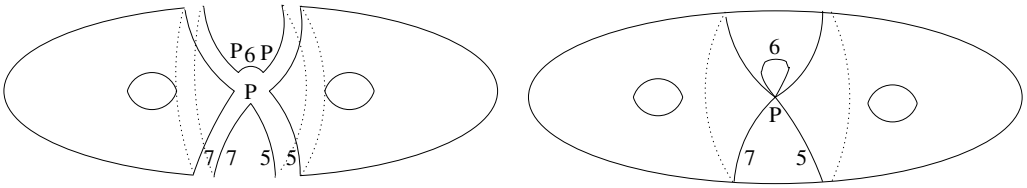


Figure 2.21: The bordered surface is being "born" by gluing a triangle along an edge (7) to the left, and another (5) to the right. Its third side becomes the closed boundary geodesic at the vertex P. The corresponding tessellated picture is fig. 2.18, but the similar procedure for the triangle given by the edges 7, 8, 13 is not indicated here.

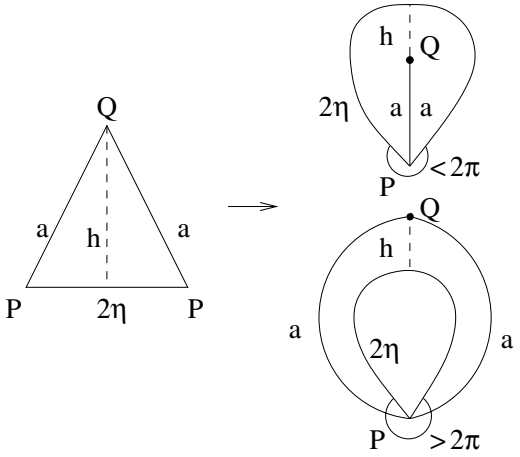


Figure 2.22: It is shown how to glue/cut off the isosceles to/from the bordered surface. On the right hand side the "drop" shape indicates the boundary geodesic on the surface at the basepoint P. The indicated angle is the sum of angles at the vertex P on the bordered surface. It is less (more) than 2π above (below) before gluing (cutting). It becomes 2π in both cases after the "operation". The boundary disappears in both cases, since the equal sides labeled by a are glued together.

of angle π. The isosceles degenerates to the zero area triangle with sides $(\eta, \eta, 2\eta)$, respectively. So we covered the interval $m \in (0, 1]$. However, we don't know how a similar construction works for negative deficit angle, which would presumably be the case for negative masses.

(ii) Presentation of S
We had a great advantage for matter free universes, namely, the surface S could be presented as a quotient: $= \mathbb{H}^2/G$, where G is the image of the fundamental group under the linear part of the holonomy homomorphism: the subgroup $SO_+(2,1)$ generated by the set of Lorentz transformations corresponding to the closed curves in the Cauchy surface. In case of massive particles with generic masses, there is no such presentation, all we have is a fundamental polygon with pasting conditions. The reason is, that elliptic elements have a fixed point in $\mathbb{H}^2$ (it is their axis).

(iii) Lorentz transformations, rest frame
Because of the previous point we have to be careful. We can still consider the basepoint of an OPT to be the unit normal vector n of the polygon and take the geodesics connecting the basepoint n to their images $\Lambda_i n$ (where Λ_i are the $SO(2,1)$ holonomies). Then the neighborhood of n is isometric to the neighborhood of the basepoint on the surface S: the geodesic arcs have the same lengths, emanate in the same order and enclose the same angles among each other as those on the surface. Conjugating the holonomies with an element $\Lambda \in SO(2,1)$ changes the data, and we would wish to prove that all we did was moving the basepoint. For this statement we would need $\mathbb{H}^2$ to be covered with isometric copies of the fundamental polygon, that is, the hyperbolic space would be the universal cover of the surface, which is not the case.

Suppose that some careful formulation of the statement is true, then the rest frame of a particle would be characterized by the coincidence of the vertex (P on the last few figures above) and the conical singularity Q. That is, the isosceles degenerates to a point, or in other words, the base point is the fixed point of the particle's holonomy.

(iv) $N > 1$, counter examples for the existence of S
If there are more particles it can happen that when we try to construct S by cutting and pasting triangles, the true angle at a vertex becomes negative. In other words we cannot construct it. An example (with

$g = 0$, $N = 4$) is depicted in fig. 2.24. The corresponding "would be" hyperbolic structure is shown in fig. 2.23. So there is a class of universes, where there is no underlying hyperbolic surface associated. How can one characterize this class? The moduli of the constructed counter example in fig. 2.24 can be varied infinitesimally to stay in that class, so it is clearly not a measure zero part of the phase space. Or is it possible, that one just needs the careful choice of the Lorentz frame (in other words the representative of the conjugacy class of G in $SO(2,1)$) to be able to construct S? If so, then what goes wrong, when we conjugate it back (cf. point 3.)?

(v) Teichmüller space for punctured surfaces.
The construction of S for $N = 0$ was useful because a hyperbolic surface can be considered as a point in the Teichmüller space $\mathcal{T}_g$, which is the configuration space of the gravity system. The Teichmüller space $\mathcal{T}_{g,N}$ of the punctured genus g surfaces can be defined as the set of punctured hyperbolic surfaces, which are complete and have finite area. The cone surfaces we construct are not complete (certain geodesics end in finite distance at the cone point(s)). However, the space of hyperbolic cone surfaces with fixed conical singularities is homeomorphic to the usual Teichmüller space [59], which correspond to the special case when all masses are equal to 1, thus the vertices of the isosceles are all at the boundary of D^2, an infinite distance away from the vertex of the triangulation. Note that in this case we still have the presentation $S = D^2/G$, since those generators of G, which correspond to the massless particles are all parabolic: they have fixed points only at the boundary of the unit disc. See [60] for more details.

2.11 Cosmological singularities

Finally, we will gain some insight into one of the original main motivations of this study: the asymptotic behaviour of the model near a Big Bang/Crunch singularity. Let us quote the conjectures of 't Hooft formulated in [44] on the basis of many numerical examples. Let us first fix the terminology. The singularity when the Cauchy surface has zero area (the sum of the area of the Euclidean polygons) is called Big Bang/Crunch, if it lies in the past/future. The time span of the universe is a finite interval I if it starts from a Big Bang and ends in a Big Crunch. It is a half line $\mathbb{R}_+$ if

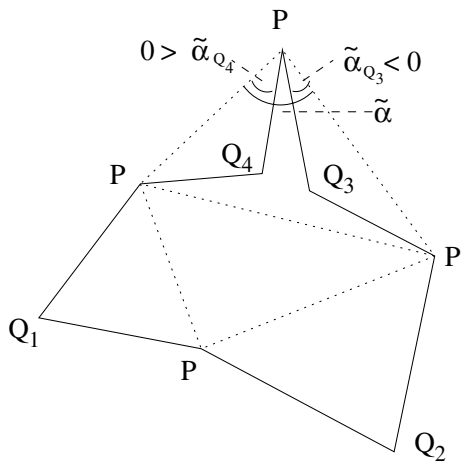


Figure 2.23: The fundamental (hyperbolic) polygon of S corresponding to a spherical topology with four particles. If we tune the masses of Q_4 and Q_3 smaller and smaller, the angle enclosed by the sides PQ_3 and PQ_4 will become negative. It means that the hyperbolic structure S cannot be constructed.

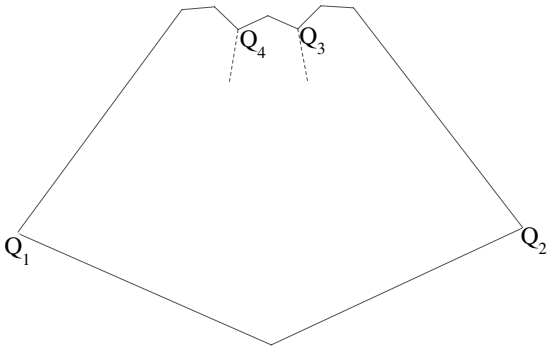


Figure 2.24: A faithful example of an OPT of a four particle universe with $g = 0$. It has two heavy ($m > 1/2$) particles Q_1 and Q_2 and two light ($m < 1/2$) particles Q_3 and Q_4. It corresponds to a situation when $\tilde{\alpha}_{Q_3}/2 + \tilde{\alpha}_{Q_4}/2 - \tilde{\alpha} < 0$ in fig. 2.23; it is the negative rotation angle between the indicated angle bisectors corresponding to a counterclockwise orientation. In this case we cannot construct the corresponding hyperbolic surface

only one of these is present and the universe is geodesically complete in the other direction of time. It is the full real line $\mathbb{R}$ if it is geodesically complete in both directions. 't Hooft has found that depending on the genus of the slice the following cases are possible[13]:

$g = 0$: I, $\mathbb{R}_+$

$g = 1$: $\mathbb{R}_+$

He conjectured that this should hold in general and extrapolated the conjecture to

$g > 1$: $\mathbb{R}_+$, $\mathbb{R}$.

For the matter-free case and $g > 0$ we saw that it is always $\mathbb{R}_+$. How does the example of I appear? First of all, we have to mention an important theorem derived in [50] stating that if all boost parameters have the same (negative) sign corresponding to contraction at $t = t^{(0)}$ then this condition remains satisfied for all times $t > t^{(0)}$. This is proven by checking all possible transitions. It means that an expanding universe with only positive boost parameters necessarily began with a Big Bang. There exist initial conditions such that a concave angle can be created in such a situation, or equivalently, a negative boost parameter. The time reversal of the mentioned theorem protects the positive sign of one boost parameter, but this is not an obstacle to end up in a shrinking universe. An example of the creation of a concave angle, hence a negative boost parameter, "surrounded" by positive ones is shown in fig. 2.25. The effect of the transition leading to this phenomenon on the level of S is the replacement of a glued triangle with a cut one.

The last thing we have to mention is the asymptotic behavior of the model near the singularities. It has been noticed by 't Hooft that approaching the Big Crunch the boost parameters grow unboundedly, which means that all edges shrink with the speed of light. Even if there has been a large number of polygons in an intermediate state of the time evolution, in this asymptotic region there remains only one, which takes an asymptotic shape of a triangle or sometimes a quadrangle. There is an infinite succession of transitions taking place, but only the *exchange* and the *hop* (first and fourth from above on page 37). We can interpret this on the level of S: those transitions are preferred in this regime which leads to longer geodesic arcs and they emanate in three "jets" from the basepoint. However, we cannot explain this behaviour.

[13]The static sectors are not taken into account. The sphere can have static configurations if the sum of the masses is 2. Then all boost parameters can be chosen to be zero. The empty torus have been discussed at the end of sec. 2.3. There is no static universe with $g > 1$ if the masses are in the range $m_i \in [0, 1)$, the Gauss-Bonnet theorem requires homogeneous vertices (with nonzero $\eta's$) to produce negative curvature.

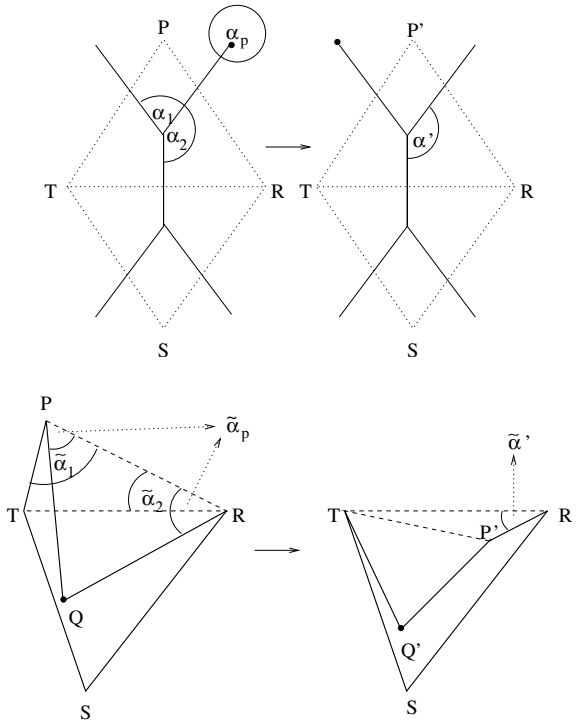


Figure 2.25: The top figure shows a hop transition leading to a concave angle. The solid lines are the edges of Γ and the dual edges are indicated by dotted lines. Below the corresponding pieces of the hyperbolic structure is shown. The isosceles corresponding to the particles are PRQ and TP'Q' on the left and right figure, respectively. The condition for the creation of a concave angle is $\alpha' = \alpha_1 + \alpha_2 + \alpha_p - 2\pi > \pi$. Rewriting this relation for the angles in the hyperbolic triangulation below we get $\tilde{\alpha}' = \tilde{\alpha}_1 + \tilde{\alpha}_2 + \tilde{\alpha}_p < 0$. We can see, that since $\tilde{\alpha} < 0$, the new triangle TRP' corresponding to the upper trivalent vertex will also be cut while the original triangle TPR was glued.

There is a great deal of knowledge now in the mathematics literature [27, 59, 61, 62, 63, 58] about classification, the cosmological time function and explicit construction of 't Hooft's polygonal spacetimes, see also appendix A.2. It seems that to be able to answer the unsolved problems discussed above and enumerated at the end of the previous section, one should explore the relation to the approach developed in this chapter and the explicit constructions of spacetimes dealt with in the above cited references. For example, the authors of [59] experienced difficulties similar to ours when trying to classify universes containing particles. Their conclusion was that the phase space of these spacetimes has many different components and it is likely, that even the dimension of some of them is less than the maximal $12g - 12 + 4N$.

Chapter 3

Polygon model from first order gravity

The previous chapter was devoted to the polygon model of classical 2+1 dimensional gravity at $\Lambda = 0$. Apart from the motivation to understand the structure of its phase space, another reason for studying it is the hope that the classically integrable model can be quantized and solved. The complete quantum theory in the polygon representation is still missing, although valuable qualitative observations are possible [47]. At the end of this chapter we will give a summary of the quantum aspects of the model and related approaches. What they have in common is their implications of spectra of geometry: time is discrete, space is continuous. How to interpret these results in a quantum theory of spacetime itself where the causal structure is expected to fluctuate, is a difficult and unsolved problem.

We will show in this section that the 't Hooft model can be derived from the first order formalism. The derivation of the phase space and the symplectic structure (2.34) will be given by means of gauge fixing and symplectic reduction. Various approaches addressing the first step, that is reducing the phase space spanned by the fields ω and e to a finite number of covariant variables, can be found in the literature [64, 34, 36, 35, 38]. These references cover the matter free case and the proper treatment of the inclusion of point particles. Another (lattice) approach has been developed by Waelbroeck et al. [32], who have compared their covariant model to the polygon model and derived the symplectic structure [65]. However, they needed the technical requirement of having the spacelike slice consisting of planar triangular faces. In this chapter we present the result of [48], whose the starting point is the smooth first order formalism, and the symplectic structure is derived in two ways. The first one is a partial gauge fixing of the covariant model of [35]. The second is a direct derivation of

the fields of the first order formalism: the boost parameters come from the spin connection, the lengths originate from the triad. The next section is devoted to summarizes the first step of the reduction from the field variables to the finite number of covariant ones: the $SO(2,1)$ holonomies and the three-vectors built from the triad, which are canonically conjugate to the holonomies. In section 3.2, the last step of the reduction is done, and we will arrive at the phase space of the polygon model.

3.1 Reduction to finite degrees of freedom

It has been explained in section 1.5 that in order to incorporate point particles with mass m, we need to add extra terms to the gravitational action. We will restrict ourselves to the case of spinless particles and choose the path of [36, 35], where the mass shell conditions of the particles of type (2.18) are added to the first order action with Lagrange multipliers. Some modifications will be needed with respect to [35], where only the special case was treated of trivial tangent bundle of the equal time surface, that is when it is the plane, the sphere or the torus. For our discussion there will be no such restriction, so Σ can have arbitrary genus.

If there are point particles present, the Chern-Simons formulation and the second order formulation of gravity are not even classically equivalent, due to the different structure of large gauge transformations versus large diffeomorphisms. In [8] an explicit example is constructed, where two configurations Φ_1 and Φ_2 in the Chern-Simons theory that are related by a non-infinitesimal local translation do not correspond to diffeomorphic spacetimes. Hence, the states Φ_1 and Φ_2 indistinguishable in Chern-Simons gravity are distinguishable in Einstein's gravity. One can understand this by considering the smooth path in a gauge orbit given by $t\,\rho^a, t \in [0,1]$, where ρ^a is a *non*-infintesimal parameter of a local translation (1.15). If Φ_1 and Φ_2 are connected by this path, then there is no guarantee that for a given value of t the configuration does not correspond to a degenerate triad/metric. The explicit example constructed in [8] is a four-particle universe with $\Sigma = \mathbb{R}^2$. We shall not deal with this issue here any further.

For the sake of simplicity, the discussion throughout this chapter will correspond to a one-polygon tessellation. Consider the equal time surface Σ as a simply connected region $P = \Sigma/\Gamma$, with the graph Γ of section 2, and pairwise identifications of the boundaries (which are the worldlines of the edges connecting vertices of Γ). Here, Γ does *not* consist of geodesic segments, but arbitrary smooth curves called edges connecting its vertices.

On shell, the spin connection and the triad are pure gauges

$$\omega^a_{jb} = (g^{-1}\partial_j g)^a{}_b \,, \quad e^a_j = (g^{-1})^a{}_b \,\partial_j f^b \,, \tag{3.1}$$

with potentials $g : \mathbb{R}^2 \to SO(2,1)$ and $f : \mathbb{R}^2 \to \mathbb{R}^{2,1}$ (the latter stands for Minkowski space). They are defined on the universal cover, and are not single valued on the surface Σ. However, the requirement that the fields should be single valued on Σ restricts the potentials. Most generally, on each pair of edges i_+ and i_- in ∂P (the boundary of the simply connected region P) which correspond to the same edge i of Γ the following conditions should hold

$$g\,|_{i_+} = \Lambda_i \, g\,|_{i_-} \,, \quad f^a\,|_{i_+} = \Lambda^a_{ib}\,(f^b\,|_{i_-} - a^b_i) \,, \tag{3.2}$$

with $\Lambda_i \in SO(2,1)$ and $a_i \in \mathbb{R}^{2,1}$. This is of course the constant gauge transformation which will lead to single valued fields (3.2) on the edges. We need to impose the consistency conditions

$$\mathbf{V}_j \equiv \Lambda^{\pm 1}_{k(j)}\,\Lambda^{\pm 1}_{l(j)}\,\Lambda^{\pm 1}_{m(j)} - \mathbf{1} = 0 \,, \tag{3.3}$$

where j runs through all 3-vertices of Γ, the tuple $(k(j), l(j), m(j))$ is in the appropriate order. Whether Λ or Λ^{-1} appears in the constraint is also straightforward when one fixes the labels i_+ and i_- for the edges. This way the constraints in (1.37) are solved everywhere, except for the 1-vertices. The reason is that the only possible locations for non-vanishing curvature of ω given by (3.2) are the vertices of Γ, but (3.3) makes the holonomy[1] around any three-vertex[2] trivial, which means that the curvature vanishes also there. In order for the 1-vertices to be point particles, we have to impose extra constraints

$$W_j \equiv \operatorname{tr}\Lambda_j - 2\cos m_j + 1 = 0 \,. \tag{3.4}$$

for each 1-vertex. To obtain these from an action principle, one needs to add the N terms of the form (3.4) to the action, where Λ is substituted with the holonomy around the particle as the functional of the spin connection. These terms have to be added to the action (1.34) with Lagrange multipliers to make sure that the massive particles move along timelike geodesics,

[1] Strictly speaking, we do not yet know if the Lorentz group element Λ_i is the holonomy along the dual closed curves corresponding to the edge i, but it is in fact the case as will be made clear below.

[2] Recall that the edges dual to the edges of Γ are closed curves and $k(j) \circ l(j) \circ m(j)$ is the closed curve (it would be the contour of the corresponding triangle, if the curves $k(j)$, $l(j)$, $m(j)$ were geodesics), which can be contracted to the three-vertex j.

massless particles[3] move along lightlike geodesics and that there are conical singularities at their locations. Since the holonomies are functionals of the connection ω, one should also check, how the equations of motion from the variation with respect to ω change. It turns out that the value of the extra Lagrange multiplier is fixed and it does not lead to new constraints. For a careful analysis of this issue, see [36].

A second set of variables is given by the following formula

$$E^a_{i_\pm} \equiv \int_{i_\pm} df^a = \int_0^1 ds\, \partial_l f^a \, \frac{d\alpha^l_{i_\pm}(s)}{ds} \,, \tag{3.5}$$

where the parametrization $s \mapsto \alpha_{i_\pm}(s)$ respects a chosen circular ordering of the boundary ∂P (that is, $\alpha_{i\pm}(1)$ follows by $\alpha_{i\pm}(0)$). The potential f defines an embedding of Σ to Minkowski space, and the "edge vector" $E^a_{i_\pm}$ is interpreted as the relative position vector in the background Minkowski space of the two ends of edge $i_\pm$ of Γ. We assume that the edges labeled by i_+ are oriented according to a global chosen orientation of ∂P and the orientation of their partners are always opposite, see fig.2.7 for an illustration. The edge vectors are not independent, but have to satisfy

$$E^a_{i_+} = -\Lambda^a_{i\,b}\, E^b_{i_-} \,, \tag{3.6}$$

where the minus sign is due to the opposite orientation. Finally there is a global constraint

$$C^a \equiv \sum_i \left(E^a_{i_+} + E^a_{i_-} \right) = 0 \,, \tag{3.7}$$

which must hold, because according to the definition (3.5), we have to integrate df around the closed contour of Γ which has no boundary, hence the integral is zero. Let us emphasize that the graph $\Gamma = \partial P$ does not consists of straight edges, but the vectors $E^a_{i\pm}$, which live in a background Minkowski space determine a closed contour due to (3.7) consisting of straight edges.

Let us remind the reader that the Poisson brackets are defined with the help of the symplectic form, which is a antisymmetric two-index tensor of maximal rank $\omega = \omega_{AB}\, dz^A dz^B$, ($z^A = (q, p)$, and the Poisson bracket of two functions on the phase space is $\{f(p,q), g(p,q)\} = \tilde{\omega}^{AB}\, \partial_A f\, \partial_B g$, where $\tilde{\omega}$ is the inverse of ω). If it is exact, that is, $\omega = d\Theta$ in the form notation, then the one-form Θ is called the symplectic potential. If the kinetic term in the action is $p\,\dot{q}$, then the symplectic potential is given by $p\,dq$ giving rise to the symplectic form $dp \wedge dq$ familiar from classical mechanics.

[3]The parameter m is proportional the deficit angle of the cone corresponding to a massive particle in its rest frame, and it is zero without any deficit angle interpretation for a massless particle as discussed in the previous chapter.

In our case, the action is given by (1.34) hence, canonical gravity in the first order formalism is an example with the above property. The symplectic potential thus reads

$$\Theta = 2\int_P d^2x\, \epsilon^{jk}\, \eta_{ab}\, d\omega_j^a\, e_k^b = \int_P d^2x\, \epsilon^{jk}\, \epsilon_{ac}^b\, \partial_j \left((dg\, g^{-1})^a{}_b\, \partial_k\, f^c \right) , \quad (3.8)$$

where we chose to work again with the units $16G\pi = 1$ for the rest of this chapter, in order to avoid writing down too many $2\pi's$. P is a "fundamental region" of the surface Σ, and the integration over it is well defined. To arrive at the right hand side, we exploited the general identity for commuting derivations ∂ and δ

$$\partial(x^{-1}\delta x) = x^{-1}\delta\, (\partial x x^{-1})x\ , \quad (3.9)$$

used Lorentz invariance and the identity $\epsilon^{jk}\partial_j\partial_k = 0$. Now, since the integrand is a total derivative, we can proceed to write

$$\Theta = \int_{\partial P} ds\, \epsilon_{ac}^b (dg\, g^{-1})^a{}_b\, \partial_s f^c\ . \quad (3.10)$$

This can be written as a sum of integrals along the edges. After introducing the notation $f_{i_\pm} \equiv f|_{i_\pm}$ and $g_{i_\pm} \equiv g|_{i_\pm}$ we find

$$\Theta = \sum_{i_+} \int_{i_+} ds\, \epsilon_{ac}^b \left((dg_{i_+}\, g_{i_+}^{-1})^a{}_b\, \partial_s f_{i_+}^c - (dg_{i_-}\, g_{i_-}^{-1})^a{}_b\, \partial_s f_{i_-}^c \right) , \quad (3.11)$$

where the integrals along the edges i_- have been replaced by those along i_+ and the minus sign is due to the opposite orientation of them. Using eq. (3.2) to express the quantities with the index + in terms of those with the index − one finds

$$\Theta = \sum_{i_+} \int_{i_+} ds\, \epsilon_{ac}^b\, (d\Lambda_i \Lambda_i^{-1})^a{}_b\, \partial_s f_{i_+}^c = \sum_{i_+} \epsilon_{ac}^b\, (d\Lambda_i \Lambda_i^{-1})^a{}_b\, E_{i_+}^c\ . \quad (3.12)$$

The equality is due to (3.5) and the fact that the Λ_i is constant. From this we can compute the Poisson brackets

$$\{\Lambda_{i\,b}^a, E_{j_+}^c\} = \delta_{ij}\, \epsilon_d^{ac} \Lambda_{i\,b}^d\ , \quad (3.13)$$

$$\{E_{i_\pm}^a, E_{j_\pm}^b\} = \delta_{ij}\, \epsilon_c^{ab} E_{i\pm}^c\ , \quad (3.14)$$

and the rest are zero. Note, that we treat the set $\{\Lambda_i, E_{i_+}\}$ as fundamental variables, only they appear in the reduced symplectic potential.

This symplectic potential coincides with that of Matschull and Welling in [36], in the context of one gravitating particle and the generalization for a multi-particle system in [35]. The calculation above "does not know" about topology, the only extra ingredient with respect to [35] was to choose the trivalent graph Γ in such a way, the $\Sigma\backslash\Gamma$ is simply connected and to impose the constraints (3.3).

The Poisson algebra above coincides also with that of Waelbroeck in [32]. His lattice model is derived by means of discretization from the first order variables. The E^a variables there are the edge vectors of the skeleton of the lattice. Two of them are associated to each edge, belonging to the two faces adjacent to the edge. The Λ variables there transform the two vectors, which belong to the same edge, into each other, the same way as in our case (cf. eq. (3.6)), and they are interpreted as parallel transport operators between the faces. Since the Poisson brackets of the basic variables are the same in our case as in the lattice model, we can conclude that if the lattice has only trivalent vertices, and, for the sake of our discussion, only one face, then the brackets of the constraints also coincide. Eq. (3.3) and (3.4) are first class constraints, they generate translations of the 3-vertices and reparametrization of the worldline of the particles The computations of the Poisson brackets can be found in [65]. The constraint (3.7) is also first class:

$$\{C^a, C^b\} = \epsilon^{ab}_{\ c} C^c \ , \tag{3.15}$$

it generates Lorentz transformations of the background Minkowski space.

What is the dynamics of the system? In order to have causal dynamics we have to impose that requirement that all edge vectors are spacelike. Then there is still no unique dynamics, since we have at hand a system with first class constraints and the Hamiltonian is a linear combination of the constraints

$$H = \sum_{i=1}^{6g-3+N} \epsilon^{b}_{ac} \mathbf{V}^{a}_{i\,b}\, N^c_i + \sum_{i=1}^{N} W_i\, N_i + N_a\, C^a. \tag{3.16}$$

Fixing the Lagrange multipliers does lead to a unique dynamics, but the actual spacetime does not depend on this choice. One may take $N^a = 0$, since this is just a Lorentz transformation of the frame or background Minkowski space. Then, H depends only on the holonomies, so they are constants of motion, since $\{\Lambda, \Lambda\} = 0$ and the evolution of the edge vectors is linear since $\{E, \Lambda\} \sim \Lambda$.

Let us briefly explain what is the geometrical meaning of a choice of the set $\{N^a_i, W_i\}$ of Lagrange multipliers in (3.16). One has to choose a time-like vector N^a_i at one of the three corners, corresponding to the 3-vertex

i, of the polygonal contour given by the edge vectors. It has to be parallel transported to the other two such corners with the appropriate Λ_j matrices (corresponding to the edges incident to the three vertex). One should "place" the axes of the Lorentz transformations Λ_i at the 1-vertices, multiplied with W_i. This way we have associated a timelike (or lightlike) vector to each corner of the polygonal contour. Then, the unique time evolution, which is generated by the Hamiltonian (3.16) amounts to sliding the corners along these vectors with an amount of proper time Δt. The slice of spacetime that has been constructed is a cylinder with a polygonal base. Its boundaries are pairwise identified by the holonomies. For a more detailed explanation of the time evolution and the fact that the constructed spacetime does not depend on the choice of the chosen Lagrange multipliers, see [33].

3.2 Gauge fixing and symplectic reduction

Waelbroeck and Zapata have shown [65] that from a triangular lattice version of their model, the polygon model can be reproduced. We will now show how the polygon model arises directly from the smooth first order formalism. First, we require by means of a gauge transformation that the initial time surface is mapped to a *planar* polygon by the potential f in Minkowski space. Then, the scalar variables will be introduced, which procedure leads to the solution of some of the constraints of the previous section. One can proceed in two ways to recover the symplectic structure (2.34) of the scalar variables of the polygon model. The first is further reduction of the symplectic potential given by formula (3.12). The second is the derivation of the Poisson brackets directly by appropriately expressing the scalar variables as functionals of the spin connection and the triad. Finally, in the last section we identify the induced constraints, which provide dynamics for the model.

The edge vectors $E^a_{i\pm}$ and the holonomies Λ_i are the finite number of covariant variables which contain the degrees of freedom of the system. One can obtain the polygon model by performing a local translation of (1.15) such that the polygonal contour is embedded in a spacelike plane in the background Minkowski space. In other words, after the gauge transformation, all edge vectors $E^a_{i_\pm}$ lie in a spacelike hyperplane. Denote the unit timelike normal to this hyperplane by n^a. Let us define the boost parame-

ters[4] and the lengths as follows

$$\cosh 2\eta_i = n_a \, \Lambda^a_{i\,b} \, n^b \ , \tag{3.17}$$

$$L_i \equiv \sqrt{E^a_{i_+} E^b_{i_+} \eta_{ab}} = \sqrt{E^a_{i_-} E^b_{i_-} \eta_{ab}}. \tag{3.18}$$

In the chosen gauge the only independent information contained in the edge vectors are their lengths. Indeed, planarity and (3.6) enables us to write in the frame where the normal of the polygon is purely timelike

$$E_{i-} = \pm L_i \, R(-\phi_i') \, e^{(1)} \ , \quad E_{i+} = \mp L_i \, R(\phi_i) \, e^{(1)} \ , \tag{3.19}$$

if the parametrization (2.8) is used and $e^{(1)} = (0, 1, 0)$. Now, we shall compute the symplectic potential departing from the rhs. of (3.12) in this coordinate system. Let us calculate the first piece containing the holonomy

$$\begin{aligned} \Lambda^{-1} d\Lambda \ &= \ R(\phi) \, B(2\eta) \, R\left(\tfrac{\pi}{2}\right) B(-2\eta) \, R(-\phi) \, d\phi' \\ &\quad + R(\phi) M R(-\phi) \, d(2\eta) + R\left(\tfrac{\pi}{2}\right) d\phi \ , \end{aligned} \tag{3.20}$$

where M is defined as the matrix with the following nonzero entries $M_{13} = M_{22} = M_{31} = 1$ and the identities $\frac{dR(x)}{dx} R(-x) = R\left(\frac{\pi}{2}\right)$ and $\frac{dB(x)}{dx} B(-x) = M$ have been used. Then, using Lorentz invariance, we arrive at the following expression for Θ

$$\begin{aligned} &\textstyle\sum_i \mp L_i \, \epsilon^b_{ac} \left(R(\pi/2)^a_{\ b} \left(B(2\eta_i) \, e^{(1)} \right)^c d\phi_i' + \right. \\ &\left. M^a_{\ b} \, e^{(1)c} d(2\eta_i) + R(\pi/2)^a_{\ b} \, e^{(1)c} \, d\phi_i \right) \ . \end{aligned} \tag{3.21}$$

The coefficients of $d\phi$ and $d\phi'$ vanish and we we are left with

$$\Theta = 2 \sum_i L_i \, d\eta_i \ . \tag{3.22}$$

The sign ambiguity can be fixed by the systematic choice of the labels $+$ and $-$ for an edge and its partner. Note that the eq. (1.36) contains an additional factor of 4π with respect to (2.34), since we used different units ($4G = 1$), convenient for the previous chapter.

[4] Let us remark that for an edge i, it is possible to choose the gauge in such a way that Λ_i is a pure boost hence $\mathrm{tr}\Lambda_i = 2\cosh 2\eta_i + 1$ as proposed in [66], but this cannot be done globally for all edges. It would mean that the (cosh of) the boost parameters are equal to traces of holonomies along closed curved, thus gauge invariant. We will see that this is not the case.

To understand better the origin of the result above, we present an alternative derivation of the Poisson brackets directly from the original fields of the theory. The Lorentz group element Λ_i is the holonomy along the dual edge ζ_i. More precisely, if the points a_- and a_+ are the starting and endpoints of a curve freely homotopic to the dual edge ζ with the same orientation, then the following relations hold

$$U^a_{\zeta\, b}(a_-, a_+) = g^{-1}(a_-)\, g(a_+) = g^{-1}(a_-)\, \Lambda\, g(a_-) = g^{-1}(a_+)\, \Lambda\, g(a_+)\ , \tag{3.23}$$

due to (3.1). Note that the index i is omitted from the above formula to keep the notation simple. The holonomy can be also expressed in the usual way as the functional of the spin connection given by eq. (1.52).

Let us now choose a fixed edge i of Γ and denote its (oriented) dual by ζ. The expression for the Poisson bracket of the associated length and boost parameter reads

$$\begin{aligned} \{L, \cosh 2\eta\} &= \{L, n_a(a_+)\, U^a_{\zeta\, b}(a_-, a_+)\, n^b(a_+)\} = \\ & \epsilon_{jk}\, \eta^{cd}\, n_a(a_+)\, n^b(a_+) \int_P d^2x\, \frac{\delta L_i}{\delta e^c_j(x)}\, \frac{\delta U^a_{\zeta\, b}(a_-, a_+)}{\delta \omega^d_k(x)}\ , \end{aligned} \tag{3.24}$$

where $n^a(a_+) = g^{-1}(a_+)^a{}_b\, n^b$. In general, the functional derivative of the holonomy can be written as

$$\frac{\delta U_\zeta}{\delta \omega^b_j(x)} = \int_\zeta ds\, \frac{dx^j(s)}{ds}\, \delta\left(x(s), x\right) U_{\zeta_1} T_b\, U_{\zeta_2}\ , \tag{3.25}$$

where ζ_1 is the segment of ζ until the point x, ζ_2 is the segment of ζ from the point x, and $x(s)$ is the parametrization of the curve ζ. The variation of the length with respect to the triad can be written using equations (3.1), (3.5) and (3.18) as

$$\frac{\delta L}{\delta e^c_j(x)} = \frac{E_{b+}}{L} \int_i d\tau\, g(\tau)^b{}_c\, \frac{dx^j}{d\tau}\, \delta(x(\tau), x)\ , \tag{3.26}$$

where $x(\tau)$ is a parametrization of the edge. Now we substitute (3.25) and (3.26) to (3.24) and integrate out one Dirac delta. If the integral for ζ is done first from the point $s_- = x(\tau)$ to its image s_+ along the curve ζ, then we have $\zeta_1 = \emptyset$ and $\zeta_2 = \zeta$, so we can write

$$\{L, \cosh 2\eta\} =$$

$$\int_i d\tau \int_\zeta ds \left[\epsilon_{jk} \frac{dx^j(\tau)}{d\tau} \frac{dx^k(s)}{ds} \delta\left(x(\tau), x(s)\right) \right] \times$$
$$\frac{E_+^b \, g(s_+)_b{}^a}{L} n_c(s_+) \left(T_a \, U(s_-, s_+)\right)^c{}_b \, n^b(s_+) \; .$$

The expression outside the square bracket is independent of s, since we can rewrite it using the notation $(v \times w)_a = \epsilon_{abc} \, v^b \, w^c$ as

$$(g^{-1}(s_+) \, E_+)^a (g^{-1}(s_+) \, n \times g^{-1}(s_+) \Lambda \, n)_a = E_+^a (n \times \Lambda \, n)_a \; , \tag{3.27}$$

because both the vector product and the Minkowski scalar product are invariant under Lorentz transformations. The integral of the square bracket above is the homotopy invariant oriented intersection number of the edge i and the closed curve ζ. It is one for the pairs (i, ζ_i) and zero otherwise. The last step is to verify the following formula

$$n \times \Lambda_i n = \sinh(2\eta_i) \, E_{i_+} \; . \tag{3.28}$$

E_{i_+} is clearly orthogonal to both n and $\Lambda_i n$ due to (3.6) and the scalar factor in the above equation is also not difficult to check using the definition (3.17). We have thus arrived at the desired result:

$$\{L_i, \cosh 2\eta_j\} = \delta_{ij} \sinh 2\eta_i \rightarrow \{L_i, 2\eta_j\} = \delta_{ij} \; . \tag{3.29}$$

At the intermediate stage of the symplectic reduction, the finite number of covariant variables Λ_i and E_i have to satisfy a number of constraints, given by (3.3), (3.4), (3.6) and (3.7). Changing to the scalar variables L_i, η_i solves the first three, if the triangle inequalities

$$|\eta_i| + |\eta_j| \geq |\eta_k| \tag{3.30}$$

are satisfied, for every vertex $\mathcal{V} = \{i, j, k\}$, and for every permutation of $\{i, j, k\}$ and also the global relation (2.38). The explanation is the following: eq. (3.3) and (3.4) determine the angles of the polygon in terms of η_i, eq. (3.6) determines the direction of the edges. Or, phrasing differently, if we choose the boost parameters and lengths as our fundamental variables, then the equations (3.3), (3.4) and (3.6) are not constraints among the variables, but fix the angles of the polygon.

In [65] it is argued that the Gauss constraint (3.7) is solved by introducing the scalar variables and there remain induced curvature constraints, but, as we explained above, the situation is on the contrary: the Gauss law and only that remains after the reduction. Eq. (3.7) becomes eq. (2.42) in

the scalar variables, with only two independent components, the real and imaginary part of $\tilde{C}$.

If the global constraint (2.33)) is taken to be the Hamiltonian H, then it generates the time evolution of the polygon model that preserves the closure of the polygon by construction, thus

$$\{H, \tilde{C}\} = \sum_i \dot{L}_i \, z_i = 0 \, , \tag{3.31}$$

where the dot now indicates the time evolution induced by H. The fact that the Poisson bracket of $\tilde{C}$ with itself yields a linear combination of the constraints is difficult to check explicitly. However, the correct number of degrees of freedom can only be recovered if all the independent constraints H, and $\tilde{C}$ are first class and generate time translation and two independent Lorentz transformations, respectively (since a rotation is factored out by using lengths and angles instead of edge vectors). In other words the constraint algebra must close and the most general time evolution is generated by the following Hamiltonian:

$$H' = a\, H + c\, \tilde{C} \, , \tag{3.32}$$

where a (c) is an arbitrary real (complex) function on the phase space. Now, similarly to the covariant description, we can take $b = c = 0$, since $\tilde{C}$ generates Lorentz transformations of the frame and $a \neq 1$ corresponds to reparametrization of the time coordinate.

3.3 Status of quantization

In this section we briefly present some features of quantization of the above models and related approaches. The covariant model of finite degrees of freedom given by $\{\Lambda_i, E_i\}$ has been quantized in [36], where the resulting quantum geometry was found and the effect of the point particle on it. That paper deals only with one (spinless) particle, which captures some essential features of quantum gravity. Let us recall the structure of the phase space. The momentum Λ lives in the group manifold $SO(2,1)$ and the position variable[5] E^a lives in Minkowski space $\mathbb{R}^{2,1}$. The Poisson brackets $\{E^a, E^b\} = \epsilon^{ab}_{\ c} E^c$ coincide with the commutator of the Lie algebra $so(2,1)$, so the full phase space is the cotangent bundle $T_*SO(2,1)$. However, in contract with the usual situation of theories whose the phase space is a cotangent bundle,

[5]It comes from the triad which is the "square root" of the configuration variable, the metric. That is why we call it position.

it is the momenta here, which live in the curved group manifold and the positions take the role of the cotangent vectors.[6]. Let us examine the direct consequences of this structure of the phase space in the quantum theory.

The Hilbert space is $L^2(SO(2,1), d\Lambda)$, $d\Lambda$ being the invariant measure on the group. We consider the momentum representation, where the momentum operators act by multiplication and the positions can be realized as the Lie algebra elements acting as (the left invariant) derivatives. They represent the Poisson brackets of their classical counterparts

$$[\hat{E}^a, \hat{E}^b] = -i\, l_P\, \epsilon^{ab}_{\ c}\, \hat{E}^c, \tag{3.33}$$

where $l_P = 4\pi G\hbar$ is a new length scale [7]. The operator of the invariant length can be defined in terms of these operators as $\sqrt{E^a E_a}$. Since $\hat{E}^a$ form a basis in the Lie algebra $so(2,1)$, the length operator is given by the Casimir of the gauge group. Let us study its spectrum. There are two series of irreducible unitary representations of $SO(2,1)$. The discrete series corresponds to negative eigenvalues $\{-j(j+1) \,|\, j \in \mathbb{N}\}$ and the principal series corresponds to positive eigenvalues $[1/4, \infty)$ of the Casimir operator[8]. Since timelike vectors have negative, spacelike vector have positive length squared, we associate the discrete series with timelike vectors, the continuous series with spacelike ones. The conclusion is that time is quantized, whereas space is continuous in the quantum theory.

Let us turn now to another promising approach to quantizing gravity, which is called loop quantum gravity. This theory is based on the idea that loops or graphs can describe excitations of quantum space. It is a canonical approach, where quantum spacetime appears as the evolution of these graphs [68]. We briefly describe the kinematical Hilbert space of the loop approach to three dimensional quantum gravity, define the length operator and describe its spectrum.

In the connection representation the states are given by the so-called *cylindrical functions*[9], which have the following properties: (i) they are

[6]Note that $SO(2,1) \cong SL(2,\mathbb{R})$ is (locally) anti de Sitter. Writing the components of the an $SL(2,\mathbb{R})$ matrix a as $a_{11} = x - w$, $a_{12} = z + y$, $a_{21} = z - y$, $a_{22} = x + w$, one finds that the unit determinant condition is $-w^2 - z^2 + x^2 + y^2 = 1$, which gives the three dimensional hyperbolic space also known as anti de Sitter embedded in $\mathbb{R}^{3,1}$.

[7]The non-commuting coordinates lead to the appearance of a shortest distance $\sim l_P$ in the quantum theory without breaking Lorentz invariance. This is a characteristic feature of theories called doubly special relativity. 2+1 gravity with one particle, the theory we are describing, is a notable example of doubly special relativity [67].

[8]Note that there is a third set of continuous representations called the supplementary series with the spectrum $[0, 1/4]$ of the Casimir. They are excluded because they do not appear in the Fourier decomposition of square integrable functions on the group.

[9]The space of cylindrical functions may seem ad hoc, but it is shown that they are dense

associated to closed graphs with E (oriented) edges, (ii) they have E arguments which depend on the holonomies $U_i(\omega)$ of the spin connection associated to the paths along the edges labeled by $i = 1..E$i, and (iii) they are invariant under the gauge transformations

$$f(g_{t(1)}\, U_1(\omega)\, g^{-1}_{s(1)}, ..., g_{t(E)}\, U_E(\omega)\, g^{-1}_{s(E)}) = f(U_1(\omega), ..., U_E(\omega)), \qquad (3.34)$$

where $s(i)$ and $t(i)$ denotes the source and the target vertex of edge i and g_x is the Lorentz group element acting in the tangent space T_xM. Since the holonomy is an element of the Lorentz gauge group, f is a function on (E copies) of the group. These functions can be decomposed into a sum (or integral) over irreducible unitary representations of the group. For example,

$$f(g) = \sum_j c^{(j)mn} D^{(j)}_{mn}(g), \qquad (3.35)$$

where j labels the irreducible representations and D is the matrix of the group element g in that representation. The Hilbert space structure is defined by the scalar product given by the invariant (Haar) measure on the group.[10] A basis in the Hilbert space is realized by the so-called *spin networks*, which are also associated to graphs. Each edge of the graph of a spin network is decorated with an irreducible representation and each vertex with an invariant tensor contracting the indices corresponding to the vertex.[11]

Two basic operators acting on the Hilbert space are the multiplication operator of the holonomy of the connection (which adds an extra disconnected loop to the graph) and the derivative operator with respect to the connection. Their quantum algebra provides the quantization of the classical Poisson brackets of the corresponding classical quantities, the holonomy and the triad, respectively. Using this algebra we can evaluate the action of the length operator on a spin network. Let us write the length of a curve embedded in the equal time surface Σ_t parametrized by $\alpha : [0, 1] \to \Sigma_t$ as

$$l = \int_0^1 d\, s \sqrt{\eta_{ab}\, e^a_i\, e^b_j\, \frac{d\alpha^i}{ds}\, \frac{d\alpha^j}{ds}}\ . \qquad (3.36)$$

in the space of gauge invariant functions of the spin connection.

[10] To compare two cylindrical functions corresponding to different graphs, one can view both as defined on the union of the two graphs but independent from the other graph.

[11] For example, for a two-valent vertex, the only possibility is that the representations on the incident edges coincide and the invariant *intertwiner* is given by the Kronecker-delta symbol. For a three valent vertex it is given by the Clebsch-Gordan coefficients, and the restriction of admissible representations on the incident edges is that one appears in the decomposition of the tensor product of the other two. For more general situations the invariant tensor is not unique.

The action of the derivative operator associated to the triad on the holonomy $\mathcal{P}\exp\int_i T_a\,\omega^a$ of edge i of the spin network is non-zero only if α intersects i. Suppose that α intersects the spin network once at edge i. Then, the action of the triad is the insertion of the generator T_a at the intersection point. Therefore, the quantum operator $\hat{l}$ corresponding to the quantity (3.36) is ($\hbar G$ times) the square root of the Casimir operator of $SO(2,1)$. The edges of the spin network are "in definite irreducible representations", so the operator $\hat{l}$ has diagonal action. Its eigenvalue is the square root of the eigenvalue of the Casimir in the representation[12] assigned to i.

We again have the situation that the spectrum of spacetime is determined by the representation theory of the Lorentz group $SO(2,1)$. One may wonder how to interpret the discrete series with negative eigenvalues in the canonical formalism at hand, where a "good" equal time surface is spacelike. However, as explained in [69], in the canonical formulation underlying the loop approach, such a requirement has nowhere been explicitly imposed. The spectrum is thus the same as the one of the Matschull/Waelbroeck approach discussed in the beginning of this section. Finding the physical Hilbert space, that is the subspace of the one just described annihilated by a Hamiltonian constraint operator, is still an unsolved problem.

Another approach, where a similar picture of quantum spacetime emerges is the so-called *spin foam* approach. It represents a generalization of the usual path integral. In these models transition amplitudes can be calculated by summing over branched colored surfaces, called spin foams. A spin foam can be thought of as the evolution of a spin network. Therefore, the coloring refers to assigning irreducible unitary representations of the gauge group to faces, invariant tensors to the edges and an amplitude (which is intimately connected to the action of the Hamiltonian constraint) to the vertices. If the action is of BF type[13], the amplitude can be derived in a systematic way from the classical action [70] by means of a discretization of the path integral $\int \mathfrak{D}\omega\,\mathfrak{D}e\,e^{iS}$ based on a triangulation which is the simplex dual to the spin foam. For topological theories without local degrees of freedom like gravity in three dimensions, the resulting amplitudes are independent of the triangulation.

What is the quantum spacetime represented by a spin foam in three

[12] In general, one has to regularize the operator $\hat{l}$ by partitioning the curve α into small pieces and add the eigenvalues corresponding to the action of $\hat{l}$ associated to the segments.

[13] If the action can be written as $S=\int(B\wedge F+\Phi(B))$, where F is the curvature of some gauge connection, B is an $D-2$ form and $\Phi(B)$ is a function of the B field then it is called of BF type.

dimensions? Since the action reads $\int \mathrm{Tr}\,(e \wedge F)$, it is of BF type so one can apply the procedure sketched above. The corresponding spin foam model is called the Ponzano-Regge model [71]. Since the face of a spin foam is dual to an edge of the triangulation in three dimensions, the irreducible representation of $so(2,1)$ assigned to it is interpreted as the quantization of the dual edge. Thus, spin foam faces give lengths to edges they intersect and the length spectrum, as we can already guess, is given by the (square root of) the spectrum of the Casimir operator.

There has also been an attempt to quantize the polygon model [47]. The canonical Poisson brackets between the boost parameters and lengths can easily be represented by operators on a Hilbert space. However, the Hamiltonian (2.33) is a non-polynomial functions of the variables and the requirement that the wave functions should be invariant under the non-physical transitions appearing in the model are serious obstacles in formulating a consistent quantum theory. Nevertheless, several important observations were made in [47] regarding three dimensional quantum gravity. One of them is that time must be quantized: since the Hamiltonian is an angle (the sum of the deficit angles), it is defined only modulo 2π. The consequence of this is that the unitary time evolution operator $U = \exp(iHt)$ is well defined only if t is integer. A similar argument shows that the length L, being canonically conjugate to a hyperbolic angle, must remain continuous in the quantum theory.

Hence, in three dimensions a consistent picture emerges from the study of possible quantization procedures. It is supported further by the works [72, 73]. In the former it is explicitly shown, that the Ponzano-Regge model, which incorporates spinning particles as well, is a generalization of Waelbroeck's quantization of [74]. In [73] it is proven that the Ponzano-Regge model is also equivalent to the Chern-Simons quantization.

Some of the above ideas are applied also to four dimensional general relativity. There the theory is not topological, and its quantization is a much more difficult task. The canonical quantization can again be based on the Hilbert-Palatini action, which has a local $SO(3,1)$-Lorentz symmetry. Its Hamiltonian formulation, is plagued with second class constraints, but it is possible to introduce a gauge fixing such that they do not appear. If one follows this path, due to Ashtekar and Barbero [75], the local gauge symmetry reduces to $SU(2)$ and the equal time surface is always spacelike. The area spectrum[14] is given by the $SU(2)$ Casimir and is thus entirely

[14] In four dimensions the spin networks intersect two dimensional surfaces in the equal time hypersurface, so the area operator is the analogue of the length operator in three dimensions. Or, using space time language, spin foam faces are dual to triangles of the triangulation, which acquire area from the representation assigned to the face.

discrete.

On the other hand, there are proposals for spin foam models based on the full Lorentz group [76]. The area spectrum predicted by the spin foam models of Lorentzian gravity is similar to what was found in three dimensions. It is given by a Casimir of the gauge group, which is now $SO(3,1)$. It is discrete for timelike surfaces, and continuous for spacelike surfaces. There is thus a potential clash between the canonical approach and the path integral. However, it was recently argued, that the gauge fixing of $SO(3,1)$ to $SU(2)$ is not allowed, because it leads to a quantum theory where the diffeomorphism symmetry is broken. In particular, this can explain the appearance of a non-physical parameter (called Immirzi parameter) in the spectrum of observables found in the standard loop quantization procedure based on the gauge group $SU(2)$. At the same time another quantization procedure was suggested, which preserves all symmetries [77]. This approach is still in its infancy and a number of issue remain to be resolved. The connection is non-commutative due to the complicated Dirac brackets (which should always replace the Poisson brackets in the presence of second class constraints), and the kinematical Hilbert space is missing. Nevertheless, a state space can be constructed and the area spectrum is computed explicitly in [78]. It is independent of the Immirzi parameter. Moreover, it was shown that a subspace of this state space coincides with that induced by spin foams[15], both in the case of spacelike [80] and timelike foliations [81].

[15]How a spin foam induces spin network states on a foliation of spacetime is described in [79].

Appendix

A.1 Models of hyperbolic space

There are two isometric models of the hyperbolic space $\mathbb{H}^2$, which we use in this thesis. We introduces them and their relations in this appendix. The first model is the upper sheet of the two-sheeted hyperboloid also known as anti de Sitter space in two dimensions given by $\{(t,x,y)|\ -t^2+x^2+y^2=-1\}$, with the metric inherited from Minkowski space. It can be shown that this metric is positive definite and its curvature $R=-1$. This space has a boundary, the space of lightlike directions, which is topologically a circle. The geodesics are hyperbolas obtained by the intersection of any plane going through the origin of Minkowski space and the hyperboloid itself. The isometry group is the (identity component of the) Lorentz group $SO_+(2,1)$.

If we use the stereographic projection from the South Pole $(-1,0,0)$ to the plane $t=0$, we obtain another model of $\mathbb{H}^2$ called the Poincaré disc $D^2=\{z\in\mathbb{C}|\ |z|<1\}$. The induced metric is given by

$$ds^2=\frac{4|dz|^2}{(1-|z|^2)^2}. \tag{A.1}$$

The geodesics of this model are diameters and Euclidean circle segments perpendicular to the boundary, which is the unit circle. The isometry group is $PSU(1,1)$, the group of two by two complex matrices g with unit determinant which satisfies the relation $gmg^\dagger=m$ where $m=\mathrm{diag}(-1,1)$ modulo multiplication of all entries of g with -1. They act on the disc as $z\mapsto(a_{11}+a_{12})/(a_{21}z+a_{22})$, where a_{ij} are the entries of the matrix. We can see that the matrix A and $-A$ have the same action.

We will need the form of the boost and rotation matrices in $SU(1,1)$.

They are given by

$$b(\xi) := \begin{pmatrix} \cosh\frac{\xi}{2} & \sinh\frac{\xi}{2} \\ \\ \sinh\frac{\xi}{2} & \cosh\frac{\xi}{2} \end{pmatrix}, \quad r(\phi) = \begin{pmatrix} \exp(i\frac{\phi}{2}) & 0 \\ \\ 0 & \exp(-i\frac{\phi}{2}) \end{pmatrix}. \quad \text{(A.2)}$$

Any $PSU(1,1)$ matrix can be written as

$$g = r(\phi)\, b(\xi)\, r(\phi')\ , \quad \text{(A.3)}$$

and the isomorphism with SO(2,1) is given by $B(\xi) \leftrightarrow b(\xi)$ and $R(\phi) \leftrightarrow r(\phi)$, cf. (2.9).

Finally, we mention some properties of hyperbolic triangles, which we use in the text. Triangles in the interior of hyperbolic space have angles $\tilde{\alpha}_i$, the sum of which is always smaller then π. The area of a triangle is $\pi - \sum_i \tilde{\alpha}_i$. If a vertex of a triangle is at the boundary then the corresponding angle is zero, and the the lengths of the sides emanating from that vertex is infinite. For illustrations and more details see e.g. [2].

A.2 Classification of globally hyperbolic spacetimes

In this appendix we summarize some of the main results of the paper [27], briefly discuss some related recent material and give the definition of some notions appearing in the main text. Classical solutions of three dimensional gravity are specified as (X, G) structures, where X is a manifold, on which the spacetime M is modeled locally by specifying homeomorphisms from open sets of M to X and the transition functions between two overlapping charts are elements of the group G, which is usually the isometry group of X. The sign of the cosmological constant Λ determines X, it is Minkowski space $\mathbb{R}^{2,1}$ for $\Lambda = 0$, (anti) de Sitter space

$$\{(w,t,x,y) \in \mathbb{R}^4 : -t^2 + x^2 + y^2 + (-)w^2 = -1\}$$

for positive (negative) Λ. The corresponding isometry groups are $ISO(2,1)$ for $\Lambda = 0$, $SO_+(2,2)$ for $\Lambda < 0$ and $SO_+(3,1)$ for $\Lambda > 0$. The *holonomy map* $\rho : \pi_1(M) \to G$ and the *developing map* $D : \tilde{M} \to X$ are associated to an (X, G) structure, where the tilde denotes the universal cover. The former is a homomorphism unique up to conjugation with an element of G, the latter is a local isometry, unique up to post composition with an element of G. They satisfy the relation

$$D(\gamma x) = \rho(\gamma) D(x) \quad \text{(A.4)}$$

where $x \in \tilde{M}$ and $\gamma \in \pi_1(M)$ acts in the usual way (γx is the endpoint of the lift of γ if x is a fixed chosen lift of the basepoint of $\pi_1(M)$). There is a large variety of such manifolds for each case. Let us restrict ourselves to spacetimes with the topology $\Sigma \times I$, which is proven to be the only possibility of any Lorentzian flat or anti de Sitter manifold with closed spacelike boundary and probably holds for the de Sitter case as well [27]. Then M is foliated by closed spacelike surfaces of the topology of Σ. One should impose the additional and also natural maximality criterion of M being a *domain of dependence*, which notion has been defined in footnote 3 on page 26.

After these preliminaries, we can discuss the classification of these spacetimes for negative or zero cosmological constant. For the flat case it is shown that the linear part $G \subset SO_+(2,1)$ of the holonomy group $\rho(\pi_1(M))$ (that is the Lorentz part of the Poincaré group elements) gives a discrete subgroup of $SO_+(2,1)$ isomorphic to $\pi_1(M)$, which yields a compact quotient space $\mathbb{H}^2/G$, (which is homeomorphic to Σ. The converse is also true, for any discrete and faithful homomorphism $f : \pi_1(\Sigma) \to SO_+(2,1)$ there are a spacetimes, whose corresponding linear holonomies are given by f. Furthermore, there exists a future (or past) complete convex domain U in Minkowski space such that $M = U/\rho(\pi_1(M))$ in the interior of the future light cone of the origin.

The spacetimes are thus parametrized by the Teichmüller space $\mathcal{T}_g$ and the set of maps: $a : \pi_1(M) \to \mathbb{R}^{2,1}$, the translation part of the Poincaré holonomy. Let us expand on this description of the phase space. A definition of Teichmüller space is the set of equivalence classes of constant curvature metrics (or conformal structures), where two metrics in a class are related by a diffeomorphism in the identity component of the diffeomorphism group.The hyperbolic surfaces S and S' of genus $g > 1$ are equivalent if, when presented as a quotient of $\mathbb{H}^2$ by the discrete groups $G \cong G'$ ($\cong \pi_1(S)$), G and G' are conjugate in $SO_+(2,1)$ [57]. The maps $a : \pi_1(M) \to \mathbb{R}^{2,1}$ satisfy the property

$$a(\alpha\beta) = a(\alpha) + f(\beta)a(\beta) \,, \tag{A.5}$$

which follows from the composition of the Poincaré group elements. a and a' is equivalent if there is is a constant vector $v \in \mathbb{R}^{2,1}$ such that

$$a(\gamma) - a'(\gamma) = v - f(\gamma)v \,. \tag{A.6}$$

It means that the two holonomies are conjugate by a constant translation yielding the same spacetime. As we explained in the first chapter, the Teichmüller space of a genus g surface is homeomorphic to $\mathbb{R}^{6g-6}$, see also

[57] for a detailed proof. The space of equivalence classes of the a maps[1] has the same dimension. Indeed, assigning a vector to every independent generator of the fundamental group means $3(2g-1)$ parameters which determines a for all elements of $\pi_1(S)$ by (A.5) and one needs to take the quotient with (A.6) to arrive at $6g-6$.

The summary of the above discussion is that the maximal globally hyperbolic spacetimes admitting a closed spacelike surface of genus $g > 1$ have two families (future complete ones and their time reversal) parametrized by $\mathbb{R}^{12g-12}$, and they are quotients of convex domains in Minkowski space by the image of the Poincaré holonomy. They contain an initial (or final) singularity with a rich structure investigated with the help of so called cosmological time function [82].

There has been a growing activity in the mathematics literature since the paper [27] appeared. Results there have been generalized to higher dimensions [63, 83], which has somewhat limited interest in physics. In a recent paper [18] it is shown, that all vacuum spacetimes or at least parts of them are related by Wick rotation or "rescaling" regardless of the sign of the constant or whether the metric is Riemannian or Lorentzian. The cosmological time has a central role in the construction and the initial singularities have the "same structure" in all cases. This result is likely to gain great importance in physics as well. Another important work in progress needs to be mentioned. Barbot studied explicit constructions of the 't Hooft slicings of vacuum spacetimes and obtained the rigorous result that each such spacetime admits an OPT [58].

Spacetimes containing particles are much less known. In general they cannot be presented as a quotient of a domain with respect to the holonomy, and the developing map is not injective. Benedetti and Guadagnini studied a special class of spacetimes with particle masses equal to integer multiples of π, and the dynamics is given by the *geodesic flow* on the cotangent bundle of Teichmüller space [61]. They consider generic multiparticle universes in [59]. They depart from a hyperbolic or flat surface with conical singularities given by the masses of the particles and obtain 3-dimensional spacetimes by means of *Minkowskian suspensions*. If dl_E^2 or dl_H^2 is the Euclidean or hyperbolic metric on the surface, then a Minkowskian suspension is given by $ds^2 = -dt^2 + dl_E^2$ or $ds^2 = -dt^2 + t^2 dl_H^2$ with $t > 0$ yielding flat spacetimes. Then appropriate deformations of these structures yield a large class of spacetimes. They reveal some facts about the moduli space of these spacetimes, but even the dimension or the number

[1] It is the *group cohomology* of *cocycles* (maps satisfying the property (A.5)) modulo *coboundaries* (maps which can be written as the rhs. of (A.6)).

of topological components are unknown. If the cone angles deviate "too much" from zero, then similar problems arise that we have faced in section 2.9, when trying to construct the invariant surface for certain configurations of too small particle masses. In these cases even the dimension of the (actual component if the) moduli space is unclear, the rough estimate $6g-6+2N \leqq d \leqq 12g-12+4N$ is given in [59].

Let us conclude with the remark that the space of physically distinct universes with particles is hardly known, and one might need a higher level of rigour to establish the polygon model of spacetimes (if all admit one) in order to trust e.g. the conjectures concerning the initial or final singularity.

A.3 Boost parameters from Teichmüller space

In this appendix we will explain how to obtain an independent set of initial values for the boost parameters η. For simplicity and illustrative purposes, we will discuss an example of genus 2. Since we will not make use of the symmetry structure of this particular case, the generalization to higher genus is immediate.

We fix the triangulation by choosing the graph Γ on fig. 2.10, leading to the triangulation of fig. 2.11. We will use the convention of multiplying loops from left to right. The numbers indicate the outgoing ends of the loops, and $i = 1, \ldots, 4$ label the generators b_i of the fundamental group satisfying $b_1\, b_2\, b_1^{-1} b_2^{-1} b_3\, b_4\, b_3^{-1} b_4^{-1} = 1$. The homotopy classes of the remaining closed curves can be obtained by composing the fundamental generators and their inverses, leading to[2]

$$\begin{array}{lcl} 5 & \rightarrow & b_2\, b_1\, b_2^{-1} b_1^{-1} \\ 6 & \rightarrow & b_1\, b_2^{-1} \\ 7 & \rightarrow & b_2\, b_1\, b_2^{-1} \\ 8 & \rightarrow & b_3\, b_4^{-1} \\ 9 & \rightarrow & b_4\, b_3\, b_4^{-1}\ . \end{array} \tag{A.7}$$

Now we use the faithful representation of $\pi_1(S)$ in $PSU(1,1) \cong SO_+(2,1)$ given explicitly in [84] in terms of the so-called Fenchel-Nielsen coordinates. They are a set of length and angle variables (l_k, τ_k), $k = 1, \ldots, 3g-3$, which parametrize the Teichmüller space $\mathcal{T}_g$ globally. One can pick an arbitrary element $(l, \tau) \in \mathbb{R}_+^{3g-3} \times \mathbb{R}^{3g-3} \cong \mathcal{T}_g$, plug it into the formulae for

[2] Our and Okai's [84] convention for multiplying elements (curves) of the fundamental group is from right to left, so the product $\nu_1\nu_2$ in $\pi_1(S)$ in our case is mapped to $g_2\, g_1$ in the corresponding Lie group.

the generators $g_i \in PSU(1,1)$ and compute the combinations for the group elements corresponding to the remaining curves in the triangulation (in our specific example, the curves labeled 5, 6, 7, 8 and 9). Now, knowing the group generators and the combinatorial information (the order of the edges going around the polygon), we can identify the boost parameters and compute the angles as well. How the boosts are supplemented by a set of length variables L_i has been described in sec. 2.8.6 above. Altogether this amounts to an explicit algorithm for constructing a set of initial data for a (2+1)-dimensional universe from any element of $\mathbb{R}_+^{6g-6} \times \mathcal{T}_g$.

A.4 The complex constraint

The proof that for an OPT there is always a Lorentz frame[3], in which the complex constraint (2.42) admits a solution rests on the following facts.

(i) The complex vector z_i defined below eq. (2.42) points to the angle bisector of the geodesic loop corresponding to Λ_i at the basepoint.

(ii) The velocity at P of the unique arc connecting the base point $P \in S$ to the unique smooth closed geodesic corresponding to Λ_i points in the same direction.

(iii) One can find a basepoint, where these velocities are not contained in a half plane.

(iv) In practice, one proceeds by finding an element $\Lambda \in SO_+(2,1)$ which corresponds to the desired change of basepoint. The original vectors z_i can be read off from $\Lambda_i \in G$. The new coefficients z_i' will be determined from the conjugated generators $\Lambda^{-1}\Lambda_i\Lambda$, and they will not lie in a half plane by the above arguments.

In the following we will use the Poincare disc model D^2 of hyperbolic space as defined in appendix A.1. How the Lorentz group elements Λ_i correspond to the $g_i \in PSU(1,1)$ and how that isometry group acts on D^2 is also explained in that appendix.

Take the universal cover where the origin $0 \in D^2$ is mapped to the base point $P \in S$. The in- and out-going ends of the loop corresponding to g_i on S can be associated with the geodesic arcs connecting 0 with $g_i 0$ and 0

[3]Equivalently, a suitable basepoint in S, or a suitable $h \in PSU(1,1)$ to conjugate all the generators with.

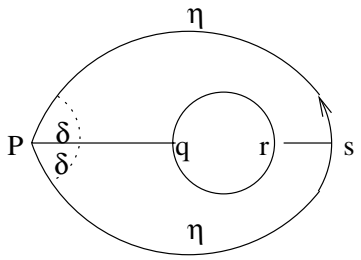


Figure A.1: On the surface S, the loop connecting the basepoint P to itself via point s is the loop in the triangulation corresponding to g_i. The circle in the middle is the unique smooth geodesic which lies in the same homotopy class as the loop. The unique arcs connecting p to the circle and the circle to the loop such that the angles at q, r and s are right angles, create two isometric quadrilaterals.

with $g_i^{-1}0$. Since the geodesics through $0 \in D^2$ are Euclidean straight lines, and since

$$g_i\, 0 = \tanh \eta_i \exp(i\,\phi_i)\ , \quad g_i^{-1}0 = \tanh \eta_i \exp\left(i\,(\pi - \phi_i')\right)\ , \tag{A.8}$$

it is clear that the angle bisector of $g_i\, 0$ and $g_i^{-1}0$ points in the same direction as z_i.

The next step is to establish the validity of fig. A.1, namely, that the two quadrilaterals $Pqrs$ are isometric. The figure shows the smooth geodesic (inner circle) freely homotopic to the geodesic loop PP and the unique smooth geodesic arcs Pq and rs connecting two, and perpendicular to them in the points q, r and s. All we need to show is that the arc Pq is the angle bisector of the velocities of the in- and outgoing ends of the loop (denoted by η in the figure). We refer to [57] for a detailed proof. The properties of the various geodesics in D^2 are illustrated by fig. A.2. The smooth geodesic of fig. A.1 is mapped to the smooth straight line on the bottom, and the geodesic loop to the periodic non-smooth curve at the top. The curves at q, r and s meet at right angles.

The situation in D^2 is as follows. We are given the set of $6g-3$ elements $\{g_i\}$ of the fundamental group G. Each g_i has its so-called *axis*, that is, the geodesic which is left invariant by g_i. On the disk D^2, an axis has the form of a circle segment whose ends are perpendicular to the disk boundary. (This fact is not reproduced in fig. A.2, where the axis is represented by the straight line at the bottom.) Under the universal cover, an axis is mapped (infinitely many times) to the smooth geodesic on S corresponding to g_i. The unique geodesic arc on D^2 from the origin which is perpendicular to

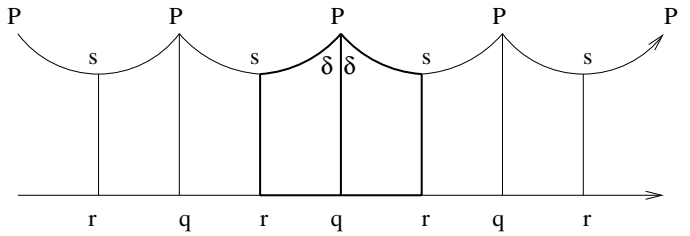


Figure A.2: Lift to D^2 of a (periodically extended) geodesic loop at P (top) and of the associated homotopic smooth geodesic (bottom). The geodesic arc rs is the unique perpendicular connecting the geodesic loop PP to the smooth closed geodesic qq. The unique geodesic arc connecting P to qq which is perpendicular at q is the angle bisector of the loop and its inverse at P. The thick lines indicate one of the (infinitely many) pairs of isometric quadrilaterals $Pqrs$ (each with three right angles) which are mapped to Fig.A.1 under the universal covering map f.

one of these axes is mapped to the arc on S connecting the basepoint to the smooth geodesic in question.

Suppose now, that the basepoint is the intersection of two smooth geodesics corresponding to two of the generators, say g_1 and g_2. In other words, the geodesic polygon obtained by cutting the surface along the geodesic loops corresponding to g_i is a normal canonical polygon. Note that such a polygon is always convex. Take a universal covering, where a lift of the basepoint is the origin of the Poincaré disc D^2. Then, the axis of g_1 and g_2 are diameters of the disc. The coefficients z_1 and z_2 are identically zero in this case, since the geodesic corresponding to g_1 and g_1^{-1} emanate exactly in the opposite direction from the basepoint (and similarly for g_2 and g_2^{-1}). However, if we move the basepoint infinitesimally from its location to a direction different from the axes, then we find nonzero z_1 and z_2. There are essentially four different location for a basepoint as shown in fig. A.3. It is easy to show that one of the four locations will always be such that z_1, z_2 and, say, z_3 are in a generic location: the complex constraint admits a solution. Furthermore, since we moved the basepoint only infinitesimally, the resulting geodesic polygon remains convex.

If the desired change of the basepoint is achieved by $h \in SU(1,1)$, (that is $h0 = T$), then we can determine the new boost parameters from the new generators $h\, g_i\, h^{-1}$.

We have thus completed the proof that one can always find a Lorentz frame in which the complex constraint admits a solution. Note that with this constructive proof one always produces convex polygons.

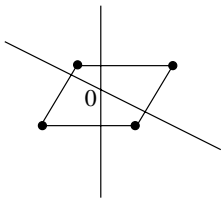


Figure A.3: The two long lines are parts of the axes of g_1 and g_2. They are diameters of the Poincaré disc, 0 denotes the origin. There is a tiny parallelogram, the edges of which are perpendicular to the axis they intersect. This parallelogram is divided into four regions by the axes. If we move the basepoint to any of the four regions, then z_1 and z_2 will point in the direction of the edges of the parallelogram, away from its corner in that region, indicated by a dot. One of the four possibilities is always sufficient for having z_1 z_2 and z_3 in a generic position, not contained in any half plane.

A.5 Eliminating polygons by gauge-fixing

In this appendix we will show how to gauge-transform a given $(\tilde{\gamma}, F)$ (a geodesic triangulation $\tilde{\gamma}$ of some F-polygon) to a configuration $(\tilde{\gamma}', F)$ which is equivalent to a configuration $(\tilde{\gamma}'', F-1)$ with one polygon fewer.[4] The induced map $(\tilde{\gamma}', F) \mapsto (\tilde{\gamma}'', F-1)$ amounts to deleting three edges and one base point from $\tilde{\gamma}'$ but does not change $\tilde{\gamma}'$ as a point set.

A gauge transformation of $(\tilde{\gamma}, F)$ is an action of $\times_F PSU(1,1)$, where each of the F copies of $PSU(1,1) \equiv SO(2,1)$ acts independently as follows. If $\tilde{e}_{ij}$ denotes an oriented edge connecting base points $\tilde{P}_i$ and $\tilde{P}_j$ on S, we will call its associated group element $g_{ij} \equiv g_{ji}^{-1}$. To give an example, for the edge $\tilde{e}_{12}$ connecting $\tilde{P}_1$ and $\tilde{P}_2$, we have $\bar{P}_2 = g_{21}\bar{P}_1$ for the inverse images in D^2. A generic gauge transformation is given by an F-tuple $(h_1, h_2, \ldots, h_F) \in \times_F PSU(1,1)$, acting by group multiplication at the end points of edges according to

$$g_{ij} \mapsto h_i \, g_{ij} \, h_j^{-1}, \quad i, j \in \{1, 2, \ldots, F\}. \tag{A.9}$$

There will usually be several edges linking a base point to itself (implying $i = j$), which can be taken care of by introducing an extra label for the edges and group elements, $e_{ii}^{(k)}$ and $g_{ii}^{(k)}$. At the level of the frames X_j, $1 < j < F$, and assuming for the moment no obstructions, this gauge transformation corresponds to a simultaneous rotation of the frames, $X_j \to \tilde{X}_j = \Lambda_j X_j$ via the canonical isomorphism $h_i \sim \Lambda_i$ of subsection 3.5. If two

[4] For the purposes of this appendix, we will mean by $\tilde{\gamma}$ a geodesic triangulation together with a definite length assignments η_i to its edges, and by γ the underlying topological triangulation.

neighbouring frames were related by a Lorentz transformation Λ_{21} before the gauge transformation, $X_2 = \Lambda_{21} X_1$, the matching condition afterwards will be $\tilde{X}_2 = \tilde{\Lambda}_{21} \tilde{X}_1$, with $\tilde{\Lambda}_{21} = \Lambda_2 \Lambda_{21} \Lambda_1^{-1}$, cf. (A.9).

Let us adopt the notation $\bar{P}$ for points in D^2 and $\tilde{P}$ for their images in S under the universal cover, and suppose that $h_i \bar{P}_i \in D^2$ is mapped to $\tilde{P}'_i$, $i = 1, 2$. Then the boost parameter $2\eta'_{21}$ read off from the group element $h_2^{-1} g_{21} h_1$ is the length of the geodesic arc connecting $\tilde{P}'_1$ to $\tilde{P}'_2$, which is freely homotopic to the original arc connecting $\tilde{P}_1$ to $\tilde{P}_2$ with length η_{21}. We conclude that also in generic multi-polygon universes gauge transformations amount to moving the base points without changing the topology of the graph γ.

Suppose now that we perform a gauge transformation on a single frame only, say, X_1. The effect on the geometry of the graph $\tilde{\gamma}$ will be a motion of the base point $\tilde{P}_1$ and a modification of the edges starting or ending at $\tilde{P}_1$. The magnitude of the change will be chosen as g_{21}, corresponding to the geodesic arc $\tilde{e}_{21} \in \tilde{\gamma}$ connecting $\tilde{P}_1$ to $\tilde{P}'_1 = \tilde{P}_2$. Its effect can be written as follows:

$$\begin{array}{llll}
\tilde{e}^{(i)}_{11} : \tilde{P}_1 \to \tilde{P}_1 & \mapsto & \tilde{e}^{(i)'}_{11} : \tilde{P}'_1 \to \tilde{P}'_1, & g^{(i)}_{11} \mapsto g_{21}\, g^{(i)}_{11}\, g_{21}^{-1} \\
\tilde{e}^{(i)}_{k1} : \tilde{P}_1 \to \tilde{P}_k, & \mapsto & \tilde{e}^{(i)'}_{k1} : \tilde{P}'_1 \to \tilde{P}_k, & g^{(i)}_{k1} \mapsto g^{(i)}_{k1}\, g_{21}^{-1} \\
\tilde{e}^{(i)}_{1k} : \tilde{P}_k \to \tilde{P}_1, & \mapsto & \tilde{e}^{(i)'}_{1k} : \tilde{P}_k \to \tilde{P}'_1, & g^{(i)}_{1k} \mapsto g_{21}\, g^{(i)}_{1k} \\
\tilde{e}^{(i)}_{kl} : \tilde{P}_l \to \tilde{P}_k, & \mapsto & \tilde{e}^{(i)'}_{kl} : \tilde{P}_l \to \tilde{P}_k, & g^{(i)}_{kl} \mapsto g^{(i)}_{kl},
\end{array} \tag{A.10}$$

assuming $k, l \neq 1$. Writing $\tilde{P}'_1$ in (A.10) is meant to emphasize that despite $\tilde{P}'_1 = \tilde{P}_2$ one has to keep track of whether an end point of a curve in $\tilde{\gamma}$ corresponds to the base point labeled by 1 or by 2. Consider now one of the two triangles in $\tilde{\gamma}$ which share the geodesic arc $\tilde{e}_{21}$ (we have dropped the counting label i for simplicity). It consists of the arcs $\tilde{e}_{21} : \tilde{P}_1 \to \tilde{P}_2$, $\tilde{e}_{1k} : \tilde{P}_k \to \tilde{P}_1$ and $\tilde{e}_{2k} : \tilde{P}_k \to \tilde{P}_2$, and we have $g_{2k} = g_{21}\, g_{1k}$ for the corresponding group elements. The action of the above transformation on these arcs and group elements reads

$$\begin{array}{llll}
\tilde{e}_{21} : \tilde{P}_1 \to \tilde{P}_2 & \mapsto & \tilde{e}'_{21} : \tilde{P}'_1 \to \tilde{P}_2, & g_{21} \mapsto g_{21}\, g_{21}^{-1} = 1 \\
\tilde{e}_{1k} : \tilde{P}_k \to \tilde{P}_1 & \mapsto & \tilde{e}'_{1k} : \tilde{P}_k \to \tilde{P}'_1 = \tilde{P}_2, & g_{1k} \mapsto g_{21}\, g_{1k} = g_{2k} \\
\tilde{e}_{2k} : \tilde{P}_k \to \tilde{P}_2 & \mapsto & \tilde{e}'_{2k} : \tilde{P}_k \to \tilde{P}_2, & g_{2k} \mapsto g_{2k}.
\end{array} \tag{A.11}$$

In other words, arc $\tilde{e}_{21}$ has shrunk to length zero (the trivial curve), arc $\tilde{e}_{1k}$ has been transformed to coincide with $\tilde{e}_{2k}$, and arc $\tilde{e}_{2k}$ has been left untouched. The new geodesic triangle with sides $\tilde{e}'_{21}$, $\tilde{e}'_{1k}$ and $\tilde{e}'_{2k}$ is degenerate. The same is true for the other triangle that shared the edge $\tilde{e}_{21}$. In order to obtain the reduced graph $(\tilde{\gamma}'', F-1)$ from $(\tilde{\gamma}', F)$, we delete the

redundant base point $\tilde{P}'_1$ and arc $\tilde{e}'_{21}$, as well as one arc of the pair $(\tilde{e}'_{1k}, \tilde{e}'_{2k})$, and one arc from the corresponding pair of the neighbouring triangle. Note that $\tilde{\gamma}'' = \tilde{\gamma}'$ as point sets, but that $\tilde{\gamma}'$ has one trivial and two double edges.

Bibliography

[1] G. 't Hooft. Quantization of space and time in 3 and in 4 space-time dimensions. In *Quantum fields and quantum space time, Cargese*, pages 151–163, 1996. arXiv:gr-qc/9608037.

[2] W. P. Thurston. *The Geometry and Topology of Three-Manifolds*. Princeton University Press, 1997.

[3] E. Witten. (2+1)-dimensional gravity as an exactly soluble system. *Nucl. Phys. B*, 311:46–78, 1988.

[4] P. A. M. Dirac. *Lectures on Quantum Mechanics*. Belfer Graduate School of Science, Yeshiva University Monograph Series 2, 1964.

[5] S. Deser, R. Jackiw, and G. 't Hooft. Three-Dimensional Einstein Gravity: Dynamics of Flat Space. *Annals of Physics*, 152:220–235, 1984.

[6] A. Achucarro and P. Townsend. Chern-Simons Action For Three-Dimensional Anti-De Sitter Supergravity Theories. *Phys. Lett. B*, 180:89, 1986.

[7] T. W. B. Kibble. Lorentz Invariance And The Gravitational Field. *J. Math. Phys.*, 2:212–221, 1961.

[8] H.-J. Matschull. On the relation between 2+1 Einstein gravity and Chern Simons theory. *Class. Quant. Grav.*, 16:2599–2609, 1999. arXiv:gr-qc/9903040.

[9] E. Witten. Topology Changing Amplitudes in (2+1)-Dimensional Gravity. *Nucl. Phys. B*, 323:113, 1989.

[10] G. T. Horowitz. Topology Change In Classical And Quantum Gravity. *Class. Quant. Grav.*, 8:587–602, 1991.

[11] S. Carlip and R. Cosgrove. Topology change in (2+1)-dimensional gravity. *J. Math. Phys.*, 35:5477–5493, 1994. arXiv:gr-qc/9406006.

[12] S. Carlip. *Quantum Gravity in 2+1 Dimensions*. Cambridge University Press, 1998.

[13] A. Hoyosa and K. Nakao. 2+1 Dimensional Quantum Gravity. Hiroshima University preprint RRK 89-16, 1989.

[14] S. Carlip. Observables, Gauge Invariance and Time in 2+1 Dimensional Quantum Gravity. *Phys. Rev. D*, 42:2647–2654, 1990.

[15] M. Banados, C. Teitelboim, and J. Zanelli. The Black hole in three-dimensional space-time. *Phys. Rev. Lett.*, 69:1849–1851, 1992. arXiv:hep-th/9204099.

[16] H.-J. Matschull. Black hole creation in 2+1-dimensions. *Class. Quant. Grav.*, 16:1069–1095, 1999. arXiv:gr-qc/9809087.

[17] K. Krasnov. Black Hole Thermodynamics and Riemann Surfaces. *Class. Quant. Grav.*, 20:2235–2250, 2003. arXiv:gr-qc/0302073.

[18] R. Benedetti and F. Bonsante. Wick rotations in 3D gravity: ML(H2)-spacetimes. arXiv:math.DG/0412470.

[19] R. Arnowitt, S. Deser, and C. W. Misner. Gravitation: an introduction to current research. In L. Witten, editor, *The dynamics of general relativity*. Wiley, New York, 1962.

[20] Bryce S. DeWitt. Quantum Theory Of Gravity. 1. The Canonical Theory. *Phys. Rev.*, 160:1113–1148, 1967.

[21] A. Ashtekar. New Variables For Classical And Quantum Gravity. *Phys. Rev. Lett.*, 57:2244–2247, 1986.

[22] T. Thiemann. QSD IV 2+1 Euclidean Quantum Gravity as a model to test 3+1 Lorentzian qunatum gravity. *Class. Quant. Grav.*, 15:1249–1280, 1998. arXiv:gr-qc/9705018.

[23] H. M. Farkas and I. Kra. *Riemann Surfaces*. Springer-Verlag, 1980.

[24] V. Moncrief. Reduction of the Einstein equations in (2+1)-dimensions to a Hamiltonian system over Teichmüller space. *J. Math. Phys.*, 30:2907–2914, 1989.

[25] A. Hosoya and K.-I. Nakao. (2+1)-dimensional pure gravity for an arbitrary closed initial surface. *Class. Quant. Grav.*, 7:163–176, 1990.

[26] R. Anderson, V. Moncrief, and A. Tromba. On the global evolution problem in 2+1 gravity. *J. Geom. Phys.*, 23:191–205, 1997. arXiv:gr-qc/9610013.

[27] G. Mess. Lorentz spacetimes of constant curvature. Institute des Hautes Etudes Scientifiques preprint IHES/M/90/28.

[28] R. Gambini and A. Trias. Gauge dynamics in the C-representation. *Nucl. Phys. B*, 278:436–448, 1986.

[29] A. Ashtekar, V. Husain, C. Rovelli, J. Samuel, and L. Smolin. (2+1)-quantum gravity as a toy model for the (3+1) theory. *Class. Quant. Grav.*, 6:L185–L193, 1989.

[30] A. Ashtekar and R. Loll. New loop representations for (2+1) gravity. *Class. Quant. Grav.*, 11:2417–2534, 1994. arXiv: gr-qc/9405031.

[31] R. Loll. Independent loop invariants for (2+1) gravity. *Class. Quant. Grav.*, 12:1655–1662, 1995. arXiv: gr-qc/9408007.

[32] H. Waelbroeck. 2+1 lattice gravity. *Class. Quant. Grav.*, 7:751–769, 1990.

[33] W. Waelbroeck and F. Zertuche. Homotopy invariants and time evolution in (2+1)-dimensional gravity. *Phys. Rev. D*, 50:4966–4981, 1994. arXiv:gr-qc/9401021.

[34] W. G. Unruh and P. Newbury. Solution to 2+1 gravity in the dreibein formalism. *Phys. Rev. D*, 48:2686–2701, 1993. arXiv:gr-qc/9307029.

[35] H.-J. Matschull. The Phase Space Structure of Multi Particle Models in 2+1 Gravity. *Class. Quant. Grav.*, 18:3497–3560, 2001. arXiv:gr-qc/0103084.

[36] H.-J. Matschull and M. Welling. Quantum Mechanics of a Point Particle in 2+1 Dimensional Gravity. *Class. Quant. Grav.*, 15:2981–3030, 1998. arXiv:gr-qc/9708054.

[37] J. Louko and H.-J. Matschull. The 2+1 Kepler problem and its quantization. *Class. Quant. Grav.*, 18:2731–2784, 2001. arXiv:gr-qc/0103085.

[38] P. de Sousa Gerbert. On Spin and (Quantum) Gravity in 2+1 Dimensions. *Nucl. Phys. B*, 346:440–472, 1990.

[39] G. 't Hooft. Nonperturbative Two Particle Scattering Amplitudes In (2+1)-Dimensional Quantum Gravity. *Comm. Math. Phys.*, 117:685–700, 1988.

[40] N. Koehler, F. Mansouri, C. Vaz, and L. Witten. Two Particle Scattering In The Chern-Simons-Witten Theory Of Gravity In (2+1)-Dimensions. *Nucl. Phys. B*, 348:373–389, 1991.

[41] L. Cantini, P. Menotti, and D. Seminara. Hamiltonian structure and quantization of 2+1 dimensional gravity. *Class. Quant. Grav.*, 18:2253–2276, 2001. arXiv: hep-th/0011070.

[42] L. Cantini, P. Menotti, and D. Seminara. Proof of Polyakov conjecture for general elliptic singularities. *Phys. Lett. B*, 517:203–209, 2001. arXiv:hep-th/0105081.

[43] Max Welling. Gravity in 2+1 dimensions as a Riemann-Hilbert problem. *Class. Quant. Grav.*, 13:653–680, 1996. arXiv:hep-th/9510060.

[44] G. 't Hooft. The evolution of gravitating point particles in 2+1 dimensions. *Class. Quant. Grav.*, 10:1023–1038, 1993.

[45] Z. Kádár and R. Loll. (2+1) gravity for higher genus in the polygon model. *Class. Quant. Grav.*, 12:2465–2491, 2003. arXiv:gr-qc/0312043.

[46] R. Franzosi and E. Guadagnini. Topology and classical geometry in (2+1) gravity. *Class. Quant. Grav.*, 13:433–460, 1996.

[47] G. 't Hooft. Canonical quantization of gravitating point particles in 2+1 dimensions. *Class. Quant. Grav.*, 10:1653–1664, 1993. arXiv: gr-qc/9305008.

[48] Z. Kádár. Polygon model from first order gravity. *Class. Quant. Grav.*, 22:809–823, 2005. arXiv:gr-qc/0410012.

[49] G. 't Hooft. Causality in (2+1)-dimensional gravity. *Class. Quant. Grav.*, 9:1335–1348, 1992.

[50] G. 't Hooft. Classical N-particle cosmology in 2+1 dimensions. *Class. Quant. Grav. Suppl.*, 10:79–91, 1993.

[51] M. Welling. *Classical and Quantum Gravity in 2+1 Dimensions*. PhD thesis, Utrecht University, 1997.

[52] S. Deser, R. Jackiw, and G. 't Hooft. Physical Cosmic Strings Do Not Generate Closed Timelike Curves. *Phys. Rev. Lett.*, 68(3):267–269, 1992.

[53] J. R. Gott III. Closed Timelike Curves Produced by Pairs of Moving Cosmic Strings: Exact Solutions. *Phys. Rev. Lett.*, 66(9):1126–1129, 1991.

[54] H.-R. Hollmann and R.-M. Williams. Hyperbolic geometry in 't Hooft's approach to (2+1) dimensional gravity. *Class. Quant. Grav.*, 16:1503–1518, 1999.

[55] L. Mosher. A user's guide to the mapping class group: once punctured surfaces. *Geometric and computational perspectives on infinite groups, DIMACS Series*, 25:101–174, 1996. arXiv: math.GT/9409209.

[56] H. Zieschang, E. Vogt, and H.-D. Coldewey. *Surfaces and planar discontinuous groups*. Springer Lecture Notes in Math. 835, 1980.

[57] P. Buser. *Geometry and spectra of compact Riemann surfaces*. Birkhäuser, 1992.

[58] T. Barbot. in preparation.

[59] R. Benedetti and E. Guadagnini. Geometric Cone Surfaces and (2+1) - Gravity coupled to Particles. *Nucl. Phys. B*, 588:436–450, 2000. arXiv:gr-qc/0004041.

[60] R. Penner. The decorated Teichmüller space of punctured surfaces. *Comm. Math. Phys.*, 113:299–339, 1987.

[61] R. Benedetti and E. Guadagnini. Classical Teichmüller Theory and (2+1) gravity. *Phys. Lett. B*, 441:60–68, 1998.

[62] Francesco Bonsante. *Deforming the Minkowskian Cone of a Closed Hyperbolic Manifold*. PhD thesis, University of Pisa, 2005.

[63] T. Barbot. Globally Hyperbolic Flat Spacetimes. arXiv:math.MG/0402257.

[64] A. Yu. Alexeev and A. Z. Malkin. Symplectic structure of the moduli space of flat connection on a Riemann surface. *Comm. Math. Phys.*, 169:99, 1995. arXiv:hep-th/9312004.

[65] H. Waelbroeck and J. A. Zapata. 2+1 covariant lattice theory and 't Hooft's formulation. *Class. Quant. Grav.*, 13:1761–1768, 1996. arXiv: gr-qc/9601011.

[66] E. R. Livine. *Loop quantum gravity and spin foam: covariant methods for the nonperturbative quantization of general relativity*. PhD thesis, Marseille, CPT, 2003. arXiv: gr-qc/0309028.

[67] L. Freidel, J. Kowalski-Glikman, and L. Smolin. 2+1 gravity and Doubly Special Relativity. *Phys. Rev. D*, 69:044001, 2004. arXiv:hep-th/0307085.

[68] C. Rovelli and L. Smolin. Loop representation of quantum general relativity. *Nucl. Phys. B*, 331:80, 1990.

[69] L. Freidel, E. R. Livine, and C. Rovelli. Spectra of length and area in 2+1 Lorentzian loop quantum gravity. *Class. Quant. Grav.*, 20:1463–1478, 2003. arXiv: gr-qc/0212077.

[70] L. Freidel and K. Krasnov. Spin Foam Models and the Classical Action Principle. *Adv. Theor. Math. Phys.*, 2:1183–1247, 1999. arXiv:hep-th/9807092.

[71] L. Freidel. A Ponzano-Regge model of Lorentzian 3-Dimnesional gravity. *Nucl. Phys. Proc. Suppl*, 88:237–240, 2000. arXiv:gr-qc/0102098.

[72] L. Freidel and D. Louapre. Ponzano-Regge model revisited. I: Gauge fixing, observables and interacting spinning particles. arXiv:hep-th/0401076.

[73] L. Freidel and D. Louapre. Ponzano-Regge model revisited. II: Equivalence with Chern-Simons. arXiv:gr-qc/0410141.

[74] H. Waelbroeck. Canonical quantization of (2+1) gravity. *Phys. Rev. D*, 50:4982–4992, 1994. arXiv:gr-qc/9401022.

[75] M. Gaul amd C. Rovelli. Loop Quantum Gravity and the Meaning of Diffeomorphism Invariance. *Lect. Notes. Phys.*, 541:277–324, 1999. arXiv:gr-qc/9910079.

[76] J. W. Barrett and L. Crane. A Lorentzian signature model for quantum general relativity. *Class. Quant. Grav.*, 17:3101–3118, 2000. arXiv:gr-qc/9904025.

[77] S. Alexandrov and E. R. Livine. SU(2) loop quantum gravity seen from covariant theory. *Phys. Rev. D*, 67:044009, 2003. arXiv:gr-qc/0209105.

[78] S. Alexandrov and D. Vassilevich. Area spectrum in lorentz covariant loop gravity. *Phys. Rev. D*, 64:044023, 2001. arXiv:gr-qc/0103105.

[79] S. Maran. Relating Spin Foams and Canonical Quantum Gravity: $(n-1)+1$ formulation of nD spin foams. arXiv:gr-qc/0412011.

[80] E. R. Livine. Projected Spin Networks for Lorentz connection: Linking spin foams and loop gravity. *Class. Quant. Grav.*, 19:5525–5542, 2002. arXiv:gr-qc/0207084.

[81] S. Alexandrov and Z. Kádár. Timelike surfaces in Lorentz covariant loop gravity and spin foam models. arXiv:gr-qc/0501093.

[82] R. Benedetti and E. Guadagnini. Cosmological Time in (2+1) Gravity. *Nucl. Phys. B*, 613:330–352, 2000.

[83] F. Bonsante. Flat Spacetimes with Compact Hyperbolic Cauchy Surfaces. arXiv:math.DG/0311019.

[84] T. Okai. An explicit representation of the Teichmüller space as holonomy representations and its applications. *Hiroshima Math. J.*, 22:259–271, 1992.

Samenvatting

Tegenwoordig speelt zwaartekracht een belangrijke rol in de theoretische fysica. De gravitationele interactie tussen macroscopische objecten wordt zeer successvol beschreven door Einstein's algemene relativiteitstheorie. Voor het beschrijven van het begin van het universum en de vorming alsook de verdamping van zwarte gaten zal echter een quantum theorie van de zwaartekracht nodig zijn. Onze hoop is dat we zo een quantumtheorie kunnen opzetten in de versimpelde context van drie ruimtetijd dimensies. De procedure om van een klassieke- naar een quantumtheorie te komen is niet uniek. Verschillende klassieke formuleringen kunnen leiden tot verschillende quantumtheorieën.

In drie dimensies zijn er verschillende modellen die min of meer equivalent zijn aan de algemene relativiteitstheorie en kunnen dienen als startpunt voor de quantumtheorie. Het eerste hoofdstuk behandelt einige voorbeelden van deze verschillende modellen en de methode van de gereduceerde fase-ruimte. Deze methode houdt in dat men expliciet de fysische vrijheidsgraden probeert te isoleren. Met andere woorden, men probeert de constraints al voor de quantizatie op te lossen. Het blijkt dat dit in drie dimensies expliciet te doen is en dat dit leidt tot slechts een eindig aantal fysische vrijheidsgraden.

Het polygoonmodel is een bepaalde formulering van drie dimensionele zwaartekracht zonder cosmologische constante maar met puntdeeltjes. Dit model is het hoofdonderwerp van dit proefschrift. De fundamentele variabelen in dit model zijn niet direkt waarneembaar maar moeten voldoen aan een aantal beperkingen zoals beschreven in het eerste deel van hoofdstuk 2. Verder wordt een beschrijving gegeven van een algorithmische oplossing voor de constraints in het geval er geen deeltjes zijn en de onderliggende ruimteachtige variëteit de topologie heeft van een Riemann oppervlak van een hoger geslacht ($g > 1$). We nemen aan dat deze procedure de volledige fysische faseruimte genereert. Het tweede deel van het hoofdstuk behandelt een mogelijke veralgemening voor het geval dat deeltjes worden toegevoegd.

Hoofdstuk 3 bevat een reductie van de eerste orde actie naar het polygoon model. De eerste stap is een reductie naar een tussenliggend, covariant model dat wordt beschreven in [32, 35]. Dit model heeft al een eindig aantal vrijheidsgraden, te weten, een aantal Lorentz vectoren die de randen van een niet planair polygoon afbakenen en de verzameling van holonomie matrices van de Lorentz groep corresponderend met gesloten niet samentrekbare krommen. De holonomieën zijn functionalen van de spin-connectie en de Lorentz vectoren zijn functionalen van het vielbein. Als we dan eisen dat het polygoon in te bedden is in $\mathbb{R}^2$ krijgen we het polygoon model. In de oorspronkelijke formulering waren de Poisson haken gepostuleerd door te eisen dat de Hamiltoniaan de goede tijdsevolutie genereert. In dit hoofdstuk worden deze echter afgeleid van de eerste orde actie van driedimensionele zwaartekracht.

Acknowledgements

I wish to thank my supervisor Gerard 't Hooft for many discussions and substantial help in improving this thesis. I am indebted to my co-supervisor, Renate Loll for motivation, collaboration, innumerable discussions and a lot of help in the last phase of writing the thesis.

I have largely profited from discussions with Dániel Nógrádi and Sergei Alexandrov, my gratitude to you and also to Balázs Szendrői, Dylan Thurston, Bernd Schroers, Mauro Carfora, Francesco Bonsante and Klaus-Dieter Semmler for invaluable help in my research. Special thanks to Willem Westra, Govind Krishnaswami, Johan Noldus, Mátyás Karádi, Bernard de Wit and Pierre van Baal for corrections of the manuscript. To all colleagues for discussions and friendship, in particular, Ivan Herger, Mario Trigiante, Emiliano Imeroni. To people I was bothering on a daily basis with practical matters, Biene, Natasja, Geertje, Wilma. To Charis for the electric plate for making coffee, for the finical like me and some of my colleagues.

During the years of my PhD research I had a great time also owing to my flatmates in Rotterdam, where I moved to, soon after I arrived in the Netherlands. My gratitude to Dani, Christina and Titusz. I also wish to thank my mother, father and brother, who always supported me with great care. Finally, I am grateful to God.

Curriculum Vitae

I was born on 26 October 1976 in Budapest, Hungary. I obtained high school education in the Budapest Lutheran Gymnasium and master's degree in theoretical particle physics in June 2000 at the Eötvös University, Budapest. My diploma thesis was about conformal field theories and the research was carried out under the supervision of Péter Bántay at the Department of Theoretical Physics. I have started to work on my PhD project at the Institute for Theoretical Physics and Spinoza Institute, Utrecht University in January 2001, under the supervision of Prof. dr. Gerard 't Hooft and dr. Renate Loll. From September 2005, I shall continue research as post-doctoral fellow at the School of Mathematical and Computer Sciences at the Henriot-Watt University, Edinburgh.