(2+1)-DIMENSIONAL GRAVITY OVER A TWO-HOLED TORUS, $T^2 \# T^2$

By

Peter R. Newbury

B.Sc.Hons., University of Manitoba, 1990

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

DEPARTMENT OF MATHEMATICS

INSTITUTE OF APPLIED MATHEMATICS

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

March, 1993

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

(Signature)

Department of Mathematics

The University of British Columbia
Vancouver, Canada

Date 12 March 1993

DE-6 (2/88)

Abstract

Research into the relationships between General Relativity, topology, and gauge theory has, for the most part, produced abstract mathematical results. This thesis is an attempt to bring these powerful theories down to the level of explicit geometric examples. Much progress has recently been made in relating Chern-Simons gauge field theory to (2+1)-dimensional gravity over topologically non-trivial surfaces. Starting from the *dreibein* formalism, we reduce the Einstein action, a functional of geometric quantities, down to a functional only of the holonomies over flat compact surfaces, subject to topological constraints. We consider the specific examples of a torus T^2, and then the two-holed torus, $T^2 \# T^2$. Previous studies of the torus are based on the fact that the torus, and only the torus, can support a continuous, non-vanishing tangent vector field. The results we produce here, however, are applicable to all higher genus surfaces. We produce geometric models for both test surfaces and explicitly write down the holonomies, transformations in the Poincaré group, ISO(2,1). The action over each surface is very nearly canonical, and we speculate on the phase space of dynamical variables. The classical result suggests the quantum mechanical version of the theory exists on curved spacetime.

Table of Contents

List of Figures

Acknowledgements

I would like to thank my supervisor, Bill Unruh, for guiding me through this project. His top-down insight met with my ground-up exploration to produce the results. I am also grateful to David Austin for reading this thesis from the point of view of a mathematician sympathetic with the pedagogical style of physics, and to Nathan Weiss for his critical reading of this thesis.

Thanks to the gang in the Audx for putting up with all my questions.

Finally, I want to thank Margaret for her unending support and confidence. Finding her was the most unexpected, but easily the greatest result of this whole endeavor.

Introduction

Using General Relativity (GR) to study of the dynamics of a spacetime generally follows a simple algorithm: Write down the action; break up the spacetime under a splitting – usually (3+1)-dimensional; re-write the action in terms of the dynamical variables on the spatial slices and constraints that govern the splitting; vary this action with respect to the dynamical variables; solve the resulting PDE's subject to the constraints for the components of the metric. In practice, even the first step of choosing an appropriate action can be daunting, let alone solving an often highly non-linear system of constrained PDE's. One approach to removing these problems is to study a different system! Rather than tackle the full 4-dimensional theory, which may describe physical space in some cases, consider the simpler case of merely 3-dimensional spacetime. This space is easily split into (2+1)-dimensions – 2 spatial and 1 temporal. While the results of these studies are clearly not physical, the techniques and problems that arise may shed light on the original (3+1)-dimensional case. fortunately, the mathematics of GR on (2+1)-dimensions is much simpler and the physics of the space is not (so) lost in mathematically difficult equations that much be examined.

With 3-dimensional spacetime sliced into (2+1)-dimensions, a system is defined by the choice of the 2 spatial dimensions. The simplest cases are those where the space is a plane or some similar infinite region. More interesting problems arise when the 2 dimensions are wrapped up into compact surfaces, especially when the surfaces are not simply (topological) spheres. In this thesis we will consider first the torus T^2 and then a genus 2 surface, the 2-holed torus $T^2 \# T^2$ (See Figure 0.1), and the interplay between the differential geometry of GR and the non-trivial topology of this surface.

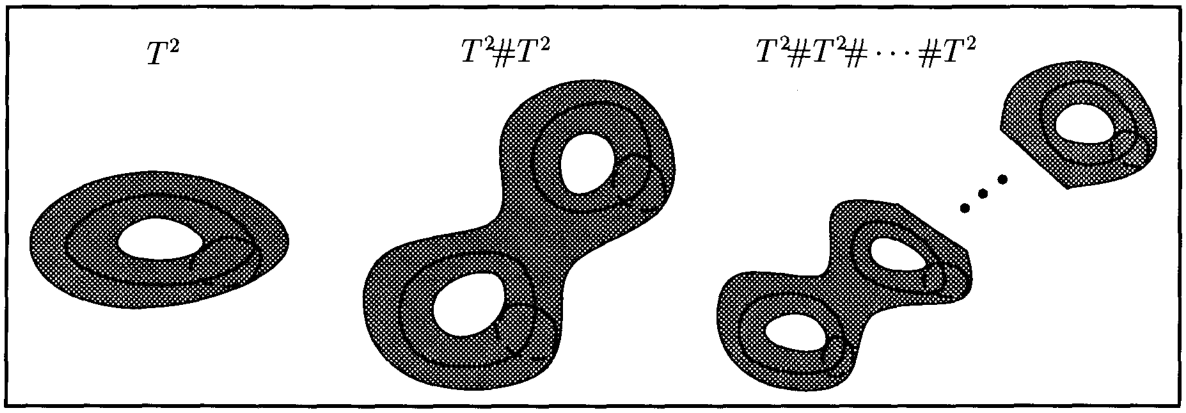

Figure 0.1: A simple torus, a 2-holed torus, and a g-holed torus, as they appear embedded in $\mathbf{R}^3$. Also shown are the non-trivial loops, 2 around each 'hole', that cannot be continuously contracted to point.

In choosing to study the simpler (2+1)-dimensional case, we have given up the possibility of producing a theory which directly describes the dynamics of the space about us. The goal, instead, is to find a mathematical result that reveals some of the subtleties of GR on non-trivial spacetimes. Nothing helps to answer questions more than a working model – where would mechanics be without the simple harmonic oscillator, or quantum mechanics with the Stern-Gerlach experiment? These idealized models answer questions about physics without the clutter of experimental error or unsolvable mathematics. Real physical systems can then be studied as perturbations of the ideal model, and generalizations to more complex models can be made from the simple ones. This is what we hope to achieve here. By constructing a simple, though unphysical, model on which we can see the roles of GR, topology, and ultimately quantum mechanics, more complex models may be built. Hopefully this will contribute to one day producing a theory of quantum gravity, one of the last stumbling blocks towards a Grand Unified Theory.

This thesis is developed as follows. We recall first, in Chapter 1, the *dreibein* formalism for GR and see how this approach, together with gauge field theory, has recently renewed interest in using (2+1)-dimensions to attempt to model quantum gravity. In Chapter 2, we reduce the Einstein action down to an (almost) canonical form. In this procedure, we see how the differential geometry drops out leaving only topological quantities: the holonomies associated with each non-trivial loop on the 2-dimensional surface. The result comes from carefully keeping the boundary terms in the action, rather than dismissing them as irrelevant, as far as the action is concerned. In Chapter 3, we construct a model for T^2. Taking this model through the prescription of Chapter 2 endows it with a frame field and connection. The resulting action depends only on the holonomies of T^2. An explicit, geometric model of the torus also shows that there are enough degrees of freedom that many different collections of holonomies are possible, some possibly admitting closed time-like curves. As a more complicated trial, Chapter 4 repeats the process over a genus 2 two-holed torus. The reduced action again depends only on the holonomies about non-trivial loops on this surface. We construct a model for $T^2 \# T^2$ by folding up an octagonal and calculate the holonomies of this particular construction. In Chapter 5, we speculate on the classical conjugate variables. The action we produce is not canonical and we make a qualitative interpretation of the phase space. While the results of Chapter 2 are classical, the form of the "conjugate variables" strongly suggests the quantum mechanical version of the model lives on curved spacetime. Finally, in Chapter 6, we see how the results may be generalized to higher genus surfaces, and where further research can be done on the link between gauge field theory and General Relativity.

This is a subject rich in both Physics, through GR and the equations of motion, and Mathematics, through topology and gauge theory. For the purpose of this thesis, we are considering (2+1)-dimensional gravity over topologically non-trivial surfaces as a Physics problem. It can equally be approached as an example of differential geometry over surfaces with non-trivial fundamental groups. The relations we encounter and choose to interpret as topological constraints imposed on the space are none other than the representation of the fundamental group of the surface in the gauge group. The particular collection of holonomies of the surface, a subset of the gauge group, obey relations which demonstrate the structure of the fundamental group. So while this thesis emphasizes the Physics interpretation of the results, the Mathematically-minded reader can translate these same results into Theorems, Proofs, and Corollaries.

Finally, a word on whom to attribute the results of this thesis. The ideas of Chapter 2 are developed as a special case of the theory of GR, and not as an interesting observation of the manipulation of the *dreibein* and connection. The difference between these two approaches is that the former requires an understanding of and experience in the mechanics of GR, gauge theory, and even index manipulation. For these reasons, Chapter 2 is a reproduction of work done by my supervisor W.G. Unruh, and to him the results should be attributed. The calculations in the rest of the thesis, the "easy part", were carried out by both of us independently, in the sense that we ended up with two stacks of paper, although I was prodded in the right direction at several stages. Therefore, this thesis should be viewed as a report of Unruh's exploration of (2+1)-dimensional gravity in the *dreibein* formalism, annotated and "demystified" to allow graduate students like myself to understand and appreciate the results.

Chapter 1

The Dreibein Formalism

The basis for studying spacetime dynamics in classical GR is the Einstein action. [1] This action is a complicated functional of the metric components $g_{\mu\nu}$. An alternative approach is to use the tetrad, or *vierbein*, formalism. In this chapter, we see how in only (2+1)-dimensions, the triad, or *dreibein*, variables greatly simplify the Einstein action. We also review recent work that recasts the *dreibein* approach to GR in a Chern-Simons gauge field theory.

1.1 The Dreibein Formalism of General Relativity

The most common approach for finding the metric on a spacetime is the Lagrangian formulation with the Einstein action

$$I = \int \sqrt{|g|} R$$

where g is the metric determinant and R, the Ricci scalar, is the twice contracted Reimannian curvature tensor. Various constants that often appear before this action, for example $\frac{1}{2}$ [1] or $\frac{1}{16\pi}$ [2], can be ignored when using variational principles. Equations of motion are found by variations of the action with respect to the metric components, $g_{\mu\nu}$. These equations of motion are generally highly non-linear PDE's, owing in part to the square root $\sqrt{|g|}$. The advantage of the *dreibein* formalism is that is supplies a kind of square root for $g_{\mu\nu}$ by using the components of the frame field as variables. The *dreibein* is

[1] This action is variously referred to as the Einstein [4], Hilbert [3], and Einstein-Hilbert [1].

the collection of three vectors $e_\mu{}^a$, where indices $\mu, \nu, \ldots = 0,1,2$ are the tangent-space components. We will use $i, j, \ldots = 1,2$ to indicate the spatial components. The indices $a, b, \ldots = 1,2,3$, sometimes known as the Lorentz indices, merely label the vectors in the frame. The metric on the spacetime is defined by[2]

$$g_{\mu\nu} = e_\mu{}^a e_{\nu a}$$

hence giving a kind of 'square root' to $g_{\mu\nu}$. The vectors must remain orthogonal in the Lorentz space:

$$e_\mu{}^a e^{\mu b} = \eta^{ab} \tag{1.1}$$

where η^{ab}=diag$\{-++\}$ is the usual Minkowski metric in (2+1)-dimensions. All raising and lowering of a, b indices is done with η^{ab}. A new connection $\omega_\mu{}^a{}_b$ is required to account for curvature in this mixed space. Covariant differentiation is defined by

$$D_\mu e_\nu{}^a = e_\nu{}^a{}_{,\mu} - \Gamma^\lambda_{\mu\nu} e_\lambda{}^a + \omega_\mu{}^a{}_b e_\nu{}^b \tag{1.2}$$

where $\Gamma^\lambda_{\mu\nu}$ are the usual Christoffel symbols. We demand that η^{ab} be constant in it's Minkowski space, which forces $\omega_\mu{}^{ab}$ to be antisymmetric:

$$\begin{aligned} 0 &= D_\mu \eta^{ab} \\ &= \eta^{ab}{}_{,\mu} + \omega_\mu{}^a{}_c \eta^{cb} + \omega_\mu{}^b{}_c \eta^{ac} \\ &= \omega_\mu{}^{ab} + \omega_\mu{}^{ba} \,. \end{aligned}$$

The spin connection $\omega_\mu{}^a = \frac{1}{2}\epsilon_{abc}\omega_\mu{}^{bc}$ is sometimes used in place of the connection, although throughout this thesis we will remain with $\omega_\mu{}^a{}_b$.

We can now re-write the Lagrangian in terms of the variables e and ω. The Ricci scaler R is found by contracting the "internal" curvature tensor $R_{\mu\nu}{}^{ab}$ with the *dreibein:*

$$R = R_{\mu\nu}{}^{\mu\nu} = e^\mu{}_a e^\nu{}_b R_{\mu\nu}{}^{ab} \,.$$

[2]Here and in all that follows, we adopt the Einstein summation convention of summing over repeated indices, both latin and greek.

While $R_{\mu\nu}{}^{\sigma\rho}$ is the curvature of spacetime, $R_{\mu\nu}{}^{ab}$ is the curvature that relates the spacetime $(\mu\nu)$ to the internal space (ab). The curvature is exchanged between the two spaces via the *dreibein*: $R_{\mu\nu}{}^{\sigma\rho} = e^{\sigma}{}_{a} e^{\rho}{}_{b} R_{\mu\nu}{}^{ab}$. A flat internal space, $R_{\mu\nu}{}^{ab} = 0$ implies a flat spacetime, $R_{\mu\nu}{}^{\sigma\rho} = 0$. In this mixed spacetime-internal space, the Lagrangian becomes

$$\sqrt{|g|}R \;=\; (\epsilon^{\mu\nu\rho} e_{\mu}{}^{a} e_{\nu}{}^{b} e_{\rho}{}^{c} \epsilon_{abc})(e^{\lambda}{}_{d} e^{\sigma}{}_{f} R_{\lambda\sigma}{}^{df}) \; .$$

By the associativity of addition, we can re-arrange the summed Lorentz indices:

$$\begin{aligned} &= \epsilon^{\mu\nu\rho} \epsilon_{adf}\, e_{\mu}{}^{a} e_{\nu}{}^{b} e_{\rho}{}^{c}\, e^{\lambda}{}_{b} e^{\sigma}{}_{c} R_{\lambda\sigma}{}^{df} \\ &= \epsilon^{\mu\nu\rho} \epsilon_{adf}\, e_{\mu}{}^{a} \delta^{\lambda}_{\nu} \delta^{\sigma}_{\rho} R_{\lambda\sigma}{}^{df} \end{aligned}$$

by the orthogonality (1.1), leaving simply

$$= \epsilon^{\mu\nu\rho} \epsilon_{abc}\, e_{\mu}{}^{a} R_{\nu\rho}{}^{bc} \; .$$

The curvature tensor can be calculated by looking at the failure of covariant derivatives to commute. Given a vector field V^a, the curvature tensor is defined by

$$[D_{\mu}, D_{\nu}]V^a = R_{\mu\nu}{}^{a}_{b} V^b \; .$$

With the definition of covariant derivative (1.2), it is simple to calculate

$$R_{\mu\nu}{}^{ab} = \omega_{\mu}{}^{ab}{}_{,\nu} + \omega_{\nu}{}^{a}{}_{d} \omega_{\mu}{}^{db} \; .$$

Finally, with volume form $\epsilon^{\mu\nu\rho} d^3x = dx^{\mu} \wedge dx^{\nu} \wedge dx^{\rho}$, we arrive at the expression

$$I = \int \epsilon_{abc} e_{\mu}{}^{a} (\omega_{\nu}{}^{bc}{}_{,\rho} + \omega_{\nu}{}^{b}{}_{d} \omega_{\rho}{}^{dc}) dx^{\mu} \wedge dx^{\nu} \wedge dx^{\rho} \tag{1.3}$$

The great advantage to this formulation of the problem is that we will nowhere encounter below a raised tangent-space index μ. This allows us to work in the more general case where $g^{\mu\nu}$ may not exist; that is, where singularities are allowed. Furthermore, on solutions to the equations of motion, the variation of the action with respect to the $e_{\mu}{}^{a}$

variables must vanish. From (1.3) we clearly see that these solutions are those connections for which $d\omega + \omega \wedge \omega$, the Ricci curvature, vanishes. Now in three dimensions, the full Reimann tensor can be written in terms of the Ricci tensor and the Ricci scalar. Hence a vanishing Ricci tensor, which immediately produces a vanishing Ricci scalar, in this particular case also implies the Riemann tensor vanishes: the solution spacetimes are truly FLAT. This is one of the main reasons why (2+1)-dimensions are studied – the geometry does not contribute to the physics of the spacetime, thereby leaving topological considerations more apparent.

In analogy with the ADM formulation [7] of (3+1)-dimensions, the action (1.3) is re-written in terms of the canonical variables and their conjugate momenta. Because of the antisymmetry of the volume form, neither $\dot{e}_o{}^a$ nor $\dot{\omega}_o{}^a{}_b$ appears in the action, where $\dot{(\ )}$ indicates differentiation with respect the time coordinate of the spacetime. Without conjugate momenta, these variables are constant in time. In the language of variational calculus, $e_o{}^a$ and $\omega_o{}^a{}_b$ are Lagrange multipliers of constraints which govern the way the 2-dimensional slices of 'space' evolve in 'time'. Explicitly, we find

$$\begin{aligned} I &= -2\int dt \int d^2x\, \epsilon^{ij} e_i{}^a \dot{\omega}_{ja} \\ &+ \int dt \int d^2x \epsilon_{abc}\{\, e_o{}^a \epsilon^{ij}(\omega_i{}^{bc}{}_{,j} - \omega_j{}^{bc}{}_{,i} + \omega_j{}^b{}_d \omega_i{}^{dc} - \omega_i{}^b{}_d \omega_j{}^{dc}) \\ &\qquad + \omega_o{}^{bc} \epsilon^{ij}(e_i{}^a{}_{,j} - e_j{}^a{}_{,i} + e_i{}^d \omega_j{}^a{}_d - e_j{}^d \omega_i{}^a{}_d)\}\ . \end{aligned}$$

The constraints, proportional to $e_o{}^a$ and $\omega_o{}^{bc}$, are recognized as

$$R_{ij}{}^{bc} = 0 \tag{1.4}$$

$$D_{[i} e_{j]}^a = 0\ , \tag{1.5}$$

respectively. These two constraints, which tell us the spatial slices are flat and torsion free, greatly simplify the system, as we will see in chapter 2.

We see that of all the ways to attempt to describe the geometry of 3-dimensional spacetime, the result produced by the theory of GR is fairly simple. In fact, most of the gymnastics of differential geometry have disappeared! It is no wonder, then, that this "simple" result can be reproduced from a quite different abstract approach, Chern-Simons gauge field theory.

1.2 Chern-Simons Gauge Field Theory over the Poincaré Group

In recent work that rekindled interest in (2+1)-dimensional GR, Witten [1] recognized the Einstein action written as (1.3) as a Chern-Simons action of the the gauge field theory for the Poincaré group. This gauge group is the collection of all Poincaré transformations, consisting of a Lorentz transformation followed by a spacetime translation:

$$V^a \to \mathcal{U}^a_{\ b} V^b + T^a \ . \tag{1.6}$$

The Lorentz subgroup of the Poincaré group, more often associated with Special Relativity, has disconnected components corresponding to proper, orthochronous transformations of the connected component of the identity, and components connected to parity-, time-, and total-inversion. The Lorentz transformations we will deal with here are restricted to the first of these components, the only component which represents physical transformations. An ISO(2,1) representation (*I*nhomogeneous *S*pecial *O*rthogonal) of the Poincaré group is the collection of 4x4 matrices

$$[\mathcal{U}|T] = \left(\begin{array}{c|c} \mathcal{U}^a_{\ b} & T^a \\ \hline 0 & 1 \end{array}\right)$$

where $\mathcal{U}^a_{\ b}$ is an SO(2,1) Lorentz transformation. Properties of these matrices $\mathcal{U}^a_{\ b}$ that will be used below are found in the Appendix A. Transformations of 3-dimensional vectors

V^a are carried out by appending a fourth component, of fixed length 1, to the 3-vectors forming $V^A = (V^a, 1)$:

$$\left(\begin{array}{c|c} \mathcal{U}^a{}_b & T^a \\ \hline 0 & 1 \end{array}\right)\left(\begin{array}{c} V^c \\ \hline 1 \end{array}\right) = \left(\begin{array}{c} \mathcal{U}^a{}_b V^b + T^a \\ \hline 1 \end{array}\right).$$

The 3-dimensional component of these resulting 4-vectors reproduces the 3-dimensional Poincaré transformation on V^a.

The Poincaré group has generators

$$J_1 = \left(\begin{array}{ccc|c} 0 & 0 & 0 & \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & \\ \hline & 0 & & 0 \end{array}\right) \quad J_2 = \left(\begin{array}{ccc|c} 0 & 0 & -1 & \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & \\ \hline & 0 & & 0 \end{array}\right) \quad J_3 = \left(\begin{array}{ccc|c} 0 & 1 & 0 & \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \\ \hline & 0 & & 0 \end{array}\right).$$

Under exponentiation, these generate, in Minkowski space where $x^a = \{t, x, y\}$, rotations in the xy-plane, boosts in the $(-y)$-direction, and boosts in the x-direction, respectively. While these 4×4 matrices generate the Lorentz subgroup of ISO(2,1), the upper left 3×3 submatrices by themselves generate the pure Lorentz group SO(2,1). The generators of translation in the t-, x-, and y-directions are, respectively,

$$P_1 = \left(\begin{array}{c|c} & 1 \\ 0 & 0 \\ & 0 \\ \hline 0 & 0 \end{array}\right) \quad P_2 = \left(\begin{array}{c|c} & 0 \\ 0 & 1 \\ & 0 \\ \hline 0 & 0 \end{array}\right) \quad P_3 = \left(\begin{array}{c|c} & 0 \\ 0 & 0 \\ & 1 \\ \hline 0 & 0 \end{array}\right).$$

The generators of this group obey the Lie algebra

$$[J_a, J_b] = \epsilon_{abc} J^c$$

$$[J_a, P_b] = \epsilon_{abc} P^c$$

$$[P_a, P_b] = 0 \ .$$

In [1], Witten observes that in taking the 1-form $A = A_\mu dx^\mu$ with components

$$A_\mu = e_\mu{}^a P_a + \omega_\mu{}^a J_a$$

as a gauge field, the Chern-Simons action

$$I_{CS} = \tfrac{1}{2} \int_M \mathbf{Tr}(A \wedge dA + \tfrac{2}{3} A \wedge A \wedge A)$$

exactly coincides with the Einstein action (1.3). We see that ISO(2,1) is the gauge group by varying the gauge field A_μ. The infinitesimal transformation of A_μ generated by a parameter $u = \rho^a P_a + \tau^a J_a$ with infinitesimal ρ^a, τ^a is defined as

$$\delta A_\mu = u_{,\mu} + [A_\mu, u] \ .$$

This variation is the covariant derivative of the field with respect to the connection A_μ. Varying the field in this way produces

$$\delta A_\mu = \delta e_\mu{}^a P_a + \delta \omega_\mu{}^a J_a$$

where

$$\begin{aligned} \delta e_\mu{}^a &= -\rho^a_{,\mu} - \epsilon^{abc} e_{\mu b} \tau_c - \epsilon^{abc} \omega_{\mu b} \rho_c \ , \\ \delta \omega_\mu{}^a &= -\tau^a_{,\mu} - \epsilon^{abc} \omega_{\mu b} \tau_c \ . \end{aligned}$$

Witten notes that by setting $\rho^a = V^\mu e^a_\mu$ for an infinitesimal spacetime vector V^μ, the difference between these transformations of e^a_μ and $\omega_\mu{}^a$, and those generated by infinitesimal Lorentz transformations along with infinitesimal diffeomorphisms (translations!) in the $-V^\mu$ direction, is simply a Lorentz transformation. Since Lorentz transformations are part of the gauge group, any gauge invariant quantity, like the action, will be unaffected by this difference. That is, the Chern-Simons action in the gauge field A_μ over the Poincaré group is the same as the Einstein action of GR. Gravity can be re-expressed as a gauge field theory, which greatly increases the chances of finding a quantum mechanical version of the theory. Though this thesis is based on classical GR, and not quantum or gauge field theory, we will see the importance of the Poincaré group in the results that follow.

Chapter 2

The Einstein Action

In this chapter we derive the main result of this thesis: the Einstein action over a compact surface can be written explicitly in terms of the holonomies on the surface. Before we produce this result, let us briefly review holonomies and boundary terms.

Intuitively, a holonomy is the failure of a vector to return to its original orientation after being parallel transported around a closed path, or loop. We will define holonomy as the transformation which carries the initial vector onto the final. While in general a holonomy is an automorphism of the tensor fields over a manifold, the holonomies encountered here are simple transformations on the tangent space. Curvature, also a measure of the changes in parallel transported vectors, is closely related to the holonomy of trivial, or contractible, loops. When a surface is flat, the holonomy of all trivial loops is simply the identity — no change occurs in a vector when it is parallel transported about a flat surface. When curvature is present, the holonomy depends both on the base point of the loops and the shape of the loops itself: parallel transport around "long and twisting" loops will alter a vector more than parallel transport around "small and simple" loops.

As we saw in Chapter 1, the surfaces we deal with here are flat, and the presence of interesting holonomies seems unlikely. However, the surfaces we deal with are also topologically non-trivial, and are covered in non-contractible loops, those paths which get "caught" on one or more holes formed by the surface (See Figure 0.1). The holonomy over these non-trivial loops may not be the identity, even though the surface is everywhere flat. The topological degrees of freedom can generate non-trivial holonomies. Assuming hereon that "surface" means a flat compact surface with interesting topology, we recall that all homotopic curves have the same holonomy. Two curves which can be smoothly deformed into each other differ by a contractible loop (See Figure 2.1). This path does not contribute to the holonomy as the surface is flat, so the holonomy of the two homotopic paths is the same. Therefore all holonomies on a surface will be known once those about a few representative non-trivial loops are known. Just as the genus 1 torus has 2 distinct classes of incontractible loops, a genus g surface, like a g-holed torus, has $2g$ generating loops. By finding, or specifying, the holonomies about each of these $2g$ loops, the geometry of the surface is completely determined, at least up to global gauge transformations. These are the transformations which transform the entire surface while leaving the action (or any other gauge-invariant function) invariant. It is these $2g$ holonomies which will play the role of the dynamical variables in the expression for the action we derive from the Einstein action (1.3) in this chapter.

Connected-sum surfaces, like those in Figure 0.1 have no boundary. How then, do we handle boundary terms that arise in the action? The answer comes by looking at the way these surfaces are constructed. Generally a genus g surface is constructed from a $4g$-sided polygon with pairs of sides identified, folding up the polygon, creating the surface. The matched points on the boundary of the polygon come together to form a seamless, boundary-less, surface. Another way to view this construction is to tile some infinite plane with these polygons, so that moving off one tile onto an adjacent one is the same as travelling off a single polygon and reappearing at the identified point on the boundary. The boundary can be re-inserted into the closed, compact surface by slicing the surface open. This does not affect integration over the surface: A continuous function has the same value at two identified points on the boundary. Since the outward-pointing normals from the two identified boundary points have opposite orientations, any contribution to

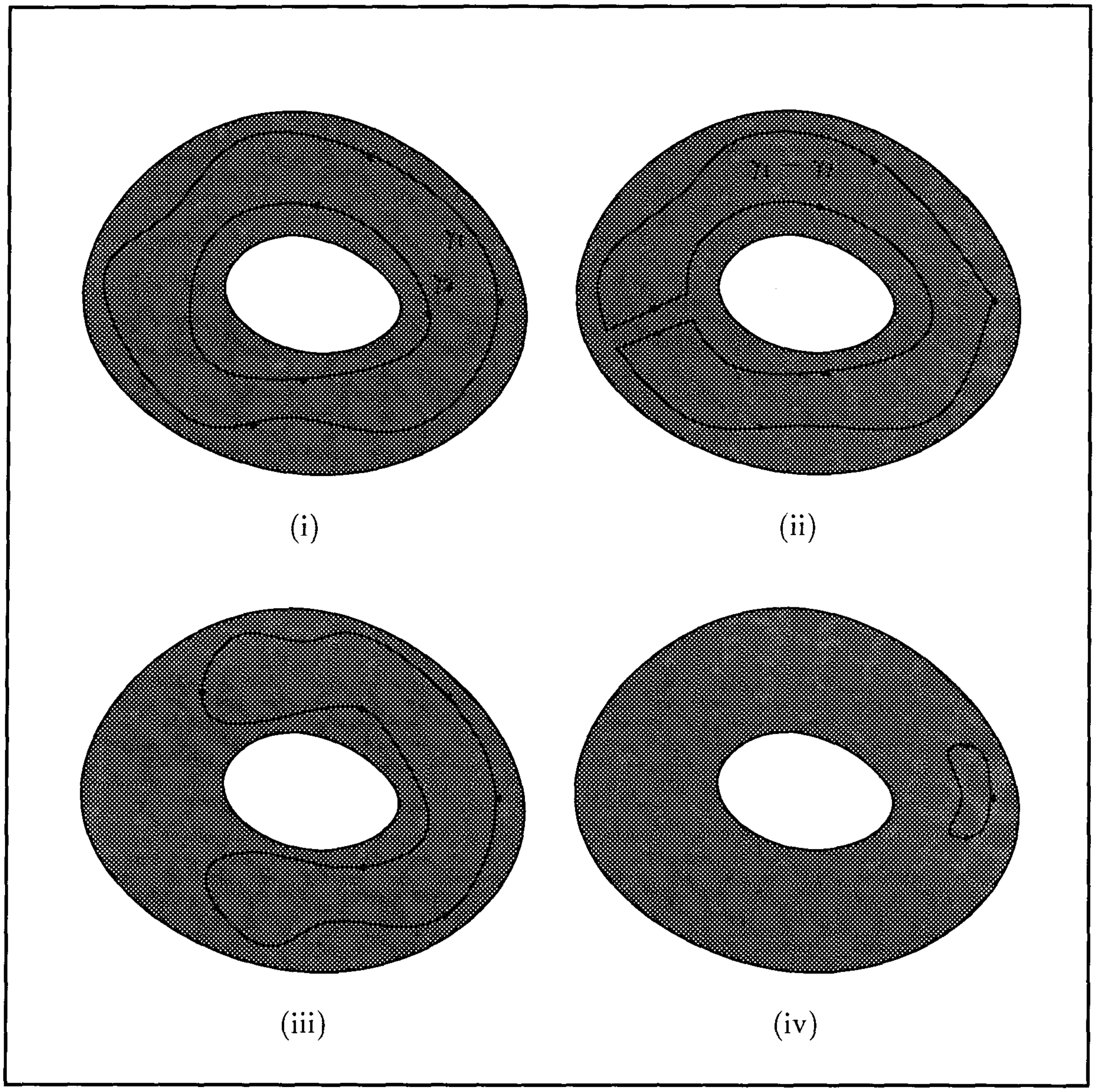

Figure 2.1: The difference $\gamma_1 - \gamma_2$ between two homotopic curves γ_1 and γ_2 is contractible to a point. The holonomy of this trivial loop is the identity, so γ_1 and γ_2 have the same holonomy.

a boundary integral along one boundary component is exactly cancelled by the matching boundary's contribution. The converse to this situation need not by true, though, as we will see below. By starting with the tile and matching pairs of sides, we are no longer guaranteed that functions are continuous across the boundary. The discontinuities we allow are exactly the holonomies that appear in the action.

2.1 Reduction of the Einstein Action

The dynamical variables that appear in the Einstein action (1.3) are the *dreibein* $e_\mu{}^a$ and the connection $\omega_\mu{}^a{}_b$. Let us assume there exists a collection of $e_\mu{}^a(x^\nu)$ and $\omega_\mu{}^a{}_b(x^\nu)$ which are continuous over the surface, or equivalently, continuous on the tile, even across the identified boundaries. Because only the torus T^2 can support a non vanishing tangent vector field [6], we must allow the possibility of the *dreibein* becoming singular on the tile.

First consider the symmetries of the action. The action, as the Correspondence Principle implies, is Lorentz invariant. Transform the variables under

$$\begin{aligned} \tilde{e}_\mu{}^a &= \mathcal{U}^a_{\ b} e_\mu{}^b \\ \omega_\mu{}^a{}_b &= \tilde{\omega}_\mu{}^c{}_d \mathcal{U}_c{}^a \mathcal{U}^d_{\ b} + \mathcal{U}_c{}^a (\mathcal{U}^c_{\ b,\mu}) \ . \end{aligned} \tag{2.1}$$

Some index manipulation along with the properties of $\mathcal{U}$ (Appendix A) show the action is the same functional of $\tilde{e}$ and $\tilde{\omega}$. Recall from the (2+1)-dimensional splitting of spacetime that the space is subject to the constraint $R_{ij}{}^{ab} = 0$. This implies that on any coordinate patch, we can find a particular Lorentz transformation $\mathcal{U}^a_{\ b}(x^\nu)$ such that $\tilde{\omega}_i{}^a_{\ b} = 0$ (See Appendix B). Note that only the spatial components of $\tilde{\omega}$ can be made to vanish because only the spatial $R_{ij}{}^{ab} = 0$.

In this new coordinate system, consider a transformation in the internal space generated by a function $\rho^a(x^\nu)$:

$$\tilde{e}_\mu{}^a = \hat{e}_\mu{}^a + \tilde{D}_\mu \rho^a \ , \tag{2.2}$$

where $\tilde{D}$ is the covariant derivative with respect to the connection $\tilde{\omega}$. While is is assumed that $e_\mu{}^a$ is continuous across the boundary, it is not necessarily true that $\hat{e}_\mu{}^a$ is continuous. With this transformation, the action becomes

$$I = \int_M \epsilon_{abc} \, (\hat{e}_\mu{}^a + \tilde{D}_\mu \rho^a) \tilde{R}_{\nu\rho}{}^{bc} \, dx^\mu \wedge dx^\nu \wedge dx^\rho$$

where M is the (2+1)-dimensional manifold and $\tilde{R}$ is the curvature, also written in terms of $\tilde{\omega}$. An integration-by-parts on the second term gives

$$\begin{aligned} I &= \int_M \epsilon_{abc} \, \hat{e}_\mu{}^a \tilde{R}_{\nu\rho}{}^{bc} \, dx^\mu \wedge dx^\nu \wedge dx^\rho \\ &\quad + \int_{\partial M} \epsilon_{abc} \, \rho^a \tilde{R}_{\mu\nu}{}^{bc} \, dx^\mu \wedge dx^\nu \\ &\quad\quad - \int_M \epsilon_{abc} \, \rho^a \tilde{D}_{[\mu} \tilde{R}_{\nu\rho]}{}^{bc} dx^\mu \wedge dx^\nu \wedge dx^\rho \ . \end{aligned}$$

The last term vanishes by the Bianchi Identity. A second constraint imposed on the system by the (2+1)-dimensional splitting is $D_{[i} e_{j]}^a = 0$ in the original coordinate system. This constraint implies we can find a $\rho^a(x^\nu)$ such that the spatial $\hat{e}_i{}^a = 0$ (See Appendix B). The first term of the above action vanishes leaving only the boundary integral. The boundary here includes the spatial boundary where the identification takes place and also the initial and final hypersurfaces in time. The integrals over these temporal boundaries, however, come from the term

$$\int_M \epsilon_{abc} \, \tilde{D}_o (\rho^a \tilde{R}_{\nu\rho}{}^{bc}) dx^o \wedge dx^\nu \wedge dx^\rho \ .$$

This total time derivative can be removed form the Lagrangian as it adds only a constant to the action and has no effect on the equations of motion found by variational principles.

This leaves only

$$I = \int_{\partial} \epsilon_{abc}\, \rho^a (\tilde{\omega}_{o,i}^{\ bc} - \dot{\tilde{\omega}}_i^{\ bc} + \tilde{\omega}_i{}^b{}_d \tilde{\omega}_o{}^{dc} - \tilde{\omega}_o{}^b{}_d \tilde{\omega}_i{}^{dc}) dx^i$$

where ∂ represents the spatial boundary. As the spatial components $\tilde{\omega}_i{}^b{}_a$ vanish, the action is further reduced to

$$I = \int_{\partial} \epsilon_{abc}\, \rho^a \tilde{\omega}_{o,i}^{\ bc} dx^i \ .$$

One more integration-by-parts gives

$$I = - \int_{\partial} \epsilon_{abc}\, \rho^a_{,i}\, \tilde{\omega}_o{}^{bc} dx^i \ .$$

The boundary term of this integration vanishes, being the boundary of a boundary. Now, each side of the $4g$-sided tile is attached to another tile by the identifications used to get the topology. Thus each distinct edge of the tile contributes twice to this integral. Call the two matching sides '1' and '2'. Summing over the $2g$ different edges gives

$$I = \sum_{\text{edges}} \int_{\text{edge}} \epsilon_{abc} \left(\rho_1{}^a{}_{,i}\, \tilde{\omega}_{1o}{}^{bc} - \rho_2{}^a{}_{,i}\, \tilde{\omega}_{2o}{}^{bc} \right) dx^i \ .$$

Recall from the transformation (2.1)

$$\tilde{\omega}_o{}^{cd} = \omega_o{}^{ab}\, \mathcal{U}^c{}_a \mathcal{U}^d_{\ b} - \dot{\mathcal{U}}^c{}_b \mathcal{U}^{db} \ . \tag{2.3}$$

Just consider the first term, in ω_o. As ω_o is assumed to be continuous across the boundary, $\omega_{1o}{}^{bc} = \omega_{2o}{}^{bc}$; we can write just $\omega_o{}^{bc}$. This first term of the action is

$$\sum_{\text{edges}} \int_{\text{edge}} \epsilon_{abc} \left(\rho_1{}^a{}_{,i}\, \mathcal{U}_1{}^b{}_d \mathcal{U}_1{}^c{}_e - \rho_2{}^a{}_{,i}\, \mathcal{U}_1{}^b{}_d \mathcal{U}_1{}^c{}_e \right) \omega_o{}^{de}\, dx^i \ .$$

As $\mathcal{U}_1$ and $\mathcal{U}_2$ are both Lorentz transformations, there is a Lorentz transformation between them. Define

$$\mathcal{U}_1{}^a{}_b = W^a{}_c \mathcal{U}_2{}^c{}_b \tag{2.4}$$

This maps $\mathcal{U}$ on one side of the boundary onto $\mathcal{U}$ on the other side, in some sense carrying $\mathcal{U}$ across the boundary. Observe that

$$\begin{aligned}
\mathcal{U}_{1c}{}^{a}\,\mathcal{U}_{1}{}^{c}{}_{b,i} &= (W_c{}^{a}\mathcal{U}_{2d}{}^{a})(W^c{}_e\mathcal{U}_{2}{}^{e}{}_{b})_{,i} \\
&= \underbrace{(W_c{}^{d}W^c{}_e)}_{\delta^d_e}\mathcal{U}_{2d}{}^{a}\mathcal{U}_{2}{}^{e}{}_{b,i} \\
&\qquad + W^d_c\,\mathcal{U}_{2d}{}^{a}\,W^c{}_{e,i}\,\mathcal{U}_{2}{}^{e}{}_{b}
\end{aligned}$$

With $\tilde{\omega}_i{}^a{}_b$ vanishing in (2.1) and $\omega_{1o}{}^{bc} = \omega_{2o}{}^{bc}$,

$$\mathcal{U}_{1d}{}^{a}\,\mathcal{U}_{1}{}^{d}{}_{b,i} = \mathcal{U}_{2d}{}^{a}\,\mathcal{U}_{2}{}^{d}{}_{b,i} \ .$$

Hence

$$W_c{}^{d}\,\mathcal{U}_{2d}{}^{a}\,W^c{}_{e,i}\,\mathcal{U}_{2}{}^{e}{}_{b} = 0 \ ,$$

or

$$W^c{}_{e,i} = 0 \ : \tag{2.5}$$

the W transition matrices are constant on each time slice. Substitute into the action for $\mathcal{U}_2$ in terms of W and $\mathcal{U}_1$. Property (A.4) then reduces this term of the action to simply

$$\sum_{\text{edges}} \int_{\text{edge}} (\rho_{1}{}^{a}_{,i} - W^a{}_b\,\rho_{2}{}^{a}_{,i})\,\mathcal{U}_{1}{}^{b}{}_{d}\,\mathcal{U}_{1}{}^{c}{}_{e}\,\omega_o{}^{de}\,dx^i \ . \tag{2.6}$$

Before we calculate the second part of the action from the $\dot{\mathcal{U}}\mathcal{U}$ term in (2.3), not the following. Recall $\dot{\omega}_o{}^{de}$ does not appear in the action by the antisymmetry of the volume form. Thus $\omega_o{}^{de}$ is a Lagrange multiplier. The $\dot{\mathcal{U}}\mathcal{U}$ term contains no $\omega_o{}^{de}$, so the constraint associated with this Lagrange multiplier comes from the first term of the action alone, namely

$$\rho_{1}{}^{a}_{,i} - W^a{}_b\,\rho_{2}{}^{b}_{,i} = 0 \ . \tag{2.7}$$

We found in (2.5) that $W^a{}_b$ is constant on each time slice, so this constraint is equivalent to

$$W^a{}_b\rho_2{}^b + \Pi^a = \rho_1{}^a \tag{2.8}$$

where Π^a is constant. Under the identification of points that generates the topology of the surface, ρ_1 and ρ_2 are two vectors sitting at the same point (in the same tangent space). The relation (2.8) shows these two vectors are related by a Poincaré transformation (see 1.6)). Another way to compare vector ρ_1 with vector ρ_2 is to parallel transport ρ_1 across the tile between the two identified points on the boundary. The resulting vector is defined to be ρ_2. Now (2.8) shows the *holonomy* of this loop in a Poincaré transformation. We have seen that the W are constant on each time-slice and that the holonomies are path independent. Thus ρ on the whole of side 1 maps onto ρ on the whole of side 2 under this Poincaré holonomy.

Now consider the remaining term in the action,

$$\sum_{\text{edges}} \int_{\text{edge}} \epsilon_{abc} \, (\rho_1{}^a_{,i} \dot{\mathcal{U}}_1{}^b{}_d \mathcal{U}_1{}^{cd} - \rho_2{}^a_{,i} \dot{\mathcal{U}}_2{}^b{}_d \mathcal{U}_2{}^{cd}) \, dx^i \ .$$

From (2.7) and (2.4) we substitute

$$\rho_2{}^a{}_{,i} = W_b{}^a \rho_1{}^b{}_{,i} \ ,$$

$$\mathcal{U}_2{}^{cd} = W_y{}^c \mathcal{U}_1{}^{yd} \ .$$

Because $\dot{W}^a_b$ does not necessarily vanish (only the spatial derivatives do), we cannot yet substitute for $\dot{\mathcal{U}}_2{}^b{}_d$. These substitutions give

$$\begin{aligned} \sum_{\text{edges}} \int_{\text{edge}} \epsilon_{abc} \, (\rho_1{}^a_{,i} \dot{\mathcal{U}}_1{}^b{}_d \mathcal{U}_1{}^{cd} & \\ - W^a_e \, \rho_1{}^e{}_{,i} \dot{\mathcal{U}}_2{}^b{}_d \, W_y{}^c \mathcal{U}_1{}^{yd}) \, dx^i \ . \end{aligned}$$

Replace ϵ_{abc} under (A.4) in the second term. Then (A.1) and (A.3) reduce the action to

$$\sum_{\text{edges}} \int_{\text{edge}} \epsilon_{abc} \, \rho_1{}^a{}_{,i} (\dot{\mathcal{U}}_1{}^b{}_d - W^b{}_x \dot{\mathcal{U}}_2{}^x{}_d) \mathcal{U}_1{}^{cd} \, dx^i \ .$$

Now

$$\begin{aligned} W^b{}_x \dot{\mathcal{U}}_2{}^x{}_d &= \dot{\overline{W^b{}_x \mathcal{U}_2{}^x{}_d}} - \dot{W}^b{}_x \mathcal{U}_2{}^x{}_d \\ &= \dot{\mathcal{U}}_1{}^b{}_d - \dot{W}^b{}_x \, W_y{}^x \mathcal{U}_1{}^y{}_d \end{aligned}$$

giving

$$\sum_{\text{edges}} \int_{\text{edge}} \epsilon_{abc}\, {\rho_1}^a{}_{,i}\, \dot{W}^b{}_x\, W^{cx}\, dx^i \ .$$

Recall the W^a_b are constant and hence can be pulled out of this integral along the edges of the tile, leaving only

$$\cdots \int_{\text{edge}} {\rho_1}^a{}_{,i}\, dx^i \ .$$

This integral, simply the difference in ρ_1 between the ends of edge 1 of the tile, we denote by $\Delta{\rho_1}^a$.

Finally, we produce the action on the tile. By imposing the constraint (2.7), which introduces the constant Π^a, the first term of the action (2.6) vanishes. All that remains for the action is the sum over representative edges

$$I = \sum_{\text{edges}} \Delta{\rho_1}^a \dot{W}^b{}_x\, W^{cx}\, \epsilon_{abc} \tag{2.9}$$

By constructing a tile and assigning on the tile fields of W and Π, we can explicitly calculate this action for a torus T^2 and later, the more complicated but interesting two-holed torus $T^2 \# T^2$.

Chapter 3

The Simple Torus

An elegant, but not particularly profound result of differential geometry is the observation, here attributed to Carlip[1], that

> "a flat connection is determined, uniquely up to gauge transformations, by its holonomies around the nontrivial loops..."

Mathematically, this is a concrete and definitive corollary, distilled from a much larger theory. It is not an explicit statement of the physics of the system, though, for it still deals with abstract ISO(2,1) transformations across some surface or region with identified points. In this Chapter, we see how the result (2.9) of Chapter 2, which asserts that the action in a function(al) of the holonomies (supporting Carlip's statement), is manifest on a simple 1-holed torus T^2. This is relatively easy to do, as the flat genus $g = 1$ torus has only two independent holonomies, and can be constructed by identifying pairs of sides of a square in the Euclidean plane $\mathbf{R}^2$.

3.1 The Action over a Torus

Each surface of constant time, the 2-dimensional spatial slices of (2+1)-dimensional spacetime, is tiled with squares, or more generally parallelograms. To calculate the action on this tile, let us first label its components as in Figure 3.1(i). We call the sides A, B, C, D and the corners 1,2,3,4. The identification of sides $A\&C$ and $B\&D$ is indicated by the

[1] [4], p. 2649.

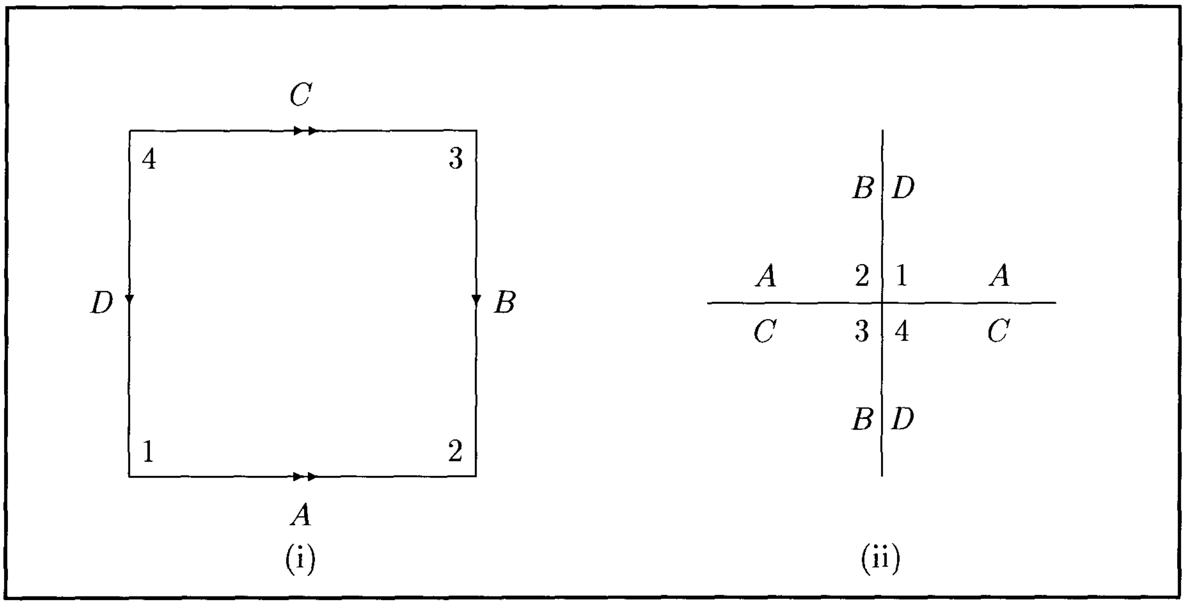

Figure 3.1: (i) Labels for the sides and corners of the tile, and identifications of sides A&C and B&D. (ii) The neighborhood of a vertex where four tiles meet.

arrows on the boundary. Folding up the tile by gluing together the identified sides with the indicated orientation creates a torus, T^2. The same gluing information can also be exhibited by looking at the neighborhood of the vertex where the 4 corners come together, or where 4 tiles meet on the plane (See Figure 3.1(ii)). We produce this Figure by "bootstrapping" around the vertex, a method we will employ frequently in what follows. Starting form the region about the vertex labelled 1, sides A and C are identified as part of the topology generating gluing. Adjacent to side C is side D, and these two sides meet at corner 4. Next, side B is glued onto D, side C is adjacent to B, and corner 3 lies at the intersection of these two boundary components. We continue in this way, labelling each side and each region of the vertex, showing all the gluing of Figure 3.1(i). Each line, like the one between A and C, is one of the boundaries we say has $\rho_1, \tilde{\omega}_1$ on one side, and $\rho_2, \tilde{\omega}_2$ on the other. The corresponding edge along which we calculate $\Delta{\rho_1}^a$ is, in this case, the line that runs out from the 1-4 region of the vertex and returns to

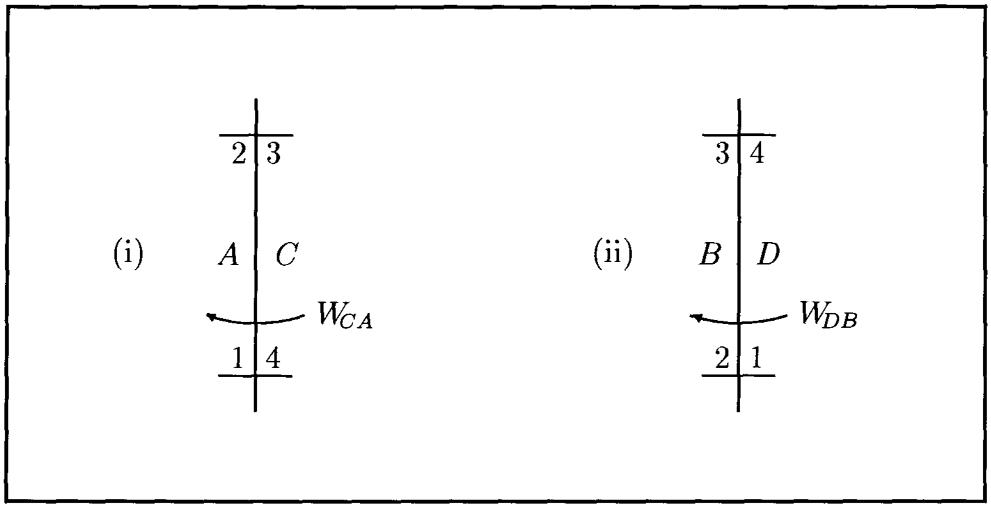

Figure 3.2: The transformations W across the identified edges (i)$C\&A$, (ii)$D\&B$.

the vertex though the 2-3 region. Across the boundary between sides A and C we find from (2.8)

$$W_{CA}{}^a{}_b\, \rho_C{}^b + \Pi_{CA}{}^a = \rho_A{}^a \ .$$

Here W_{CA} is the Lorentz transformation which relates $W_{CA}\mathcal{U}_C = \mathcal{U}_A$ and Π_{CA} is the constant defined by (2.7). This relation holds all along the edge between sides A and C so we push the result down to the 1-4 region of the vertex. Evaluated at the vertex we write

$$W_{CA}{}^a{}_b\, \rho_C{}^b(4) + \Pi_{CA}{}^a = \rho_A{}^a(1) \ .$$

The term $\rho_C{}^a(1)$ is really the limiting value of $\rho_C{}^a$ as the vertex is approached along edge C. There is a similar relation at the other end of this CA boundary, the end in the 3-2 region of the vertex. The relations across each of the four edges of the tile can be written down by looking at Figure 3.2:

$$
\begin{array}{ll}
\text{Figure 3.2 (i)} & \left\{ \begin{array}{l} W_{CA}{}^{a}{}_{b}\, \rho_C{}^{b}(4) + \Pi_{CA}{}^{a} = \rho_A{}^{a}(1) \\ W_{CA}{}^{a}{}_{b}\, \rho_C{}^{b}(3) + \Pi_{CA}{}^{a} = \rho_A{}^{a}(2) \end{array} \right. \\
\text{Figure 3.2 (ii)} & \left\{ \begin{array}{l} W_{DB}{}^{a}{}_{b}\, \rho_D{}^{b}(4) + \Pi_{DB}{}^{a} = \rho_B{}^{a}(3) \\ W_{DB}{}^{a}{}_{b}\, \rho_D{}^{b}(1) + \Pi_{DB}{}^{a} = \rho_B{}^{a}(2) \end{array} \right.
\end{array}
\tag{3.1}
$$

To produce the action we must evaluate the $\Delta\rho^a$ by integrating along the identified edges. Each edge contributes twice, recall, but the return integral along the 2-side is taken care of with the $\dot{W}W$ terms in the action. We need only consider the 1-sides, from corners $1 \to 2$ and $2 \to 3$. The action is simply

$$
\begin{aligned}
I \;=\; & \Big\{ (\rho_A{}^a(2) - \rho_A{}^a(1))\, \dot{W}_{CA}{}^{b}{}_{x} W_{CA}{}^{cx} \\
& + (\rho_B{}^a(3) - \rho_B{}^a(2))\, \dot{W}_{DB}{}^{b}{}_{x} W_{DB}{}^{cx} \Big\}\, \epsilon_{abc}\ .
\end{aligned}
$$

Now it is not true that, say, $\rho_C(4) = \rho_A(1)$, even though these two functions are evaluated at the same point, under the identification of sides A and C. There is the discontinuity is ρ^a defining W_{CA} and Π_{CA}. It is true, however, that *at vertex* 4, $\rho_D(4) = \rho_C(4)$, as these are evaluated at the same point on the tile, without any identification of points required. Call the corner value just ρ_4. Similar relations hold in each region of the vertex:

$$
\begin{array}{ll}
\rho_A(1) = \rho_D(1) = \rho_1 & \rho_B(2) = \rho_A(2) = \rho_2 \\
\rho_C(3) = \rho_B(3) = \rho_3 & \rho_D(4) = \rho_C(4) = \rho_4
\end{array}
\tag{3.2}
$$

With these relations, along with the W and Π above, we can express ρ in each region of the vertex in terms of, say, ρ_1. By bootstrapping around the vertex from region 1, jumping across boundaries with (3.1) and around corners with (3.2), we evaluate $\rho_3(\rho_1)$, for instance, in vector- rather than component-form, as

$$
\begin{aligned}
\rho_3 &= W_{CA}^{-1}(\rho_2 - \Pi_{CA}) \\
&= W_{CA}^{-1}(W_{DB}\rho_1 + \Pi_{DB} - \Pi_{CA})\ .
\end{aligned}
$$

Note that just as we define W_{CA} to jump from side C to side A, W_{AC} jumps from side A to side C. Each of the W is invertible, though, so $W_{AC} = W_{CA}^{-1}$. In this way we find

$$\left.\begin{array}{rcl} \rho_2 &=& W_{DB}(\rho_1 + \Pi_{DB}) \\ \rho_3 &=& W_{CA}^{-1}(W_{DB}\rho_1 + \Pi_{DB} - \Pi_{CA}) \\ \rho_4 &=& W_{CA}^{-1}(\rho_1 - \Pi_{CA}) \end{array}\right\} \tag{3.3}$$

We arbitrarily chose to write $\rho_3 = \rho_3(\rho_2(\rho_1))$, but we equally could have chosen $\rho_3(\rho_4(\rho_1))$, or even $\rho_3(\rho_4(\rho_1(\rho_2(\rho_3(\rho_4(\rho_1(\cdots))$. In order that the results be consistent, it must be true that the transformation giving a complete circuit of the vertex is the Identity:

$$\rho_1(\rho_4(\rho_3(\rho_2(\rho_1)))) = \rho_1 \ .$$

As the transformation is in general a Poincaré transformation, this means the Lorentz part is the Identity, while the translation vanishes. Jumping all the way around the vertex produces

$$\begin{array}{rcl} \rho_1 &=& W_{CA}W_{DB}^{-1}W_{CA}^{-1}W_{DB}\rho_1 \\ && + \ W_{CA}W_{DB}^{-1}W_{CA}^{-1}\Pi_{DB} \\ && - \ W_{CA}W_{DB}^{-1}W_{CA}^{-1}\Pi_{CA} \\ && - \ W_{CA}W_{DB}^{-1}\Pi_{DB} \\ && + \ \Pi_{CA} \ . \end{array}$$

Hence the Lorentz transformations must obey the closure relation

$$W_{CA}W_{DB}^{-1}W_{CA}^{-1}W_{DB} = \mathbf{1} \ . \tag{3.4}$$

This "constraint" is the representation of the fundamental group $\pi_1(T^2)$ in the gauge group ISO(2,1). It is equivalent to the condition

$$W_{CA}W_{DB} = W_{DB}W_{CA}$$

so that the two Lorentz transformations commute. Applying the relation (3.4), we reduce the translation constraint to simply

$$(\mathbf{1} - W_{CA})\,\Pi_{DB} = (\mathbf{1} - W_{DB})\,\Pi_{CA}\ . \tag{3.5}$$

We will further analyse this constraint below, and how much, or how little, it further constrains the system. First though, we calculate the action.

With the relations (3.2) and (3.3), we can express the action in terms of ρ_1 (or simply ρ), W, and Π, plus the constraints (3.4) and (3.5). We first introduce notation to remove the cumbersome term

$$\cdots\ \dot{W}^b{}_x W^{cx}\epsilon_{abc}\ .$$

Define

$$\{\dot{W}\}_a = \dot{W}^b{}_x W^{cx}\epsilon_{abc}\ .$$

We now adopt a 0-index notation so that

$$\Pi^a \dot{W}^b{}_x W^{cx}\epsilon_{abc} = \Pi^a\{\dot{W}\}_a \ \longrightarrow\ \Pi\{\dot{W}\}\ .$$

The algebra of reducing the action is greatly simplified by the following properties of $\{\dot{W}\}$, the details of which are found in Appendix C:

$$(\mathcal{U}\Pi)\,\{\mathcal{U}\dot{W}\} = \Pi\{\dot{W}\} \tag{C.1}$$

$$\Pi\{\dot{\overline{W_1W_2}}\} = \Pi\{W_1\dot{W}_2\} + \Pi\{\dot{W}_1\} \tag{C.2}$$

$$\Pi\left(\{\dot{W}_1\} - \{\dot{W}_2\}\right) = \left(W_2^{-1}\Pi\right)\{\dot{\overline{W_2^{-1}W_1}}\} \tag{C.3}$$

With this new notation, the action, *without* the constraints yet imposed, reads

$$\begin{aligned} I \;=\; & (W_{CA}^{-1}\rho)\{\dot{\overline{W_{CA}^{-1}W_{DB}^{-1}W_{CA}W_{DB}}}\} \\ & -\,(W_{CA}^{-1}\Pi_{CA})\{\dot{W}_{DB}\} \\ & \quad +\,(W_{DB}^{-1}\Pi_{DB})\{\dot{\overline{W_{DB}^{-1}W_{CA}W_{DB}}}\}\ . \end{aligned}$$

An example of the $\{\dot{W}\}$ algebra is shown below in the $T^2\#T^2$ case. If we now impose the constraints on the system, the first term above vanishes, as $\{\dot{\mathbf{1}}\} \propto \dot{\mathbf{1}} = 0$. The action is finally reduced to

$$\begin{aligned} I &= (W_{DB}^{-1}\Pi_{DB})\{\dot{W_{CA}}\} - (W_{CA}^{-1}\Pi_{CA})\{\dot{W_{DB}}\} \\ &\quad + \xi_a{}^b(W_{CA}W_{DB} - W_{DB}W_{CA})^a{}_b \\ &\quad\quad + \zeta_a((\mathbf{1} - W_{CA})\Pi_{DB} - (\mathbf{1} - W_{DB})\,\Pi_{CA})^a \end{aligned} \tag{3.6}$$

where the constraints are included with the Lagrange multipliers ξ and ζ.

The most important feature of this result in that it is written entirely in terms of the holonomies $[W_{CA}|\Pi_{CA}]$ and $[W_{DB}|\Pi_{DB}]$. All reference to the geometry of the tile has been removed. This is exactly the property that flat connections are completely specified by a collection of holonomies. Furthermore, we see that the action is very nearly in the canonical $p\dot{q}$ form, except that $\{\dot{W}\}$ is not merely $\frac{d}{dt}W$. Before we attempt to extract the canonical variables, we will look at the $T^2\#T^2$ case for more insight.

3.2 Consistency of the Constraints

It appears from (3.6) that 6 degrees of freedom will be removed by imposing the constraints on the system: 3 from the vector constraint, and just 3 from the matrix constraint, due to the symmetry of Lorentz transformations. In fact, explicitly writing out the constraints shows that these six equations are not independent, and only four degrees of freedom are directly removed. Also, we will see that under a interesting identification of the Π with translations in the solution space of W's, the relations (3.4) and (3.5) contain the same constraints. This association is not precise, but merely suggestive of the roles of W and Π as canonically conjugate variables.

The first constraint of the T^2-system is the closure relation

$$W_{CA}{}^a{}_b W_{DB}{}^b{}_c - W_{DB}{}^a{}_b W_{CA}{}^b{}_c = 0 \ .$$

The W's are forced by this constraint to lie on a surface in W-space. Hence variations of the solution W_{CA} & W_{DB} must also lie on the this surface. That is, variations of the constraint with respect to the the coordinates W^a_b must vanish. We will see that this new condition is a copy of the second constraint on the system, subject to some interpretation of the variables. The variation of $W_{CA}W_{DB} - W_{DB}W_{CA}$ is calculated by

$$\begin{aligned}\Delta^a{}_c \;=\; & \frac{\delta W_{CA}{}^a{}_b}{W_{CA}{}^x{}_y}\delta W_{CA}{}^x{}_y W_{DB}{}^b{}_c - W_{DB}{}^a{}_b \frac{\delta W_{CA}{}^b{}_c}{W_{CA}{}^x{}_y}\delta W_{CA}{}^x{}_y \\ & + W_{CA}{}^a{}_b \frac{\delta W_{DB}{}^b{}_c}{W_{DB}{}^x{}_y}\delta W_{DB}{}^x{}_y - \frac{\delta W_{DB}{}^a{}_b}{W_{DB}{}^x{}_y}\delta W_{DB}{}^x{}_y W_{CA}{}^b{}_c \ ,\end{aligned}$$

where

$$\frac{\delta W^a_b}{\delta W^x_y} = \delta^a_x \delta^y_b \ .$$

Thus

$$\begin{aligned}\Delta^a{}_c \;=\; & \delta W_{CA}{}^a{}_b W_{DB}{}^b{}_c - W_{DB}{}^a{}_b \delta W_{CA}{}^b{}_c \\ & + W_{CA}{}^a{}_b \delta W_{DB}{}^b{}_c - \delta W_{DB}{}^a{}_b W_{CA}{}^b{}_c\end{aligned}$$

or, in matrix form,

$$\begin{aligned}\Delta \;=\; & \delta W_{CA} W_{DB} - W_{DB}\delta W_{CA} \\ & + W_{CA}\delta W_{DB} - \delta W_{DB} W_{CA} \ .\end{aligned}$$

By varying the constant metric η^{ab}, observe that $(\delta W)W^{-1}$ is antisymmetric:

$$\begin{aligned} 0 &= \delta(\eta^{ab}) \\ &= \delta(W^a_c W^{bc}) \\ &= \delta W^a_c W^{bc} + W^a_c \delta W^{bc} \\ &= \delta W^a_c W^{bc} + \delta W^b_c W^{ac} \ . \end{aligned} \tag{3.7}$$

Insert factors of $\mathbf{1} = W^{-1}W$ where necessary into the expression for Δ to get all δW into this antisymmetric form:

$$\begin{aligned}
\Delta &= (\delta W_{CA} W_{CA}^{-1}) W_{CA} W_{DB} - W_{DB} (\delta W_{CA} W_{CA}^{-1}) W_{CA} \\
&\quad + W_{CA} (\delta W_{DB} W_{DB}^{-1}) W_{DB} - (\delta W_{DB} W_{DB}^{-1}) W_{DB} W_{CA} \\
&= (\delta W_{CA} W_{CA}^{-1}) W_{CA} W_{DB} - W_{DB} (\delta W_{CA} W_{CA}^{-1}) W_{DB}^{-1} W_{DB} W_{CA} \\
&\quad + W_{CA} (\delta W_{DB} W_{DB}^{-1}) W_{CA}^{-1} W_{CA} W_{DB} - (\delta W_{DB} W_{DB}^{-1}) W_{DB} W_{CA}
\end{aligned}$$

where the last step puts each $(\delta W\, W^{-1})$ into a similarity transformation. We right-multiply by $W_{CA} W_{DB}$, or equivalently $W_{DB} W_{CA}$, without changing the vanishing variation and apply the relation (3.4) to find

$$\begin{aligned}
0 &= (\delta W_{CA} W_{CA}^{-1}) - W_{DB} (\delta W_{CA} W_{CA}^{-1}) W_{DB}^{-1} \\
&\quad + W_{CA} (\delta W_{DB} W_{DB}^{-1}) W_{CA}^{-1} - (\delta W_{DB} W_{DB}^{-1}) \ .
\end{aligned}$$

Now as $(\delta W\, W^{-1})$ is antisymmetric, we can replace this matrix by an equivalent vector λ^a defined as

$$\lambda^a = \epsilon^a{}_{bc} \delta W^b{}_d\, W^{cd} \ .$$

The similarity transformations $\mathcal{U}(\delta W\, W^{-1})\mathcal{U}^{-1}$ become linear transformations $\mathcal{U}\lambda$ of the vector. "Rotate" the antisymmetric matrix $(\delta W\, W^{-1})$ with a Lorentz transformation $\mathcal{U}$:

$$\delta W^b{}_d W^{cd} \longrightarrow \mathcal{U}^b{}_x \delta W^x{}_d W_y{}^d \mathcal{U}^{cy} \ .$$

This new matrix is also antisymmetric in b, c: interchange b and c to find

$$\begin{aligned}
\mathcal{U}^c{}_x \left(\delta W^x{}_d W_y{}^d\right) \mathcal{U}^{by} &= \mathcal{U}^c{}_x \left(\delta W^x{}_d W^{yd}\right) \mathcal{U}^b{}_y \\
&= -\, \mathcal{U}^c{}_x \left(\delta W^y{}_d W^{xd}\right) \mathcal{U}^b{}_y
\end{aligned}$$

as $\delta W\, W^{-1}$ is antisymmetric. Now relabel the dummy indices, interchanging x and y:

$$\begin{aligned}
&= -\, \mathcal{U}^c{}_y \left(\delta W^x{}_d W^{yd}\right) \mathcal{U}^b{}_x \\
&= -\, \mathcal{U}^b{}_x \left(\delta W^x{}_d W_y{}^d\right) \mathcal{U}^{cy} \ .
\end{aligned}$$

Define a new vector λ'^a from this antisymmetric matrix:

$$\lambda'^a = \epsilon^a{}_{bc}\left(\mathcal{U}^b{}_x\mathcal{U}^c{}_y\delta W^x{}_d W^{yd}\right) \ .$$

By (A.4),

$$\epsilon^a{}_{bc}\mathcal{U}^b{}_x\mathcal{U}^c{}_y = \mathcal{U}^a{}_z\epsilon^z{}_{xy} \ ,$$

so

$$\lambda'^a = \mathcal{U}^a{}_z\epsilon^z{}_{xy}\delta W^x{}_d W^{yd}$$

which is simply

$$\lambda'^a = \mathcal{U}^a{}_z\lambda^z \ .$$

Under this substitution of λ^a, the variation of the constraint on the W's is equivalent to the constraint

$$\begin{aligned} 0 &= \lambda_{CA} - W_{DB}\lambda_{CA} + W_{CA}\lambda_{DB} - \lambda_{DB} \\ &= (\mathbf{1} - W_{DB})\lambda_{CA} - (\mathbf{1} - W_{CA})\lambda_{DB} \ . \end{aligned}$$

This is exactly the second constraint (3.5) under the identification

$$\lambda^a \longleftrightarrow \Pi^a \ .$$

The λ^a are infinitesimal translations in the space of W's, suggesting the holonomy components Π are related to translation-generating momenta, conjugate to the configuration variables W. We will expand on this idea further, after considering the $T^2\#T^2$ case, where we reproduce the almost canonical $p\dot{q}$ action.

3.3 A Model of T^2

To transform the action in $\{\dot{W}\}$ and Π into something that can be written on the back of an envelope, we construct an explicit model for T^2 and find the Poincaré transformations

$[W|\Pi]$ across its boundaries. We will see that even a very simple model reveals interesting details.

It is easy to construct a region representing the torus because the plane $\mathbf{R}^2$ can be tiled in unit squares. By identifying points on opposite sides of the square, the torus' topology is produced. To bring this tile into the arena where we can study the action, we must attach the holonomies $[W_{CA}|\Pi_{CA}]$ and $[W_{DB}|\Pi_{DB}]$ between the identified sides. The constraints (3.4) and (3.5) make this a fairly simple procedure.

Because the Lorentz transformations W_{CA} and W_{DB} commute, they must be boosts in the *same direction*. Let us choose coordinates $(x^1, x^2, x^3) = (t, x, y)$ over the tiled $\mathbf{R}^2$ plane so that this direction is the x-direction, with the origin at corner 1 of the tile. Recall from §1.2 that boosts in the x-direction are generated by exponentiation of the matrix J_3. We can choose

$$W_{CA} = e^{\mu J_3} = \begin{pmatrix} \cosh(\mu) & \sinh(\mu) & 0 \\ \sinh(\mu) & \cosh(\mu) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$W_{DB} = e^{\alpha\mu J_3} = \begin{pmatrix} \cosh(\alpha\mu) & \sinh(\alpha\mu) & 0 \\ \sinh(\alpha\mu) & \cosh(\alpha\mu) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

where μ is some boost parameter and $\alpha\mu$ ensures W_{DB} is parallel to W_{CA}.

It is interesting to look at the translation constraint (3.5) with respect to this choice of W_{CA} and W_{DB}. The constraint reads

$$\begin{pmatrix} 1-\cosh(\mu) & -\sinh(\mu) & 0 \\ -\sinh(\mu) & 1-\cosh(\mu) & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} {\Pi_{DB}}^1 \\ {\Pi_{DB}}^2 \\ {\Pi_{DB}}^3 \end{pmatrix}$$

$$= \begin{pmatrix} 1-\cosh(\alpha\mu) & -\sinh(\alpha\mu) & 0 \\ -\sinh(\alpha\mu) & 1-\cosh(\alpha\mu) & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} {\Pi_{CA}}^1 \\ {\Pi_{CA}}^2 \\ {\Pi_{CA}}^3 \end{pmatrix} .$$

While ${\Pi_{DB}}^1, {\Pi_{DB}}^2, {\Pi_{CA}}^1, {\Pi_{CA}}^2$ are coupled by two equations, ${\Pi_{DB}}^3$ and ${\Pi_{CA}}^3$ are completely unconstrained. That is, for any translations Π^3 in the y-direction, the sequence of transformations that circumvents the tile vertex is still the Identity. While initially it appeared that the constraints would directly remove six degrees of freedom (df) from the system, in fact only four are removed. This is not merely an artifact of our particular choice of the x-direction for the boosts, for a rotation will not affect the indeterminacy of the matrices $\mathbf{1} - W_{DB}$ or $\mathbf{1} - W_{CA}$. We will see below in the more complicated $T^2 \# T^2$ case that this failure of the six constraints to remove six df is unique among the non-zero genus surfaces to the torus. Furthermore, whereas the Lorentz component of the holonomies confine the transformations to a surface of constant $\sqrt{t^2 - x^2}$, the translation carries vectors off this hyperbolic plane. With some translation in the t- and x-directions and arbitrary y-translation, it seems possible that the sequence of transformations around the vertex can be a loop with time-like sections, carrying vectors along closed time-like curves.

To attach these Lorentz transformations to the unit square recall that W is defined at the link between the values of the field $\mathcal{U}^a_{\ b}$ across the identification boundary: $W_{21}\mathcal{U}_2 = \mathcal{U}_1$. The field $\mathcal{U}^a_{\ b}$ which is consistent with this choice of W's is found by treating the transformations W as a sort of "phase difference" between identified points. Referring again to Figure 3.1, first consider just the corners labelled 1,2,3,4. As the vertex at corner 1 on side D maps to the vertex at corner 2 on side B under W_{DB}, $\mathcal{U}(2) = W_{DB}\mathcal{U}(1)$. To find $\mathcal{U}(3)$, we see that $\mathcal{U}(3)$ maps onto $\mathcal{U}(2)$ under W_{CA}, so $\mathcal{U}(2) = W_{CA}\mathcal{U}(3)$ or

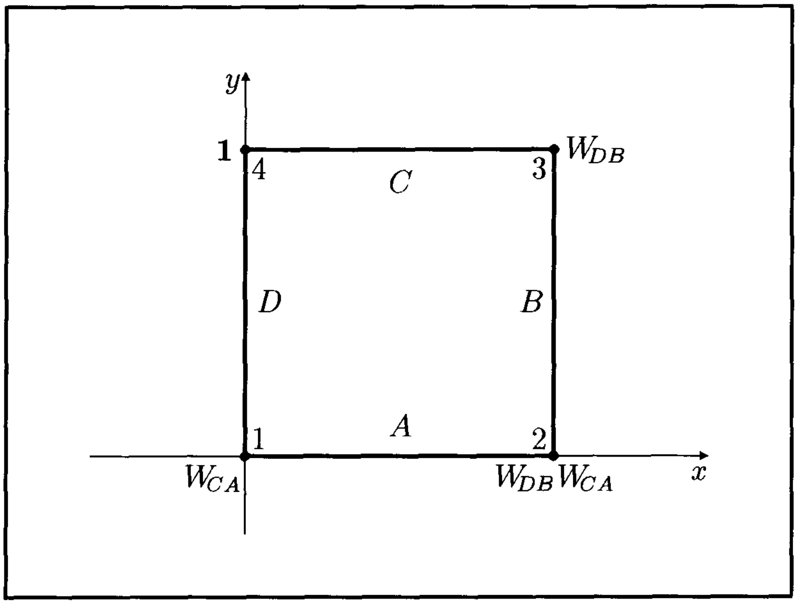

Figure 3.3: Values for the 'phase' $\mathcal{U}$ at the corners of the tile.

$\mathcal{U}(3) = W_{CA}^{-1}\mathcal{U}(2) = W_{CA}^{-1}W_{DB}\mathcal{U}(1)$. Similarly, $\mathcal{U}(4) = W_{CA}^{-1}\mathcal{U}(1)$. Now if we calculate $\mathcal{U}(3)$ from $\mathcal{U}(4)$, instead of $\mathcal{U}(2)$, we write $\mathcal{U}(3) = W_{DB}\mathcal{U}(2) = W_{DB}W_{CA}^{-1}\mathcal{U}(1)$. So that the result is independent of the choice of evaluation, it must be true that $W_{CA}^{-1}W_{DB} = W_{DB}W_{CA}^{-1}$ or $W_{DB}W_{CA} = W_{CA}W_{DB}$, exactly the constraint we encountered earlier. Since the transformations W are determined only by the phase *difference* between the values of $\mathcal{U}$ at identified points, there is an over all arbitrary choice for $\mathcal{U}(1)$. Setting $\mathcal{U}(1) = W_{CA}$ gives simple values to the corners of the tile (See Figure 3.3). These corner values are now interpolated smoothly over the tile, taking care to keep the correct W phase difference between identified points on the boundary. The particular choice of interpolation does not change the action, which depends only on the difference across the tile (or across the boundary). One simple example of an interpolation is

$$\mathcal{U}(x, y) = \mathrm{e}^{(1-f(y)+\alpha f(x))\mu J_3}$$

where f is linear between $f(0) = 0$ and $f(1) = 1$. Smoother interpolations (quadratic, cubic, ...) can be used where continuity of derivatives of $\mathcal{U}$ is required in further calculations.

The final step in the construction of our model is finding a set of Π's. These translations entered the calculation through $W_{21}\rho_2 + \Pi_{21} = \rho_1$, where $\rho^a(x^b) = F^a(x^b)$ is the function chosen to eliminate the $\hat{e}_i{}^a$. A simple choice is to give the translations only y-components, $\Pi_{CA} = (0,0,a)$ and $\Pi_{DB} = (0,0,b)$, and set $\rho^a(x^b) = x^a$. The Lorentz components of the holonomies preserve $\sqrt{t^2 - x^2}$ while the translations shift the y-components of parallel transported vectors by a constant. The tangent space over the tile is a parallelogram on the surface of constant $\sqrt{t^2 - x^2}$ (See Figure 3.4). The relation $W_{21}\rho_2 + \Pi_{21} = \rho_1$ tells us which points are identified in the tangent space.

3.4 An Alternative Approach to the Torus

Another, more "standard", method for studying (2+1)-dimensional GR on a torus is based on the fact that of all the compact surfaces, only the genus 1 torus can support a continuous, non-vanishing tangent vector field. The approach, therefore, cannot be generalized to higher genus surfaces.

Because the plane $\mathbf{R}^2$ can be tiled in unit squares, we break spacetime into $\mathbf{R}^2 \otimes \mathbf{R}$. By identifying opposite sides of a unit square, or defining spatial coordinates x and y to be periodic with period 1, the spacetime $T^2 \otimes \mathbf{R}$ is generated. With this geometry, Carlip [4] proceeds by specifying the holonomies of this surface, two commuting Poincaré transformations:

$$\Lambda_1 : (t,x,y) \rightarrow (t\cosh\lambda + x\sinh\lambda,\ x\cosh\lambda + t\sinh\lambda,\ y + a)$$

$$\Lambda_2 : (t,x,y) \rightarrow (t\cosh\mu + x\sinh\mu,\ x\cosh\mu + t\sinh\mu,\ y + b)$$

A *dreibein* and connection which exhibit these holonomies under a path-ordered integral

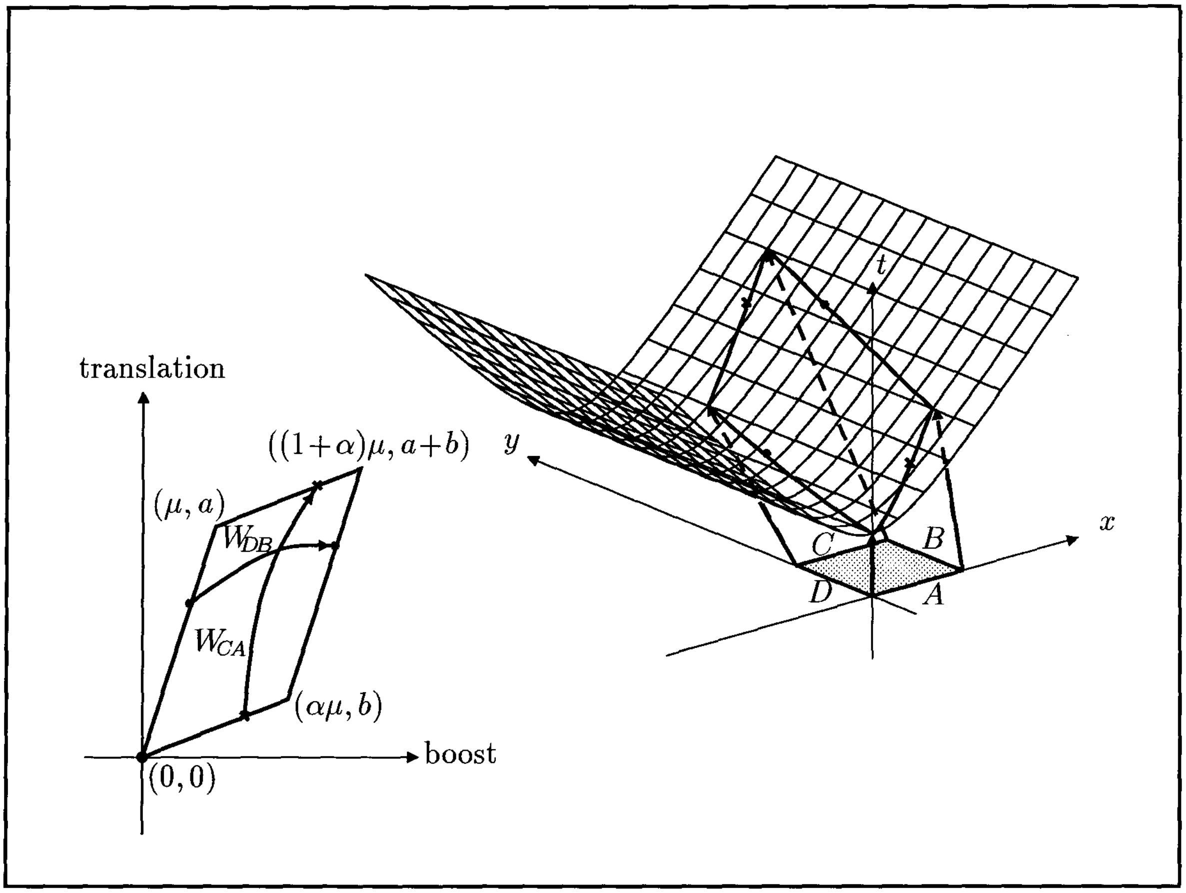

Figure 3.4: Tiles of the tangent space over the torus are parallelograms in boost-translation space.

or Wilson line [5] calculation of the holonomy are $e^1 = e^2 = 0$, $\omega^1 = \omega^2 = 0$ and

$$e^3 = (0, a, b) \qquad \omega^3 = (0, \lambda, \mu) \ .$$

This *dreibein* produces a singular metric, but we can gauge transform to a non-singular system:

$$\begin{aligned} e^1 &= (-\dot{\beta}, 0, 0) & \omega^1 &= (0,0,0) \\ e^2 &= (0, \beta\lambda, \beta\mu) & \omega^2 &= (0,0,0) \\ e^3 &= (0, a, b) & \omega^3 &= (0, \lambda, \mu) \end{aligned}$$

where $\beta(t)$ is a function only of the time on the slice. On each slice of constant time, a constant, continuous non-vanishing tangent vector field is realized. The metric arising from this choice of *dreibein* is

$$\begin{aligned} ds^2 = {} & \dot{\beta}^2 dt^2 - (a^2 + \beta^2\lambda^2)dx^2 \\ & -2(ab + \beta^2\lambda\mu)dx\, dy - (b^2 + \beta^2\mu^2)dy^2 \ . \end{aligned}$$

Now Carlip observes that the spatial part of the metric, the metric on the torus, is unchanged by the two coordinate transformations

$$\begin{aligned} &\left(x + \left(\frac{ab + \beta^2\lambda\mu}{a^2 + \beta^2\lambda^2}\right)y,\ \frac{\beta(a\mu - \lambda b)}{a^2 + \beta^2\lambda^2}\, y\right) \longrightarrow \\ &\qquad \left((x+1) + \left(\frac{ab + \beta^2\lambda\mu}{a^2 + \beta^2\lambda^2}\right)y,\ \frac{\beta(a\mu - \lambda b)}{a^2 + \beta^2\lambda^2}\, y\right) \\ &\left(x + \left(\frac{ab + \beta^2\lambda\mu}{a^2 + \beta^2\lambda^2}\right)y,\ \frac{\beta(a\mu - \lambda b)}{a^2 + \beta^2\lambda^2}\, y\right) \longrightarrow \\ &\qquad \left(x + \left(\frac{ab + \beta^2\lambda\mu}{a^2 + \beta^2\lambda^2}\right)(y+1),\ \frac{\beta(a\mu - \lambda b)}{a^2 + \beta^2\lambda^2}(y+1)\right) \end{aligned}$$

The geometry of the space, therefore, is characterized by these two coordinate translations in a, b, λ, μ, the parameters that fix the holonomies. Treating a, b, λ, μ as a new set of coordinates, the Hamiltonian produced is simply

$$H = \beta(a\mu - \lambda b) \ .$$

One interpretation of the canonical variables is to take the boost parameters λ, μ as coordinates and the translations a, b as conjugate momenta.

This result is based on the existence of a non-singular spatial metric on the surface. Its spatial periodicity is equivalent to the periodic tiling of the tangent space over the torus. The holonomies (λ, a) and (μ, b) completely determine the action because of the flatness of the space.

A model for the genus $g = 1$ torus is simple to construct because the plane $\mathbf{R}^2$ can be tiled in regular $(4g = 4)$-gons, or squares. This surface is also easier to study than other genus surfaces because it is the only one that can support a non-vanishing tangent vector field, immediately giving the surface a non-singular metric. We now turn a to more complicated surface, the two-holed torus, where the results for T^2 are closely mimicked.

Chapter 4

A Two-Holed Torus

In this chapter we apply the results of Chapter 2 to the more complicated genus 2 two-holed torus, $T^2 \# T^2$. One would suspect that this surface is more difficult to study than T^2, for it is impossible to put a *continuous* non-vanishing tangent vector field onto this surface whose non-zero Euler characteristic $\chi = 2 - 2g = -2$ is non-zero.[6] We will find, however, that $T^2 \# T^2$ is simply the connected sum of two tori T^2, and the action is more complicated only because it is a functional of twice as many variables. This surface has $2g = 4$ generating non-trivial loops. It can be constructed by identifying pairs of sides of a $(4g = 8)$-sided polygon. Or equivalently, the two-holed torus can be conceived by a tiling of a plane with octagonal tiles, with eight tiles meeting at each vertex. As we shall see, but intuitively understand already, the plane $\mathbf{R}^2$ *cannot* be tiled in regular octagons without leaving gaps in the tiling. Hence we look to hyperbolic geometry where the condition that triangles have 180^o no longer applies. Before we construct such a tile and its collection of holonomies W and Π, we first consider the case of a general octagonal tiling, and translate the results of Chapter 2 into the $T^2 \# T^2$ variables.

4.1 The Action over a Two-Holed Torus

We calculate the action over the two-holed torus exactly as we did for T^2. Cover each surface of constant time with octagonal tiles. Label the sides of the octagon $A, B, \ldots, H$ and the corners $1, 2, \ldots, 8$ (See Figure 4.1(i)). The identification of sides is indicated by the arrows on each side. Gluing the matching sides together generates the

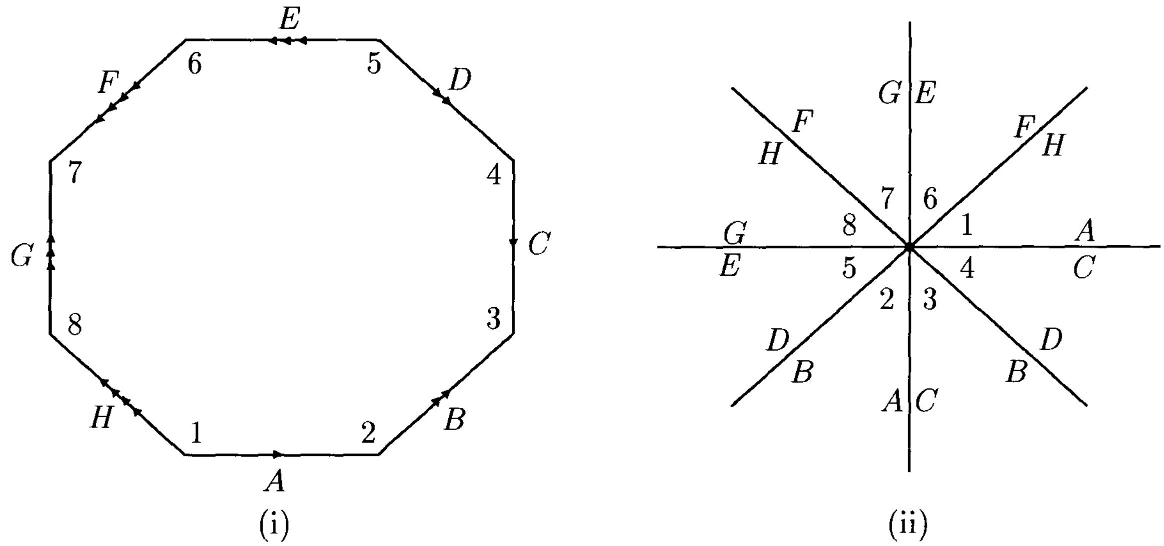

Figure 4.1: (i) Label for the sides and corners of the tile, and identifications of sides. (ii) The neighborhood of a vertex where eight tiles meet.

boundary-less two-holed torus. Note the reversed orientation in the pairs of identified sides $C\&A, D\&B, G\&E, H\&F$. The neighborhood of a vertex where eight tiles meet also shows the identifications and labels (See Figure 4.1(ii)). As before, we bootstrap around the vertex using the identifications and label the sides and regions of the vertex. Figure 4.1(ii) is a truer representation of the tile because the neighborhood of a vertex is a patch of $\mathbf{R}^2$, and the angular contribution of each tile is $\frac{\pi}{4}$. By drawing the *whole* tile on paper ($\mathbf{R}^2$) as opposed to the hyperbolic plane where the tile really sits, we are forced to stretch the angles out to $\frac{3\pi}{4}$.

Across the boundary between sides A and C we find from (2.8)

$$W_{CA}{}^a{}_b\, \rho_C{}^a + \Pi_{CA}{}^a = \rho_A{}^a \ .$$

Here W_{CA} is the Lorentz transformation which relates $W_{CA}\mathcal{U}_C = \mathcal{U}_A$ and Π_{CA} is the constant defined by (2.7). This relation holds all along the edge between sides A and C

as $W^a_{\ b,i} = 0$ (2.5), so we push the result down to the 3-2 region of the vertex where

$$W_{CA}{}^a{}_b \, \rho_C{}^a(3) + \Pi_{CA}{}^a = \rho_A{}^a(2) \ .$$

The term $\rho_C{}^a(3)$ is the limiting value of $\rho_C{}^a$ as the vertex is approached along edge C. Similar relations across each of the identified edges are defined in Figure 4.2:

$$\begin{array}{ll}
\text{Figure 4.2 (i)} & \left\{ \begin{array}{l} W_{CA}{}^a{}_b \, \rho_C{}^b(4) + \Pi_{CA}{}^a = \rho_A{}^a(1) \\ W_{CA}{}^a{}_b \, \rho_C{}^b(3) + \Pi_{CA}{}^a = \rho_A{}^a(2) \end{array} \right. \\
\text{Figure 4.2 (ii)} & \left\{ \begin{array}{l} W_{DB}{}^a{}_b \, \rho_D{}^b(4) + \Pi_{DB}{}^a = \rho_C{}^a(3) \\ W_{DB}{}^a{}_b \, \rho_D{}^b(5) + \Pi_{DB}{}^a = \rho_C{}^a(2) \end{array} \right. \\
\text{Figure 4.2 (iii)} & \left\{ \begin{array}{l} W_{GE}{}^a{}_b \, \rho_G{}^b(7) + \Pi_{GE}{}^a = \rho_E{}^a(6) \\ W_{GE}{}^a{}_b \, \rho_G{}^b(8) + \Pi_{GE}{}^a = \rho_E{}^a(5) \end{array} \right. \\
\text{Figure 4.2 (iv)} & \left\{ \begin{array}{l} W_{HF}{}^a{}_b \, \rho_H{}^b(8) + \Pi_{HF}{}^a = \rho_F{}^a(7) \\ W_{HF}{}^a{}_b \, \rho_H{}^b(1) + \Pi_{HF}{}^a = \rho_F{}^a(6) \end{array} \right.
\end{array} \tag{4.1}$$

We must evaluate $\Delta\rho_1{}^a$ by integrating along identified edges to produce the action. The return integrals along 2-sides are accounted for by the $\dot{W}W$ terms, so we need only consider the 1-sides, from corners $1 \to 2$, $2 \to 3$, $5 \to 6$, and $6 \to 7$. The action becomes

$$\begin{aligned} I \ = \ \{ \ & (\rho_A{}^a(2) - \rho_A{}^a(1)) \dot{W}_{CA}{}^b{}_x W_{CA}{}^{cx} \\ & + (\rho_B{}^a(3) - \rho_B{}^a(2)) \dot{W}_{DB}{}^b{}_x W_{DB}{}^{cx} \\ & + (\rho_E{}^a(6) - \rho_E{}^a(5)) \dot{W}_{GE}{}^b{}_x W_{GE}{}^{cx} \\ & + (\rho_F{}^a(7) - \rho_F{}^a(6)) \dot{W}_{HF}{}^b{}_x W_{HF}{}^{cx} \ \} \, \epsilon_{abc} \ . \end{aligned} \tag{4.2}$$

The field ρ^a is discontinuous across the boundaries, the discontinuities related to the holonomies W and Π. But in region 4 of the vertex, for instance, $\rho_D(4) = \rho_C(4)$, as this is the limiting value of a function ρ^a, continuous on the *tile*, without any identification of points required. Call the common value just ρ_4. An analogous relation holds in each

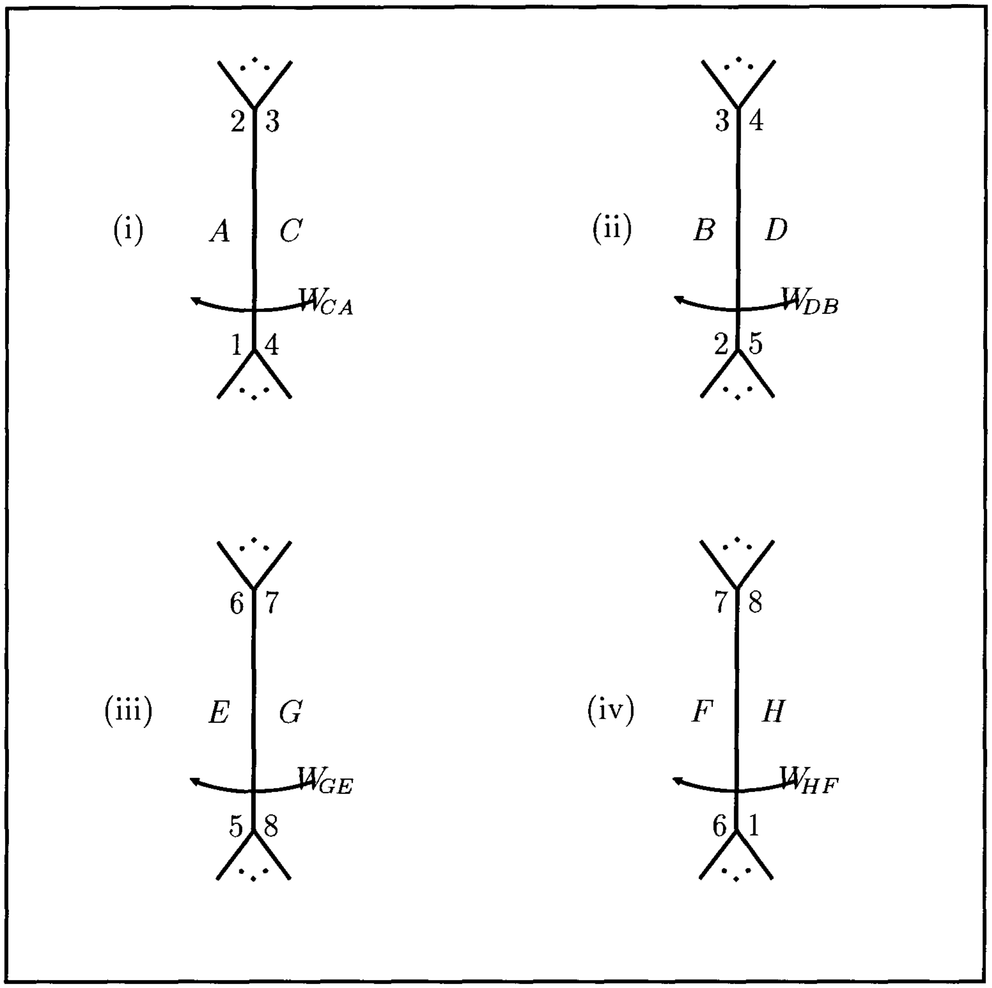

Figure 4.2: The transformations W across the identified edges (i)$C\&A$, (ii)$D\&B$, (iii)$G\&E$, and (iv)$H\&F$.

region of the vertex:

$$\begin{array}{ll} \rho_A(1) = \rho_H(1) = \rho_1 & \rho_A(2) = \rho_B(2) = \rho_2 \\ \rho_B(3) = \rho_C(3) = \rho_3 & \rho_C(4) = \rho_D(4) = \rho_4 \\ \rho_D(5) = \rho_E(5) = \rho_5 & \rho_E(6) = \rho_F(6) = \rho_6 \\ \rho_F(7) = \rho_G(7) = \rho_7 & \rho_G(8) = \rho_H(8) = \rho_8 \end{array}$$

By bootstrapping around the vertex from region 1 with these relations and the W, Π above (4.1) we can express all the ρ_i in terms of ρ_1. As vectors and matrices rather than in components, we see

$$\begin{aligned} \rho_2 &= W_{CA}\,\rho_3 + \Pi_{CA} \\ &= W_{CA}(W_{DB}\,\rho_4 + \Pi_{DB}) + \Pi_{CA} \\ &= W_{CA}(W_{DB}(W_{CA}^{-1}(\rho_1 - \Pi_{CA})) + \Pi_{DB}) + \Pi_{CA} \end{aligned}$$

Repeating the process for each corner at the vertex gives

$$\left.\begin{aligned} \rho_2 &= W_{CA}(W_{DB}(W_{CA}^{-1}(\rho_1 - \Pi_{CA})) + \Pi_{DB}) + \Pi_{CA} \\ \rho_3 &= W_{DB}(W_{CA}^{-1}(\rho_1 - \Pi_{CA})) + \Pi_{DB} \\ \rho_4 &= W_{CA}^{-1}(\rho_1 - \Pi_{CA}) \\ \rho_5 &= W_{GE}W_{HF}^{-1}(W_{GE}^{-1}(W_{HF}\rho_1 + \Pi_{HF} - \Pi_{GE}) - \Pi_{HF}) + \Pi_{GE} \\ \rho_6 &= W_{HF}\rho_1 + \Pi_{HF} \\ \rho_7 &= W_{GE}^{-1}(W_{HF}\rho_1 + \Pi_{HF} - \Pi_{GE}) \\ \rho_8 &= W_{HF}^{-1}(W_{GE}^{-1}(W_{HF}\rho_1 + \Pi_{HF} - \Pi_{GE}) - \Pi_{HF}) \end{aligned}\right\} \qquad (4.3)$$

We arbitrarily chose to write $\rho_8 = \rho_8(\rho_7(\rho_6(\rho_1)))$ in finding the last result in this list, but we equally could have bootstrapped the other way around the vertex, writing $\rho_8 = \rho_8(\rho_5(\rho_2(\rho_3(\rho_4(\rho_1)))))$. For the ρ_i to be well-defined, it must be true that a complete circuit of the vertex is the Identity:

$$\rho_1(\rho_4(\rho_3(\rho_2(\rho_5(\rho_8(\rho_7(\rho_6(\rho_1)\cdots)) = \rho_1 \ .$$

The Lorentz part of this Poincaré transformation must be the Identity and the translation must vanish. The transformation carrying ρ_1 around the vertex back onto ρ_1 is given by

$$\begin{aligned}
\rho_1 \;=\;& W_{CA}W_{DB}^{-1}W_{CA}^{-1}W_{DB}W_{GE}W_{HF}^{-1}W_{GE}^{-1}W_{HF}\,\rho_1 \\
&+ W_{CA}W_{DB}^{-1}W_{CA}^{-1}W_{DB}W_{GE}W_{HF}^{-1}W_{GE}^{-1}\Pi_{HF} \\
&- W_{CA}W_{DB}^{-1}W_{CA}^{-1}W_{DB}W_{GE}W_{HF}^{-1}W_{GE}^{-1}\Pi_{GE} \\
&- W_{CA}W_{DB}^{-1}W_{CA}^{-1}W_{DB}W_{GE}W_{HF}^{-1}\Pi_{HF} \\
&+ W_{CA}W_{DB}^{-1}W_{CA}^{-1}W_{DB}\Pi_{GE} \\
&+ W_{CA}W_{DB}^{-1}W_{CA}^{-1}\Pi_{DB} \\
&- W_{CA}W_{DB}^{-1}W_{CA}^{-1}\Pi_{CA} \\
&- W_{CA}W_{DB}^{-1}\Pi_{DB} \\
&+ \Pi_{CA}\;. \qquad (4.4)
\end{aligned}$$

The Lorentz transformations $W_{CA}, W_{DB}, W_{GE}, W_{HF}$ must obey the closure relation

$$W_{CA}W_{DB}^{-1}W_{CA}^{-1}W_{DB}W_{GE}W_{HF}^{-1}W_{GE}^{-1}W_{HF} = \mathbf{1}\;. \qquad (4.5)$$

Again we see the fundamental group $\pi_1(T^2\#T^2)$ represented by the W in the gauge group. Define

$$\begin{aligned}
\Omega_1 &= W_{CA}W_{DB}^{-1}W_{CA}^{-1}W_{DB}\;, \\
&= W_{HF}^{-1}W_{GE}W_{HF}W_{GE}^{-1}\;.
\end{aligned} \qquad (4.6)$$

The constraint (4.5) is equivalent to

$$\Omega_1 - \Omega_2 = 0\;.$$

While $\Omega_1 = \Omega_1(W_{CA}, W_{DB})$ but $\Omega_2 = \Omega_2(W_{GE}, W_{HF})$, they are the same transformation: Ω_1 is a transformation halfway around the vertex in one direction, Ω_2 is a transformation halfway around in the other direction, and the two results coincide there. The decoupling of the transformations into $ABCD$ terms and $EFGH$ terms is an indication of the connected sum construction of $T^2 \# T^2$. The constraint (4.5) shows that the two tori glue together smoothly. By inserting the factors missing from the cycle (4.5) and removing the resulting factors of $\mathbf{1}$, we can reduce the translation part of (4.4) to the constraint

$$\begin{aligned} 0 &= \Pi_{CA} - W_{CA} W_{DB}^{-1} W_{CA}^{-1} \Pi_{CA} \\ &\quad - W_{CA} W_{DB}^{-1} \Pi_{DB} + W_{CA} W_{DB}^{-1} W_{CA}^{-1} \Pi_{DB} \\ &\quad - W_{HF}^{-1} \Pi_{GE} + W_{HF}^{-1} W_{GE} W_{HF} W_{GE}^{-1} \Pi_{GE} \\ &\quad + W_{HF}^{-1} \Pi_{HF} - W_{HF}^{-1} W_{GE} \Pi_{HF} \ . \end{aligned} \tag{4.7}$$

We can now write the action in terms of ρ_1 (or simply ρ), W, and Π, plus the constraints (4.5) and (4.7) with the help of (4.3). Again we introduce

$$\{\dot{W}\}_a = \dot{W}^b{}_x W^{cx} \epsilon_{abc}$$

to more easily write

$$\Pi^a \dot{W}^b{}_x W^{cx} \epsilon_{abc} = \Pi\{\dot{W}\} \ .$$

We use the properties of $\{\dot{W}\}$ found in Appendix C to reduce the action to a simple form:

$$\begin{aligned} I &= -\rho\{\dot{\overline{W_{CA} W_{DB}^{-1} W_{CA}^{-1} W_{DB}}}\} + \rho\{\dot{\overline{W_{HF}^{-1} W_{GE} W_{HF} W_{GE}^{-1}}}\} \\ &\quad + \left(W_{DB}^{-1} \Pi_{CA}\right) \{\dot{\overline{W_{DB}^{-1} \Omega_1}}\} - \left(W_{DB}^{-1} \Pi_{DB}\right) \{\dot{\overline{W_{CA}^{-1} \Omega_1}}\} \\ &\quad - \Pi_{GE}\{\dot{\overline{W_{HF} \Omega_2}}\} + \left(W_{GE} W_{HF}^{-1} \Pi_{HF}\right) \{\dot{\overline{W_{GE} \Omega_2}}\} \ . \end{aligned}$$

The term in Π_{CA}, for example, is found as follows: Upon substituting the relations (4.3) into the action (4.2), the terms containing Π_{CA} are

$$\begin{aligned} &- \left(W_{CA} W_{DB} W_{CA}^{-1} \Pi_{CA}\right) \{\dot{W}_{CA}\} + \Pi_{CA}\{\dot{W}_{CA}\} \\ &\quad - \left(W_{DB} W_{CA}^{-1} \Pi_{CA}\right) \{\dot{W}_{DB}\} + \left(W_{CA} W_{DB} W_{CA}^{-1} \Pi_{CA}\right) \{\dot{W}_{DB}\} \\ &\quad - \Pi_{CA}\{\dot{W}_{DB}\} \ . \end{aligned}$$

Rewrite this using (C.1) to give the terms the form of the right-hand side of (C.3):

$$
\begin{aligned}
= &\left(\underbrace{-W_{CA}W_{DB}W_{CA}^{-1}}\Pi_{CA}\right)\{\dot{\overline{\underbrace{W_{CA}W_{DB}W_{CA}^{-1}}W_{CA}W_{DB}^{-1}W_{CA}^{-1}W_{CA}}}\} \\
&\Pi_{CA}\{\dot{W}_{CA}\} - \Pi_{CA}\{\dot{W}_{DB}\} - \left(\underbrace{W_{DB}W_{CA}^{-1}}\Pi_{CA}\right)\{\dot{\overline{\underbrace{W_{DB}W_{CA}^{-1}}W_{CA}W_{DB}^{-1}W_{DB}}}\} \\
&+\left(\underbrace{W_{CA}W_{DB}W_{CA}^{-1}}\Pi_{CA}\right)\{\dot{\overline{\underbrace{W_{CA}W_{DB}W_{CA}^{-1}}W_{CA}W_{DB}^{-1}W_{CA}^{-1}W_{DB}}}\}\ .
\end{aligned}
$$

Now apply (C.3) to remove the under-braced terms:

$$
\begin{aligned}
= &-\Pi_{CA}\{\dot{\overline{W_{CA}W_{DB}^{-1}}}\} + \Pi_{CA}\{\dot{\overline{W_{CA}W_{DB}^{-1}W_{CA}^{-1}}}\} \\
&+\Pi_{CA}\{\dot{W}_{CA}\} - \Pi_{CA}\{\dot{W}_{DB}\} - \Pi_{CA}\{\dot{W}_{CA}\} + \Pi_{CA}\{\dot{\overline{W_{CA}W_{DB}^{-1}}}\} \\
&+\Pi_{CA}\{\dot{\overline{W_{CA}W_{DB}^{-1}W_{CA}^{-1}W_{DB}}}\} - \Pi_{CA}\{\dot{\overline{W_{CA}W_{DB}^{-1}W_{CA}^{-1}}}\}\ .
\end{aligned}
$$

The only two remaining terms are grouped together with (C.3) to give

$$
= \left(W_{DB}^{-1}\Pi_{CA}\right)\{\dot{\overline{W_{DB}^{-1}W_{CA}W_{DB}^{-1}W_{CA}^{-1}W_{DB}}}\}\ .
$$

Substituting Ω_1 from (4.6) produces the Π_{CA} term in the action above.

By imposing the constraint $\Omega_1 = \Omega_2$, the terms in ρ cancel in the action, leaving the constrained action

$$
\begin{aligned}
I \;=\; &\left(W_{DB}^{-1}\Pi_{CA}\right)\{\dot{\overline{W_{DB}^{-1}\Omega_1}}\} - \left(W_{DB}^{-1}\Pi_{DB}\right)\{\dot{\overline{W_{CA}^{-1}\Omega_1}}\} \\
&-\Pi_{GE}\{\dot{\overline{W_{HF}\Omega_2}}\} + \left(W_{GE}W_{HF}^{-1}\Pi_{HF}\right)\{\dot{\overline{W_{GE}\Omega_2}}\} \\
&+\xi_a{}^b(\Omega_1 - \Omega_2)^a{}_b \\
&+\zeta_a(\Pi_{CA} - W_{CA}W_{DB}^{-1}W_{CA}^{-1}\Pi_{CA} + \cdots - W_{HF}^{-1}W_{GE}\Pi_{HF})^a\ .
\end{aligned}
\tag{4.8}
$$

The Lagrange multipliers ξ and ζ have been included to account for the constraints imposed by (4.5) and (4.7), respectively. Observe again that the terms in $ABCD$ are decoupled from those in $EFGH$.

As in the T^2 case, this action is written entirely in terms of the holonomies W and Π. The geometry of the tile has been removed. The action is very nearly in the canonical

$p\dot{q}$ form, except that the term $\{\dot{W}\}$ is not merely $\frac{d}{dt}W$. The only difference between the "simple" T^2 and the "difficult" $T^2\#T^2$ is that more variables have appeared to account for the increased number of incontractible loops and holonomies.

4.2 Consistency of the Constraints

Recall the first constraint on the system is that the Lorentz transformations W obey the closure condition (4.5)

$$W_{CA}W_{DB}^{-1}W_{CA}^{-1}W_{DB}W_{GE}W_{HF}^{-1}W_{GE}^{-1}W_{HF} = \mathbf{1} \ ,$$

or with (4.6),

$$\Omega_1 - \Omega_2 = 0 \ .$$

The second constraint forces the sequence of translations to vanish (4.7), closing the circuit of Poincaré transformations around the two-holed torus:

$$\begin{aligned} 0 \ = \ & \Pi_{CA} - W_{CA}W_{DB}^{-1}W_{CA}^{-1}\Pi_{CA} \\ & -W_{CA}W_{DB}^{-1}\Pi_{DB} + W_{CA}W_{DB}^{-1}W_{CA}^{-1}\Pi_{DB} \\ & -W_{HF}^{-1}\Pi_{GE} + W_{HF}^{-1}W_{GE}W_{HF}W_{GE}^{-1}\Pi_{GE} \\ & +W_{HF}^{-1}\Pi_{HF} - W_{HF}^{-1}W_{GE}\Pi_{HF} \ . \end{aligned}$$

The "configuration" variables W (they are not quite the configuration variables because the action is of the form $\Pi\{\dot{W}\}$, not merely $\Pi\dot{W}$) are forced by the first constraint to lie on a surface defined by the relation (4.5). Variations, or nearby solutions, must also lie in this surface. We will see that this new condition is the second constraint (4.7), under an identification of the "conjugate" variables Π (as above, the Π are not quite conjugate to W) with infinitesimal Lorentz generators. We will work both in tensor and component form in showing this result.

The variation of $\Omega_1 - \Omega_2$ is defined as

$$\begin{aligned}\Delta &= \frac{\delta(\Omega_1-\Omega_2)}{\delta W_{CA}}\delta W_{CA} + \frac{\delta(\Omega_1-\Omega_2)}{\delta W_{DB}}\delta W_{DB} \\ &\quad + \frac{\delta(\Omega_1-\Omega_2)}{\delta W_{GE}}\delta W_{GE} + \frac{\delta(\Omega_1-\Omega_2)}{\delta W_{HF}}\delta W_{HF}\end{aligned}$$

where

$$\frac{\delta W^a_{\ b}}{\delta W^x_{\ y}} = \delta^a_x \delta^y_b$$

$$\frac{\delta W_b^{\ a}}{\delta W^x_{\ y}} = -W_p^{\ a} W_b^{\ q} \delta^p_x \delta^y_q \ ,$$

the latter coming from $\delta(W\,W^{-1}) = 0$. In component form, the constraint reads

$$\begin{aligned}0 &= W_{CA}{}^a{}_b W_{DB}{}_c{}^b W_{CA}{}_d{}^c W_{DB}{}^d{}_e \\ &\quad - W_{HF}{}_b{}^a W_{GE}{}^b{}_c W_{HF}{}^c{}_d W_{GE}{}_e{}^d \ .\end{aligned}$$

The variation, with the δ-functions evaluated, is

$$\begin{aligned}\Delta^a{}_e &= (\delta W_{CA}{}^a{}_b) W_{DB}{}_c{}^b W_{CA}{}_d{}^c W_{DB}{}^d{}_e - W_{CA}{}^a{}_b W_{DB}{}_c{}^b W_{CA}{}_p{}^c W_{CA}{}_d{}^q (\delta W_{CA}{}^p{}_q) W_{DB}{}^d{}_e \\ &\quad - W_{CA}{}^a{}_b W_{DB}{}_p{}^b W_{DB}{}_c{}^q (\delta W_{DB}{}^p{}_q) W_{CA}{}_d{}^c W_{DB}{}^d{}_e + W_{CA}{}^a{}_b W_{DB}{}_c{}^b W_{CA}{}_d{}^c (\delta W_{DB}{}^d{}_e) \\ &\quad - W_{HF}{}_b{}^a (\delta W_{GE}{}^b{}_c) W_{HF}{}^c{}_d W_{GE}{}_e{}^d + W_{HF}{}_b{}^a W_{GE}{}^b{}_c W_{HF}{}^c{}_d W_{GE}{}_p{}^d W_{GE}{}_e{}^q (\delta W_{GE}{}^p{}_q) \\ &\quad + W_{HF}{}_p{}^a W_{HF}{}_b{}^q (\delta W_{HF}{}^p{}_q) W_{GE}{}^b{}_c W_{HF}{}^c{}_d W_{GE}{}_e{}^d - W_{HF}{}_b{}^a W_{GE}{}^b{}_c (\delta W_{HF}{}^c{}_d) W_{GE}{}_e{}^d \ .\end{aligned}$$

This is merely the $(\ )^a{}_e$ element of the matrix

$$\begin{aligned}\Delta &= (\delta W_{CA}) W_{DB}^{-1} W_{CA}^{-1} W_{DB} - W_{CA} W_{DB}^{-1} W_{CA}^{-1} (\delta W_{CA}) W_{CA}^{-1} W_{DB} \\ &\quad - W_{CA} W_{DB}^{-1} (\delta W_{DB}) W_{DB}^{-1} W_{CA}^{-1} W_{DB} + W_{CA} W_{DB}^{-1} W_{CA}^{-1} (\delta W_{DB}) \\ &\quad - W_{HF}^{-1} (\delta W_{GE}) W_{HF} W_{GE}^{-1} + W_{HF}^{-1} W_{GE} W_{HF} W_{GE}^{-1} (\delta W_{GE}) W_{GE}^{-1} \\ &\quad + W_{HF}^{-1} (\delta W_{HF}) W_{HF}^{-1} W_{GE} W_{HF} W_{GE}^{-1} - W_{HF}^{-1} W_{GE} (\delta W_{HF}) W_{GE}^{-1} \ ,\end{aligned}$$

which is clearly the variation of $W_{CA} W_{DB}^{-1} W_{CA}^{-1} W_{DB} \ - \ W_{HF}^{-1} W_{GE} W_{HF} W_{GE}^{-1}$ under $\delta(W^{-1}) = -W^{-1}(\delta W) W^{-1}$.

Recall that $(\delta W)W^{-1}$ is antisymmetric (3.7). Insert factors of $\mathbf{1} = W^{-1}W$ where necessary into the expression for Δ to get all δW into this antisymmetric form. As the variation Δ must vanish, we can right-multiply Δ by $W_{DB}^{-1}W_{CA}W_{DB}W_{CA}^{-1}$, or equivalently $W_{GE}W_{HF}^{-1}W_{GE}^{-1}W_{HF}$, to find the following:

$$\begin{aligned}
0 \;=\;& (\delta W_{CA}\, W_{CA}^{-1}) - W_{CA}W_{DB}^{-1}W_{CA}^{-1}(\delta W_{CA}\, W_{CA}^{-1})W_{CA}W_{DB}^{-1}W_{CA}^{-1} \\
& - W_{CA}W_{DB}^{-1}(\delta W_{DB}\, W_{DB}^{-1})W_{DB}W_{CA}^{-1} + W_{CA}W_{DB}^{-1}W_{CA}^{-1}(\delta W_{DB}\, W_{DB}^{-1})W_{CA}W_{DB}W_{CA}^{-1} \\
& - W_{HF}^{-1}(\delta W_{GE}\, W_{GE}^{-1})W_{HF} + W_{HF}^{-1}W_{GE}W_{HF}W_{GE}^{-1}(\delta W_{GE}\, W_{GE}^{-1})W_{GE}W_{HF}^{-1}W_{GE}^{-1}W_{HF} \\
& + W_{HF}^{-1}(\delta W_{HF}\, W_{HF}^{-1})W_{HF} - W_{HF}^{-1}W_{GE}(\delta W_{HF}\, W_{HF}^{-1})W_{GE}^{-1}W_{HF}\;.
\end{aligned}$$

As $(\delta W)W^{-1}$ is antisymmetric, it has only three independent components and can be replaced by a vector λ^a:

$$\lambda^a = \epsilon^a{}_{bc}\, \delta W^b{}_d\, W^{dc}.$$

Again, the similarity transformations $\mathcal{U}(\delta W\, W^{-1})\mathcal{U}^{-1}$ in the variation are simple linear transformations $\mathcal{U}\lambda$ of the vector λ. The variation of the constraint becomes

$$\begin{aligned}
0 \;=\;& \lambda_{CA} - W_{CA}W_{DB}^{-1}W_{CA}^{-1}\lambda_{CA} \\
& - W_{CA}W_{DB}^{-1}\lambda_{DB} + W_{CA}W_{DB}^{-1}W_{CA}^{-1}\lambda_{DB} \\
& - W_{HF}^{-1}\lambda_{GE} + W_{HF}^{-1}W_{GE}W_{HF}^{-1}W_{GE}^{-1}\lambda_{GE} \\
& + W_{HF}^{-1}\lambda_{HF} - W_{HF}^{-1}W_{GE}\lambda_{HF}\;.
\end{aligned}$$

Under the identification

$$\lambda_{XY}{}^a \longleftrightarrow \Pi_{XY}{}^a\;,$$

this is exactly the second constraint (4.7). The same notion of the Π being momenta conjugate to the configuration variables W is suggested. We offer an interpretation of these "canonical variables" in Chapter 5.

We arrive at the question of how many degrees of freedom are directly removed by these constraints. The condition $\Omega_1\Omega_2^{-1} = \mathbf{1}$ implies

$$W_{CA} W_{DB}^{-1} W_{CA}^{-1} = \Omega_2 W_{DB}^{-1} \ . \tag{4.9}$$

Taking the trace of this matrix equation shows

$$\mathrm{Tr} W_{DB}^{-1} = \mathrm{Tr}\,(W_{CA} W_{DB}^{-1} W_{CA}^{-1}) = \mathrm{Tr}\,(\Omega_2 W_{DB}^{-1})$$

Whatever form Ω_2 takes, the components of W_{DB} must satisfy this (scalar) equation, removing one df.

Using the fact that the Lorentz transformations form a Lie group, we can write each element in exponential form $\mathcal{U} = \mathrm{e}^{\eta^a J_a}$, where the generators J_a introduced in §1.2 are a basis for this vector space. We re-write (4.9) as

$$\mathrm{e}^{-\eta_{DB}{}^a W_{CA} J_a W_{CA}^{-1}} = \Omega_2\, \mathrm{e}^{-\eta_{DB}{}^a J_a} \ .$$

Now $-\eta_{DB}{}^a W_{CA} J_a W_{CA}^{-1}$ is a vector in the space spanned by the basis vectors $W_{CA} J_a W_{CA}^{-1}$, the original basis J_a rotated by W_{CA}. In the original, non-rotated basis, this is the vector $-(W_{CA}\eta_{DB})^a J_a$. This is the same as the earlier result of §3.2 that showed $\lambda' = \mathcal{U}\lambda$ when the matrices $\delta W\, W^{-1}$ are rotated by $\mathcal{U}$. Thus we have

$$\mathrm{e}^{-(W_{CA}\eta_{DB})^a J_a} = \Omega_2\, \mathrm{e}^{-\eta^a_{DB} J_a} \ . \tag{4.10}$$

The right-hand side is W_{DB}^{-1} transformed by Ω_2. The vector representing this new Lorentz transformation is $-(W_{CA}\eta_{DB})^a$, a rotation of η^a_{DB}. The equation specifies the direction about which this rotation must occur (two equations) and its magnitude (one equation). We see, however,

$$\begin{aligned} W_{CA} W_{DB}^{-1} W_{CA}^{-1} &= W_{CA}(W_{DB} W_{DB}^{-1}) W_{DB}^{-1} W_{CA}^{-1} \\ &= (W_{CA} W_{DB}) W_{DB}^{-1} (W_{CA} W_{DB})^{-1} \end{aligned}$$

so that rotations about the η_{DB}-direction are inconsequential. Therefore, the rotation specified by (4.10) is determined by only one parameter. Together with the magnitude of the rotation, two df are removed. Coupled with the trace relation, a full three df are directly removed by the relation $\Omega_1\Omega_2^{-1} = \mathbf{1}$. The translation constraint (4.7) likewise removes three df, directly reducing the dimension of the phase space by six. The failure of the constraints to remove a full six df from the T^2 system is due to the triviality of the trace relation: The condition $W_{CA}W_{DB}^{-1}W_{CA}^{-1}W_{DB} = \mathbf{1}$ shows $W_{DB}^{-1} = W_{CA}W_{DB}^{-1}W_{CA}^{-1}$ so that

$$\begin{aligned}\mathrm{Tr}\, W_{DB}^{-1} &= \mathrm{Tr}\,(W_{CA}W_{DB}^{-1}W_{CA}^{-1}) \\ &= \mathrm{Tr}\, W_{DB}^{-1} \,.\end{aligned}$$

This equation put no conditions of the transformation W_{DB}. This, together with the corresponding translation "non-constraint", supplies the unexpected extra df in the torus system.

Each constraint also removes a gauge degree of freedom, that gauge transformation generated by the constraint. Thus the dimension of the phase space over all genus $g > 1$ surfaces is ($2g$ holonomies @ 6 df per holonomy) $-$ (6 constraints + 6 gauge choices), giving dimension $12g - 12$, except for the genus 1 torus, which has an unexpected 12−8=4 degrees of freedom.

4.3 A Model of $T^2\#T^2$

While the results we have found are quite explicit, they are still based on some unspecified octagonal tiling and an abstract collection of Lorentz transformations and translations. We now propose to build an actual physical model (as physical as (2+1) can be...) for the tiling and the holonomies. We will see, though, that to construct a model, we have to simplify the geometry with high symmetry, eliminating the translation components Π^a of the holonomies altogether.

4.3.1 Tilings

To construct a closed surface without boundary, like $T^2 \# T^2$, one abstractly thinks of a plane or similar infinite region modulo some identification. Concretely this can mean covering the infinite region with tiles and identifying sides of the tile in pairs. The plane $\mathbf{R}^2$, however, cannot be tiled with all regular polygons. Suppose the tiles are regular p-sided polygons. The interior angle at each vertex of a p-gon is $\frac{\pi(p-2)}{p}$. If q such tiles meet at every vertex then each tile contributes $\frac{2\pi}{q}$ so that $\frac{\pi(p-2)}{p} = \frac{2\pi}{q}$ or $(p-2)(q-2) = 4$. The solutions to this condition are $\{p = 3, q = 6\}$, $\{p = 6, q = 3\}$, and $\{p = 4, q = 4\}$ which correspond to covering the plane in triangles, hexagons and squares, respectively. The latter is the tiling we use to construct the torus T^2. Clearly there is no integral value of q for which $p = 8$ is a solution, meaning the $\mathbf{R}^2$ cannot be tiled in octagons.

Instead we look to a hyperbolic plane where the sum of the angles in a triangle is less than π, and the interior angle of regular p-gon is less than $\frac{\pi(p-2)}{p}$. The neighborhood of a vertex where q hyperbolic p-gons meet is a patch of $\mathbf{R}^2$ and still requires a full 2π radians. Each tile contributes an angle of $\frac{2\pi}{q}$. Since the p-gon is hyperbolic, $\frac{2\pi}{q} < \frac{\pi(p-2)}{p}$, or $(p-2)(q-2) > 4$. One of infinitely many solutions to this condition is $\{p = 8, q = 8\}$. While it is not clear how an octagon can regularly cover a hyperboloid, one must recall that the hypersurface of constant time is embedded in Minkowski, not Euclidean, space so that every point on the hypersurface is the same (proper) distance from the origin, much like the 2-dimensional surface of a sphere in Euclidean $\mathbf{R}^3$.

4.3.2 Construction of an Octagonal Tile

The constraints tell us ${R_{ij}}^{ab} = 0$, so the tile we construct must be flat. Consider the simplest flat 3-dimensional space, Minkowski. In polar coordinates, the 3-dimensional Minkowski metric is

$$ds^2 = -dt^2 + dR^2 + R^2 d\theta^2.$$

Change coordinates (t, R, θ) to (τ, ρ, θ) defined by

$$\begin{aligned} t &= \tau \cosh \rho \\ R &= \tau \sinh \rho. \end{aligned} \tag{4.11}$$

Inverting this transformation shows

$$\begin{aligned} \tau &= \pm\sqrt{t^2 - R^2} \\ \rho &= \tanh^{-1} \frac{R}{t}. \end{aligned}$$

Observe that τ is invariant under Lorentz transformations, while ρ is the magnitude of the Lorentz boost which takes $R = 0$ out to $R = \tau \sinh \rho$. We exploit the invariance of τ under Lorentz transformations. By building the tile on a surface of constant τ, points on the tile will be connected with merely Lorentz, rather than full Poincaré, transformations.

Finally, define a new coordinate

$$r = \tau \tanh(\frac{\rho}{2}) \tag{4.12}$$

and consider for simplicity the $\tau = 1$ hypersurface. This coordinate transformation projects the $\tau = 1$ hypersurface onto a unit disk with infinity at $r = 1$, much the same way the stereographic map projects $\mathbf{R}^2$ onto the 2-sphere. The metric on this surface is conformally flat:

$$d\sigma^2 = \frac{4}{(1 - r^2)^2}(dr^2 + r^2 d\sigma^2) .$$

Because this disk is conformally flat, angles are preserved between the unit disk and the (τ, ρ, θ) coordinate system.

To construct the tile, we piece together 8 identical curves chosen in the following way. In the original Minkowski space, consider the intersection of the $\tau = 1$ hyperboloid and the "vertical" plane $y = 0$. This curve lies on the hyperboloid "above" the y-axis and can be parametrized by

$$\gamma(\lambda) = \left(t(\lambda) = \lambda, R(\lambda) = \sqrt{\lambda^2 - 1}, \theta(\lambda) = \pm\frac{\pi}{2}\right)$$

for $\lambda \geq 1$. Now Lorentz-boost every point in the y=0 plane in the x-direction by some magnitude ρ; the points which lie on the y-axis are unaffected by this transformation and the $y = 0$ plane is 'tilted' in the x-direction. With increasing boosts the intersection curve $\gamma(\lambda)$ moves away from lying over the y-axis until finally with an infinite boost, the plane has tilted by $\pi/4$ and is just tangent to the hyperboloid at $x = \pm\infty$. With arbitrary but finite boosts by ρ, the family of curves $\gamma(\lambda; \rho)$ is parametrized by

$$\begin{aligned}
t(\lambda) &= \lambda \cosh \rho \\
R(\lambda) &= \sqrt{\lambda^2 \cosh^2 \rho - 1} \\
\theta(\lambda) &= \tan^{-1}\left(\frac{\pm\sqrt{\lambda^2 - 1}}{\lambda \sinh \rho}\right).
\end{aligned}$$

One can check that $y(\lambda) = R(\lambda) \sin \theta(\lambda)$ remains unchanged under the ρ-boost, so that the plane is tilted without any stretching.

On the unit disk with coordinates (r, θ) found by projecting down the $\tau = 1$ hypersurface, consider the collection of circles centred outside the disk which intersect $r = 1$ orthogonally (See Figure 4.3). The arcs within the unit disk are the geodesics of this Poincaré Disk model of hyperbolic geometry [8]. Parametrize in θ the arc within the unit disk of a circle of radius $\hat{r}$:

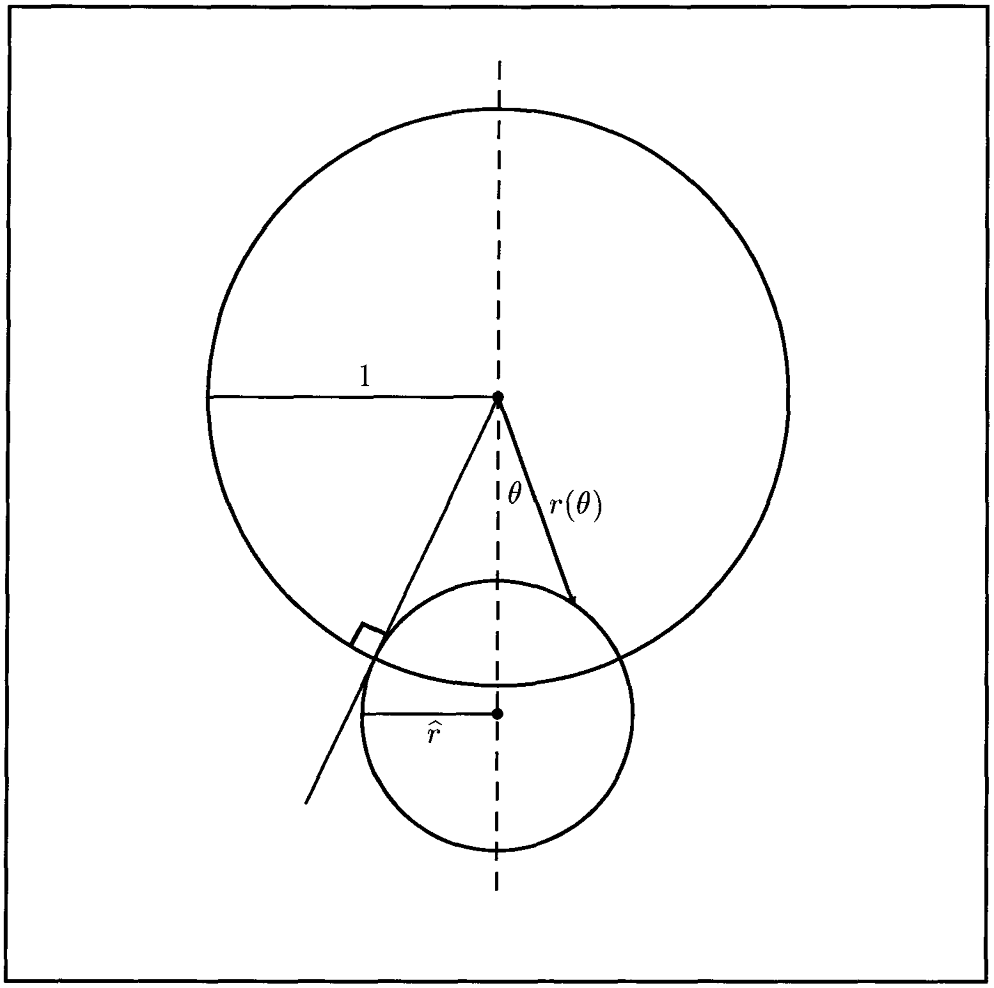

Figure 4.3: Circles of radius $\hat{r}$ intersect the unit disk orthogonally. The arcs within the unit disk are parametrized in θ.

$$\begin{aligned} r(\theta) &= \sqrt{1+\hat{r}^2}\cos\theta - \sqrt{(1+\hat{r}^2)\cos^2\theta - 1} \\ \tan\theta &\in [-\hat{r}, \hat{r}] \,. \end{aligned}$$

This circle is centered a distance $\sqrt{1+\hat{r}^2}+\hat{r} \geq 1$ from the origin.

Transforming this family of curves $r(\theta;\hat{r})$ with parameter $\hat{r}$ back to Minkowski coordinates, we find

$$\begin{aligned} t(\theta) &= \frac{1+r(\theta)^2}{1-r(\theta)^2} \\ R(\theta) &= \frac{2r(\theta)}{1-r(\theta)^2} \\ (\ \theta &= \theta\ ) \end{aligned}$$

Comparing this with the family of intersection curves $\gamma(\lambda;\rho)$ we see

$$t(\lambda)^2 - R(\lambda)^2 = 1 = t(\theta)^2 - R(\theta)^2$$

so that both families of curves lie on the $\tau = 1$ hypersurface. Furthermore, by comparing the $\theta = 0$ points of both families we find the correspondence between ρ and $\hat{r}$ and finally that these two families of curves are identical. That is, the curve in Minkowski space where the $y = 0$ plane, tilted by $\tanh(\rho)$, intersects the $\tau = 1$ hypersurface becomes the arc of a circle of radius $\hat{r}(\rho)$ which orthogonally intersects the boundary of the unit Poincaré Disk.

By simply rotating these curves about the origin, we can piece together arcs in the unit disk to form an 8-sided figure. We must now find the value of the parameter $\hat{r}$, or equivalently ρ, which produces the correct tile. The polygonal tile we are constructing is regular, so the 8-sides must be spaced at equal intervals of $\pi/4$. Consider the figure produced by laying down 8 arcs of radius $\hat{r}$ centered at radius $\sqrt{1+\hat{r}^2}+\hat{r}$ on the 8 $\pi/4$-'spokes' (See Figure 4.4). When $\hat{r} \sim 0$ the arcs belong to small circles centered just beyond $r = 1$ and the arcs do not intersect (Figure 4.4(i)). At some larger $\hat{r}$ when each

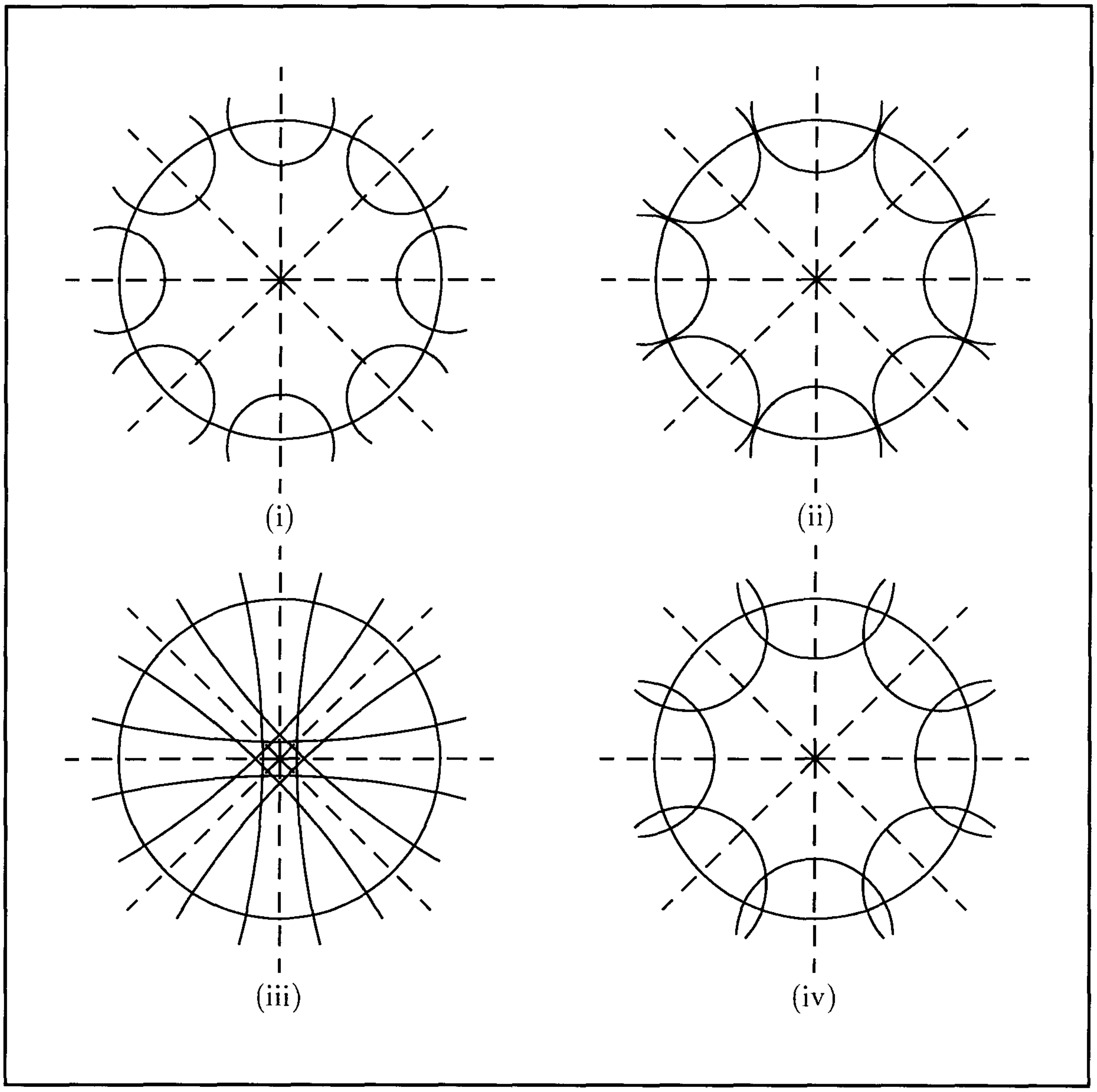

Figure 4.4: For various values of $\hat{r}$, arcs (i) do not intersect, (ii) are tangent to one another, (iii) form an almost regular octagon about the origin, and (iv) intersect with angle $\pi/4$.

arc just intersects the two neighboring arcs, the angle at the intersection of two adjacent arcs is 0 because all arcs intersect orthogonally with $r = 1$ (Figure 4.4(ii)). When $\hat{r} \to \infty$, the arcs become diameters of the unit circle, and at large $\hat{r}$, the 8 arcs intersect to form an 8-sided figure about the origin which is very nearly a regular octagon (Figure 4.4(iii)). The angle between adjacent arcs of this figure is almost $\frac{3\pi}{4}$, the interior angle of a regular plane octagon. For each value of $\hat{r}$ we amputate the legs of the 8-sided figure about the origin and call the result an octagon. We must choose the value of $\hat{r}$ which generates an octagon whose adjacent sides intersect at an angle $\frac{\pi}{4}$ (Figure 4.4(iv)) so that 8 such tiles will supply the 2π radians about the vertex. On the conformally flat Poincaré Disk, we can use plane geometry to find

$$\hat{r}\,(\frac{\pi}{4}) = \sqrt{\frac{1}{2+2\sqrt{2}}} \quad . \tag{4.13}$$

This $\hat{r}$ corresponds to a boost magnitude of

$$\rho(\frac{\pi}{4}) = \ln\left[\frac{1+r}{1-r}\right] = \ln\left[\frac{1+\sqrt{1+\hat{r}(\frac{\pi}{4})^2} - \hat{r}(\frac{\pi}{4})}{1-\sqrt{1+\hat{r}(\frac{\pi}{4})^2} + \hat{r}(\frac{\pi}{4})}\right] \quad . \tag{4.14}$$

The magnitude $\rho(\frac{\pi}{4})$ generates a tile on the $\tau = 1$ hypersurface. The construction can be repeated for hypersurfaces at arbitrary τ, but the result is *the same.* This magnitude is actually independent of τ and generates curves on all hypersurfaces of constant τ from which these octagonal tiles can be constructed.

4.3.3 Holonomies

The 8-sided figure on the hypersurface of constant time τ is the tile we represent schematically in Figure 4.1(i). For simplicity, suppose the centre of the octagon lies over the origin and that the $\theta = 0$ ray bisects side A. Corner 1 lies at $\theta = -\frac{\pi}{8}$, corner 2 at $\theta = \frac{\pi}{8}$, corner 3 at $\theta = \frac{3\pi}{8}$, and so on (See Figure 4.5). The transformations W between the identified sides can be easily found by recalling the procedure used to construct the curves which

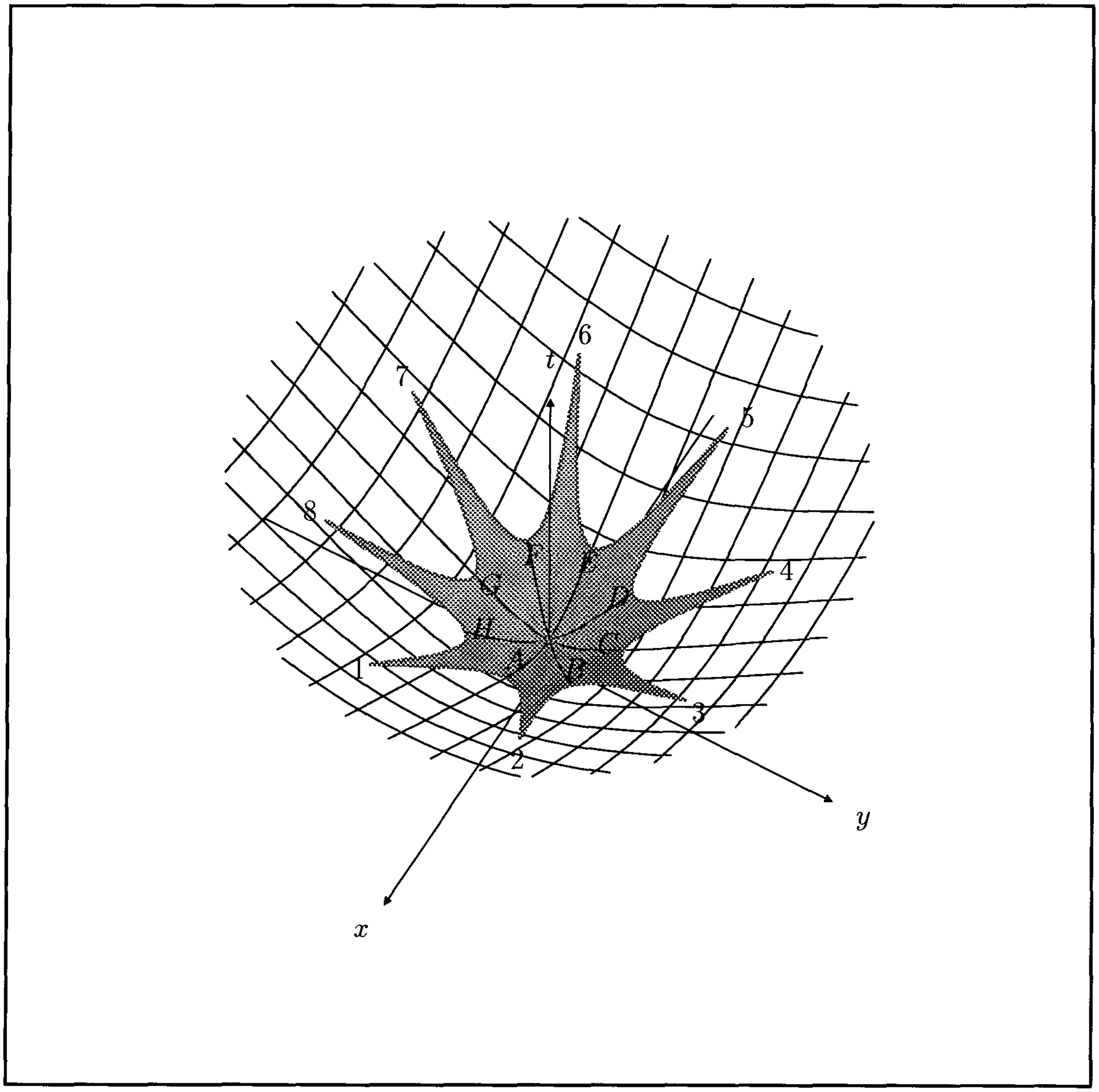

Figure 4.5: The octagonal tile lying on a hyperbolic hypersurface of constant time τ. In Minkowski space, every point on the tile is the same distance τ from the origin.

became the sides of the octagon. Hereon, ρ will refer to the value $\rho(\frac{\pi}{4})$ produced from $\hat{r}(\frac{\pi}{4})$. Just as a boost in the y-direction by ρ drops the curve over the x-axis down to form side C of the tile, a boost in the $(-y)$-direction by ρ (or equivalently a boost in the y-direction by $-\rho$) will lift side C back up to a curve lying over the x-axis. Rotate this curve by $+\frac{\pi}{2}$ (+ to generate the right orientation), and boost it by ρ in the x-direction to drop it back down onto side A. This sequence of SO(2,1) transformations maps side C onto side A. As SO(2,1) is a group, the composition of the three is a single Lorentz transformation, which we call W_{CA}.

A general Lorentz boost of magnitude μ in the ϕ-direction is $\vec{x} \to \Lambda(\mu,\phi)\vec{x}$ where

$$\Lambda(\mu,\phi) = \begin{pmatrix} \cosh\mu & \sinh\mu\,\cos\phi & \sinh\mu\,\sin\phi \\ \sinh\mu\,\cos\phi & (\cosh\mu-1)\cos^2\phi+1 & (\cosh\mu-1)\cos\phi\,\sin\phi \\ \sinh\mu\,\cos\phi & (\cosh\mu-1)\cos\phi\,\sin\phi & (\cosh\mu-1)\sin^2\phi+1 \end{pmatrix}.$$

Rotations about the origin in the xy-plane by angle ψ are produced under the transformation

$$R(\psi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\psi & -\sin\psi \\ 0 & -\sin\psi & \cos\psi \end{pmatrix}.$$

The transformations W_{CA}, then, is given by

$$\begin{aligned} W_{CA} &: \text{maps side } C \text{ onto side } A \\ W_{CA} &= \Lambda(\rho,0)R(\frac{\pi}{2})\Lambda(\rho,\frac{3\pi}{2}) \\ &= \begin{pmatrix} \cosh 2\rho & 0 & -\sinh 2\rho \\ \sinh 2\rho & 0 & -\cosh 2\rho \\ 0 & 1 & 0 \end{pmatrix}. \end{aligned}$$

This method of lift-rotate-drop gives each of the W transformations:

$$
\begin{aligned}
W_{DB} &: \text{ maps side } D \text{ onto side } B \\
W_{DB} &= \Lambda(\rho, \frac{\pi}{4}) R(\frac{\pi}{2}) \Lambda(\rho, -\frac{\pi}{4}) \\
&= \begin{pmatrix} \cosh 2\rho & \frac{1}{\sqrt{2}} \sinh 2\rho & -\frac{1}{\sqrt{2}} \sinh 2\rho \\ \frac{1}{\sqrt{2}} \sinh 2\rho & \sinh^2 \rho & -\cosh^2 \rho \\ \frac{1}{\sqrt{2}} \sinh 2\rho & \cosh^2 \rho & -\sinh^2 \rho \end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
W_{GE} &: \text{ maps side } G \text{ onto side } E \\
W_{GE} &= \Lambda(\rho, \pi) R(\frac{\pi}{2}) \Lambda(\rho, \frac{\pi}{2}) \\
&= \begin{pmatrix} \cosh 2\rho & 0 & \sinh 2\rho \\ -\sinh 2\rho & 0 & -\cosh 2\rho \\ 0 & 1 & 0 \end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
W_{HF} &: \text{ maps side } H \text{ onto side } F \\
W_{HF} &= \Lambda(\rho, \frac{5\pi}{4}) R(\frac{\pi}{2}) \Lambda(\rho, \frac{3\pi}{4}) \\
&= \begin{pmatrix} \cosh 2\rho & -\frac{1}{\sqrt{2}} \sinh 2\rho & \frac{1}{\sqrt{2}} \sinh 2\rho \\ -\frac{1}{\sqrt{2}} \sinh 2\rho & \sinh^2 \rho & -\cosh^2 \rho \\ -\frac{1}{\sqrt{2}} \sinh 2\rho & \cosh^2 \rho & -\sinh^2 \rho \end{pmatrix}
\end{aligned}
$$

4.3.4 The Fields $\mathcal{U}^a{}_b$ and ρ^a

The symmetry of the model for $T^2 \# T^2$ allows us to find a single transformation W_{CA} that maps side C onto side A. Because W_{CA} is independent of the pair of identified points on the matching sides, we see $W_{CA,i} = 0$ as required by (2.5). The same applies to W_{DB}, W_{GE}, and W_{HF}. The W transformations, recall, are the link between the values of the field $\mathcal{U}$ on opposites sides of a boundary. We attach the W onto the octagon as we did for the square tile of T^2, by regarding the holonomy as a "phase difference" between identified points.

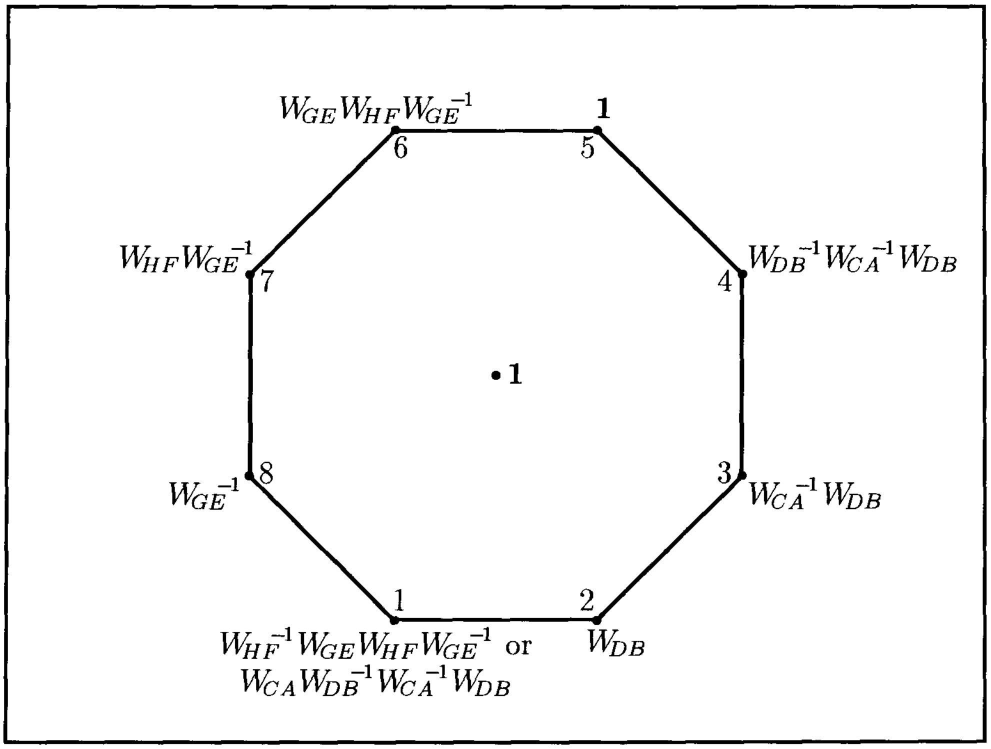

Figure 4.6: Values of the phase $\mathcal{U}$ on the corners of the $T^2\#T^2$'s octagonal tile. The constraints guarantee $\mathcal{U}(1)$ is well defined.

Again we start on the corners of the tile (Refer to Figure 4.1(i)). Since only the phase difference between identified points matters, there is an arbitrary constant phase. If we set $\mathcal{U}(3) = W_{CA}^{-1}W_{DB}$, the $\mathcal{U}$ field nicely decouples into $ABCD$ and $EFGH$, as we found before (See Figure 4.6). By starting on corner 3 and transforming to each corner under the W's, we find on corner 1 both $\mathcal{U}(1) = W_{HF}^{-1}W_{GE}W_{HF}W_{GE}^{-1}$ and $\mathcal{U}(1) = W_{CA}W_{DB}^{-1}W_{CA}^{-1}W_{DB}$. The constraint (4.5) ensures $\mathcal{U}(1)$ is well defined. It is now a simple matter of interpolating these vertex values over the whole tile, while preserving the phase difference between identified points. A simple choice is a linear interpolation between the corners along the boundary, coupled with a radial interpolation under which the $\mathcal{U}$ field, now defined on the boundary, decays down to the identity at the center of the tile. All the transformations W are in SO(2,1) and connected continuously to the identity. These two interpolations together give a continuous field over the tile with points on the boundary differing by the appropriate transformations W. Smoother interpolations can be used, if necessary.

The last component of this model is the field ρ^a, introduced to write the *dreibein* as $e_\mu{}^a = D_\mu \rho^a$. The choice we make is a very simple one; too simple, perhaps, for it sets the translations Π to 0. Consider the choice $\rho^a = x^a = (t, x, y)$ in Minkowski space, giving $e_x = (0,1,0)$ and $e_y = (0,0,1)$. Recall the solutions to the equations of motion are those for which $R_{\mu\nu}{}^{ab} = 0$, so in the solution space we can take $e_t = (1,0,0)$ as well. The two fields $\mathcal{U}$ and ρ completely determine the holonomies. We found in (2.8)

$$W_{21}{}^a{}_b \rho_2{}^a + \Pi_{21}{}^a = \rho_1{}^a$$

where ρ_1 and ρ_2 are the values of ρ on either side of the identified boundary. The transformations W_{CA}, for instance, maps (t, x, y) on side C onto (t, x, y) on side A:

$$\begin{pmatrix} t_A \\ x_A \\ y_A \end{pmatrix} = \begin{pmatrix} & & \\ & W_{CA} & \\ & & \end{pmatrix} \begin{pmatrix} t_C \\ x_C \\ y_C \end{pmatrix} .$$

But this is exactly

$$\rho_A{}^a = W_{CA}{}^a{}_b \rho_C{}^b ,$$

showing the translation $\Pi_{CA}{}^a$ is not needed to reproduce the discontinuity in ρ across the CA-boundary. This choice of ρ^a is "elegant" because it so easily exhibits the discontinuities required for the holonomies to be non-trivial. The solution $\Pi = 0$ just means the configuration is (momentarily) stationary. This is consistent with the geometry of the tile being independent of the time τ on the spatial slice, a property we found in (4.14) when we built the tile.

Chapter 5

Canonically Conjugate Variables and Quantization

We have seen from the two surfaces studied here that the Einstein action, on the solutions to the equations of motion and with the constraints imposed, is of the form

$$\begin{aligned} I_{T^2} &= (W_{DB}^{-1}\Pi_{DB})\{\dot{W_{CA}}\} - (W_{CA}^{-1}\Pi_{CA})\{\dot{W_{DB}}\} \\ I_{T^2\#T^2} &= \left(W_{DB}^{-1}\Pi_{CA}\right)\{\dot{\overline{W_{DB}^{-1}\Omega_1}}\} - \left(W_{DB}^{-1}\Pi_{DB}\right)\{\dot{\overline{W_{CA}^{-1}\Omega_1}}\} \\ &\quad -\Pi_{GE}\{\dot{\overline{W_{HF}\Omega_2}}\} + \left(W_{GE}W_{HF}^{-1}\Pi_{HF}\right)\{\dot{\overline{W_{GE}\Omega_2}}\} \end{aligned}$$

This result is not the canonical $p\dot{q}$ form of the action because $\{\dot{W}\}$ is not $\frac{d}{dt}W$ but instead, recall,

$$\{\dot{W}\}_a = \dot{W}^b{}_x W^{cx}\epsilon_{abc}\ .$$

While this is not the vector $\frac{d}{dt}W$ tangent to the space of W's, $\{\dot{W}\}$ is still tangent to some W-space, related in a one-to-one way with the tangents of the space of Lorentz transformations. The translations Π are not (quite) the momentum conjugate to the Lorentz transformations W, although the association or the Π with infinitessimal translations of the W is very suggestive. It is possible to bring the result even closer to the canonical form to reveal more about the phase space.

Every Lorentz transformation can be written in the form

$$W = \mathrm{e}^{\eta^a J_a}$$

where the J_a are the 3×3 Lorentz generators introduced in §1.2. The components η^a can be determined as follows. Write $W = \mathrm{e}^{\mathcal{A}}$ for some $\mathcal{A}$. As $\mathcal{B}\mathrm{e}^{\mathcal{A}}\mathcal{B}^{-1} = \mathrm{e}^{\mathcal{B}\mathcal{A}\mathcal{B}^{-1}}$,

$$\begin{aligned} \mathrm{e}^{\mathcal{A}} &= W \\ &= W\,W\,W^{-1} \\ &= W\mathrm{e}^{\mathcal{A}}W^{-1} \\ &= \mathrm{e}^{W\mathcal{A}W^{-1}} \end{aligned}$$

so that $\mathcal{A} = W\mathcal{A}W^{-1}$ or $[\mathcal{A}, W] = 0$. If we write $\mathcal{A} = \eta^a J_a$, then

$$0 = \eta^a [J_a, W] \ .$$

There are only three distinct components in this matrix equation due to the symmetry of the J_a. The three equations are not linearly independent, though (the right-hand-side has vanishing determinant, as $\det(J_a)$=0). Another relation is needed to solve for the components η^a. Determinants will not suffice: As $\det(W)$=1

$$1 = \det(W) = \mathrm{e}^{\mathrm{Tr}\mathcal{A}} \ .$$

Thus $\mathrm{Tr}\mathcal{A}$=0 and any constant times the matrix $\mathcal{A}$ will not change the determinant of W. Instead, we consider

$$\mathrm{Tr}\, W = \mathrm{Tr}(\mathrm{e}^{\mathcal{A}}) = \mathrm{e}^{\lambda_1} + \mathrm{e}^{\lambda_2} + \mathrm{e}^{\lambda_3}$$

where $\lambda_1, \lambda_2, \lambda_3$, the eigenvalues of $\mathcal{A}$, are functions of η^a. This comes from the property that $\mathcal{A}$ can be written $\mathcal{D}\Lambda\mathcal{D}^{-1}$ for some $\mathcal{D}$ and diagonal Λ, so that $\mathrm{Tr}\,\mathrm{e}^{\mathcal{A}} = \mathrm{Tr}(\,\mathrm{e}^{\mathcal{D}\Lambda\mathcal{D}^{-1}}) = \mathrm{Tr}(\,\mathcal{D}\mathrm{e}^{\Lambda}\mathcal{D}^{-1}) = \mathrm{Tr}\ \mathrm{e}^{\Lambda}$ by the cyclic nature of the trace. Since Λ is diagonal, $\mathrm{Tr}\ \mathrm{e}^{\Lambda} = \mathrm{e}^{\lambda_1} + \mathrm{e}^{\lambda_2} + \mathrm{e}^{\lambda_3}$. This relation, along with the two equations arising from $[\mathcal{A}, W] = 0$, suffice to completely determine η^a.

Now we can write

$$\dot{W} = \frac{d}{dt}\left(\mathrm{e}^{\eta^a J_a}\right)$$

$$\begin{aligned} &= (\frac{\partial W}{\partial \eta^a})\,\dot{\eta}^a \quad \Longrightarrow \\ \dot{W}^b{}_x &= (\frac{\partial W}{\partial \eta^a})^b{}_x\,\dot{\eta}^a \,. \end{aligned}$$

Therefore the action can be re-written as

$$\sum_{\text{holonomies }[W|\Pi]} \left(\Pi^a (\frac{\partial W}{\partial \eta^z})^b{}_x\, W^{cx} \epsilon_{abc} \right) \dot{\eta}^z \,.$$

This action is now in canonical $p\dot{q}$-form. This expression, however, is not as well behaved as one would like. In a neighborhood of each W in the space of Lorentz transformations, we can attach η coordinates and consider $(\partial W/\partial \eta^z)$. Writing $W = \mathrm{e}^{\eta^a J_a}$ is only shorthand for the power series $\mathbf{1} + \eta^a J_a + \frac{1}{2}[\eta^a J_a, \eta^b J_b] + \cdots$, so derivatives with respect to η^z may not even exist. Furthermore, there is no guarantee that the coordinate patches around each W are part of some global coordinates over which we can compare the values of $(\partial W/\partial \eta^z)$ for two different holonomies.

With these problems in mind, let us speculate on the canonical coordinates. The configuration space variables are η^a, the coordinates in the space of Lorentz transformations. This space is now represented as a vector space with basis J_a. The relations $[J_a, J_b] = \epsilon_{ab}{}^c J_c$ show the structure constants are $\epsilon_{ab}{}^c$, suggesting a non-trivial geometry. The momenta conjugate to these coordinates are related to the translations Π^a. Under this association, we would expect the Π^a, as momenta, to generate translations in the configuration space. And this is what we find in looking at the equivalence of the constraints: we make the association $\lambda^a \leftrightarrow \rho^a$ for $\lambda^a = \epsilon^a{}_{bc}\, \delta W^b{}_d\, W^{dc}$. It is easy to verify that $\epsilon_a{}^b{}_c = (J_a)^b{}_c$ — these λ^a are (infinitesimal) translations about the vector space of W's. While the interpretation is by no means rigorous, it is very suggestive: The Lorentz components W of the holonomies are the configuration coordinates. Translations in this space, generated by the momenta $\Pi^a \sim (\delta W\, W) J^a$, are infinitesimal Lorentz transformations. This is just what we expect for translation in a space of Lorentz transformations.

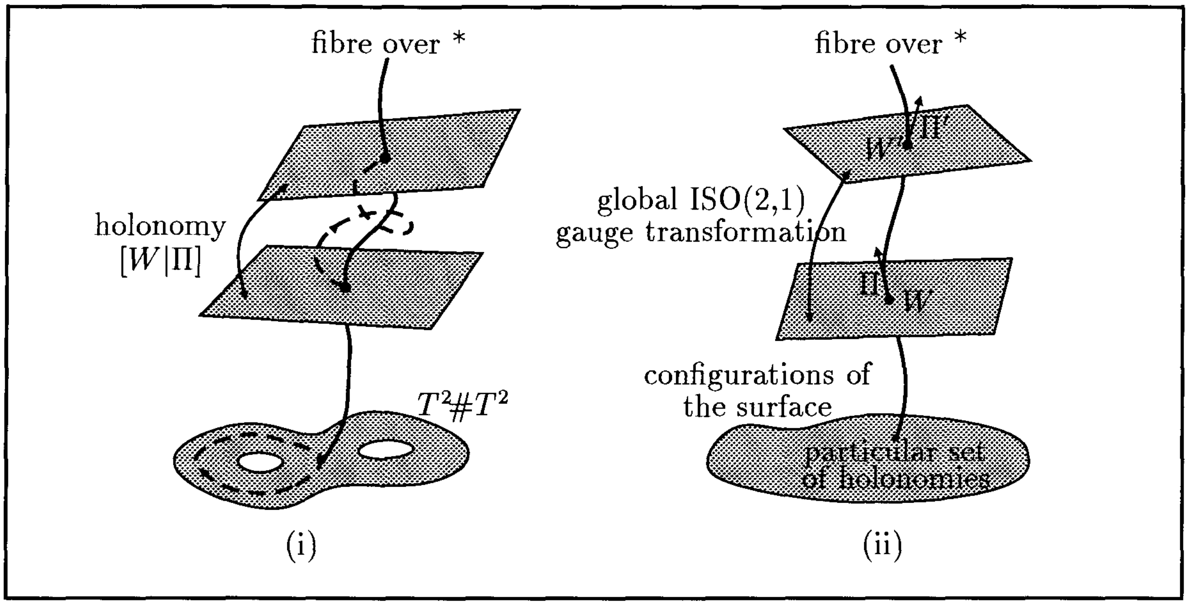

Figure 5.1: (i) A principle ISO(2,1)-bundle over the base space $T^2\#T^2$. (ii) The 'phase space' is like an ISO(2,1)-bundle over the space of configurations of the surface.

There is another, more mathematical description of this system. Vectors in the tangent space over each point of the surface, T^2 or $T^2\#T^2$, are subject to the action of the Poincaré group ISO(2,1). This group leaves the Einstein action invariant, and its generators obey a Lie algebra. These are the ingredients needed to define a principle ISO(2,1)-bundle over the surface. The group action moves us along each fibre over the surface without changing the Einstein action. The holonomy at a point on one of the incontractible loops is the element $[W|\Pi]$ of the gauge group ISO(2,1) relating two distinct points in the fibre over this base point (See Figure 5.1(i)). In the language of principle bundles, the phase space is also like an ISO(2,1)-bundle, this one over a base space consisting of configurations of the 2-dimensional surface (See Figure 5.1(ii)). The one point in the base space we have found represents the regular octagonal tile constructed above. Other points represent asymmetric tiles and their corresponding collections of holonomies. The momenta Π^a are infinitesimal generators of the group action in each fibre. The W and Π vary as we move about in each fibre from one horizontal lift to the next under global gauge transformations. Yet the projections down to the base space of holonomies remains unchanged. If we rotate the tile, or globally boost it to a new location, a similarly transformed collection of holonomies is produced. The tile representing this new surface, though, is essentially unchanged, merely displaced. The canonical variables W and Π are phase space coordinates in this bundle: The W are configuration coordinates in the fibre over a basepoint, a particular configuration of the tile representing the surface. The momenta Π generate translations along this fibre under the infinitesimal group action. Different horizontal lifts, all differing by global gauge transformations, have different coordinates (W, Π), but project down to the same model of the surface, perhaps displaced but leaving the Einstein action invariant.

In the full ISO(2,1) representation, the Poincaré holonomies can be written as

$$[W(\eta^a)|\Pi(\sigma^a)] = \mathrm{e}^{\eta^a J_a + \sigma^a P_a}$$

where P_a, introduced with J_a in §1.2, generates translation in the x^a direction. In the case of the torus T^2, we can see ${\eta_{CA}}^a = (0, 0, \mu)$, ${\sigma_{CA}}^a = (0, 0, a)$ and ${\eta_{DB}}^a = (0, 0, \alpha\mu)$, ${\sigma_{DB}}^a = (0, 0, b)$. For the two-holed torus $T^2 \# T^2$, the transformations $W_{CA}, \ldots, W_{HF}$ are compositions of boosts and rotations and the corresponding η^a must be evaluated by the method outlined above. The conjugate translations $\Pi_{CA}, \ldots, \Pi_{HF}$ all vanish so ${\sigma_{CA}}^a = (0, 0, 0), \ldots,$ ${\sigma_{HF}}^a = (0, 0, 0)$. This exponential ISO(2,1) representation of the holonomies matches the Chern-Simons approach. There, the gauge field is the collection of flat connections A_μ that transform under the action of an infinitesimal parameter $u = \tau^a J_a + \rho^a P_a$. That is, the difference between two horizontal lifts along the same fibre, (W, Π) and (W', Π'), which project down to the same model of the surface, is generated by an infinitesimal Poincaré transformation u. The coordinates W and Π determine the point on the fibre over the base space.

When the canonically conjugate variables of a classical Hamiltonian system are known, quantization in the Schrodinger picture involves expressing the variables as operators on a Hilbert space of wave functions Ψ. The observed values of the coordinates and momenta are eigenvalues of coordinate and momentum operators. Unfortunately our choice of canonical variables W and Π have ill-defined operators, for the following reason. We suggest that the configuration space has coordinates η^a, the components of the vector representing W in the vector space spanned by $\{J_1, J_2, J_3\}$. The quantum mechanical η^a are the spectrum of a position operator N^a:

$$N^a \Psi = \eta^a \Psi \ .$$

While J_3 and J_2 generate boosts in the x- and $(-y)$-directions, respectively, and have spectra $\eta^3, \eta^2 \in (-\infty, \infty)$, recall that J_1 generates rotations in the xy-plane. The eigenvalue of this operator is the angle of rotation, $\eta^1 \sim \theta$:

$$N^1(\theta)\, \Psi = \theta \Psi \ .$$

Because rotations differing by 2π give the same reading θ, we require both

$$N^1(\theta)\, \Psi = \theta \Psi \ ,$$

$$N^1(\theta + 2\pi)\, \Psi = \theta \Psi \ .$$

Thus N^1 is no longer a linear operator on the Hilbert space of Ψ's. This tells us that θ, or η^1, by itself cannot be an observable. Instead, some function of the operator N^1, e^{iN^1} for instance, is needed for a well-defined operator. At the same time, the Hilbert space is no longer a vector space, but has a "cylindrical" shape (See Figure 5.2).

Because the rotations are a subgroup of SO(2,1), which is itself a subgroups of the gauge group ISO(2,1), perhaps this problem can be circumvented by re-defining the configuration space modulo S^1: $\mathbf{R}^3/_{S^1} \otimes \mathbf{R}^3/_{S^1} \otimes \cdots \otimes \mathbf{R}^3/_{S^1}$, one term for each of the

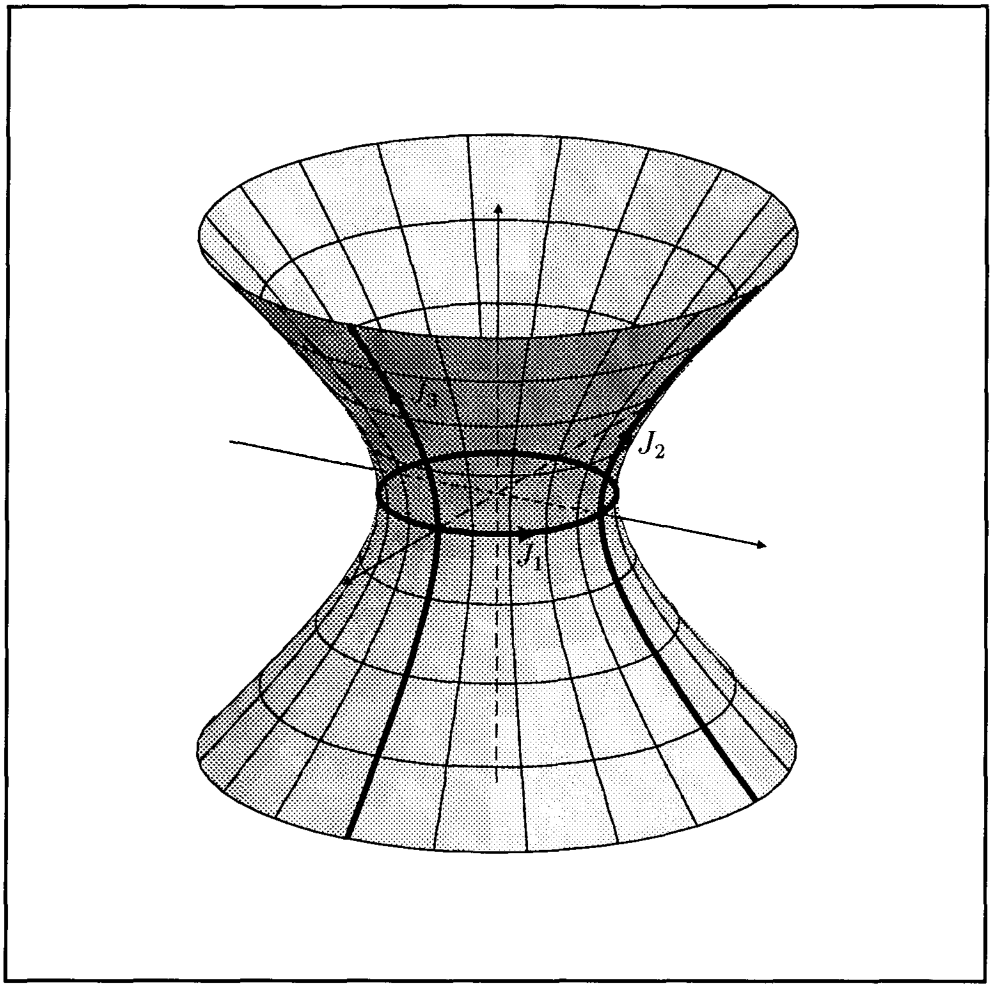

Figure 5.2: J_2 and J_3 are generators of Lorentz boosts while J_1 generates rotations. Because the same eigenvalue θ results from rotations by $\theta + 2\pi n$ for all integers n, the coordinate $\eta^1 \sim \theta$ is periodic and the Hilbert space of wave functions has a "cylindrical" geometry.

$2g$ holonomies of the flat genus g surface. In any event, this complex structure of the phase space arises from the choice of W and Π as canonically conjugate variables, and this interpretation is only speculative, based on qualitative observations.

Chapter 6

Conclusions

6.1 Generalization of the Result

The ease with which we jumped from the torus to the two-holed torus suggests this formulation of the Einstein action can be applied to all higher genus surfaces. By choosing an appropriate $\hat{r}$ on the Poincaré Disk to produce a regular $(4g)$-sided polygon with an interior angle of $\frac{2\pi}{4g}$, tilings of the hyperbolic plane are produced. Algebraic topology describes a genus g surface, a connected sum of g tori, in terms of $2g$ cycles [9]:

$$a_1 b_1 a_1^{-1} b_1^{-1} a_2 b_2 a_2^{-1} b_2^{-1} \cdots a_g b_g a_g^{-1} b_g^{-1} \ .$$

We use this expression to read off the identification of sides of the tile. In the $T^2 \# T^2$ case, $a_1 b_1 a_1^{-1} b_1^{-1} a_2 b_2 a_2^{-1} b_2^{-1} \leftrightarrow ABCDEFGH$ tells us to identify sides $A\&C$, $B\&D$, $E\&G$, $F\&H$, each with the orientation of the 2 identified sides reversed. That is, choose a representation of the fundamental group of the surface in the gauge group ISO(2,1). Each additional "hole" formed by the surface simply adds two more holonomies, $[W_{2g+1}|\Pi_{2g+1}]$ and $[W_{2g+2}|\Pi_{2g+2}]$, to the collection that determines the geometry of the surface. The constraint that the transformation producing a complete circuit of the vertex is the Identity ensures each additional torus glues smoothly to the rest:

$$W_1 W_2^{-1} W_1^{-1} W_2 \cdots W_{2g+1} W_{2g+2}^{-1} W_{2g+1}^{-1} W_{2g+2} = \mathbf{1} \quad \Longrightarrow$$

$$\begin{aligned} W_{2g+2}^{-1} W_{2g+1} W_{2g+2} W_{2g+1}^{-1} &= W_1 W_2^{-1} W_1^{-1} W_2 \cdots W_{2g-1} W_{2g}^{-1} W_{2g-1}^{-1} W_{2g} \\ T^2_{g+1} \ &\# \ \left(T_1^2 \# \cdots \# T_g^2\right) \ . \end{aligned}$$

These transformations can be attached to the tile with the "phase difference" approach. The closure relation on the W's guarantees a well-defined $\mathcal{U}$ field. Analogously, two more terms, $(\Pi)_{2g+1}\{\dot{W}\}_{2g+1}$ and $(\Pi)_{2g+2}\{\dot{W}\}_{2g+2}$, are added to the action to account for these new degrees of freedom. The problems of conjugate variables and quantization are still present, but not further obscured by the increase in genus.

6.2 Future Research

The Chern-Simons action based on gauge field theory and the usual Einstein action of GR are two different representations of the same system. The former deals with the gauge group theory of a principle ISO(2,1)-bundle over a (compact) surface, while the latter looks at invariants of the Einstein action over a (2+1)-dimensional splitting of spacetime. By comparing the ISO(2,1) gauge invariance of the Chern-Simons field $A_\mu = e_\mu{}^a P_a + \omega_\mu{}^a J_a$ with the usual Lorentz + diffeomorphism invariance of the *dreibein* $e_\mu{}^a$ and connection $\omega_\mu{}^a$, Witten [1] shows that the 2 representations differ only by a transformation that is part of the gauge group, and thus is inconsequential. Therefore, the general results of principle bundles and gauge theory can be used to study the specific case of (2+1)-dimensional spacetime. There is reason to suspect, however, that the two approaches are equivalent only under special circumstances.

Recall from §1.2 that the variation of A_μ is given by $\delta A_\mu = \delta e_\mu{}^a P_a + \delta\omega_\mu{}^a J_a$ where

$$\begin{aligned} \delta e_\mu{}^a &= -\rho^a_{,\mu} - \epsilon^{abc} e_{\mu b}\tau_c - \epsilon^{abc}\omega_{\mu b}\rho_c \ , \\ \delta\omega_\mu{}^a &= -\tau^a_{,\mu} - \epsilon^{abc}\omega_{\mu b}\tau_c \ . \end{aligned}$$

Under the substitution $\rho^a = V^\mu e_\mu{}^a$ these transformations coincide, on the solutions to the equations of motion, to infinitesimal diffeomorphisms and Lorentz transformations, the invariants of the Einstein action. In the case where $e_\mu{}^a$ is everywhere non-vanishing, this gauge transformation is a physical coordinate transformation generated by the vector field $V^\mu = \rho^a e^\mu{}_a$. Note that $V^\mu = V^\mu(e^\mu{}_a)$, where $e^\mu{}_a = g^{\mu\nu} e_{\nu a}$.

Now on all but the torus T^2, the tangent vector field to the spatial surface must have at least one singularity. Suppose the spatial $e_i{}^a$ fail to span the tangent space at the point $\bar{x}^\nu$. Where does this point "slide" to under the diffeomorphism generated by $\rho^a = V^\mu e_\mu{}^a$? We cannot say, as the vector field $V^\mu(\bar{x}^\nu)$ cannot be determined! The perfectly acceptable gauge transformation is no longer a coordinate transformation and hence is no longer physical. When we drop the requirement that the results be physical, we are left with an exercise in mathematics, not a theory of spacetime dynamics.

There is another problem related to this singularity. Suppose we are at a point $\bar{x}^\nu$ in flat spacetime where the *dreibein* is *not* singular. At this point, we perform the gauge transformation with parameter ρ^a chosen such that $\rho^a_{,\mu} = e_\mu{}^a$, so that $\delta e_\mu{}^a = -e_\mu{}^a$ and the *dreibein* becomes singular. This simple gauge transformation does not have a corresponding coordinate transformation. Furthermore, we cannot perform a gauge invariant *coordinate* transformation to get away from this singularity, for the diffeomorphism must be generated by $V^\mu e_\mu{}^a$ which vanishes for all V^μ. This suggests we can take two very different spacetime (Euclidean and Minkowski $\mathbf{R}^3$, for instance) and glue them together at this point $\bar{x}^\nu$. Moreover, there are *gauge* transformations that allow us to pass through this point from one spacetime to the other. It may even be possible to extend this "phenomenon" to a whole region, allowing us to construct a manifold whose metric changes signature. Clearly the equivalence of the Chern-Simons and Einstein actions has interesting details as yet unexplored.

One more question raised by the results deals with a subtlety of the *dreibein* approach that suggests this formalism is somehow "larger" than GR. On all but the torus, the vector field tangent to the spatial surfaces must vanish at one or more points. This means the full spacetime metric must either be singular, or at least time-like. One of the gauge transformations allowed by the *dreibein* representation of the Einstein action is the internal transformation $e_\mu{}^a \to \hat{e}_\mu{}^a + D_\mu \rho^a$. In general, this transformation will change the metric components $\hat{g}_{\mu\nu} = \hat{e}_\mu{}^a \hat{e}_{\nu a}$. At the point(s) where the metric $g_{\mu\nu}$ is time-like, a *gauge* equivalent metric $\hat{g}_{\mu\nu}$ may be space-like or null — clearly this is not a coordinate transformation. This phenomenon is prohibited in the usual form of GR written in terms of the metric and its derivatives. There is a freedom allowed by the *dreibein* formalism not allowed in GR. It may be possible to construct an explicit model and study it in analogy with the gauge theory explanation for the Bohm-Aharonov effect.[10] Overlapping coordinate patches may be related by a gauge transformation, but not a coordinate transformation.

General Relativity is simpler in (2+1)-dimensions in the *dreibein* formalism because the conditions that describe the slicing of 3-dimensional spacetime force the spatial hypersurfaces to be flat. We have considered the cases where these 2-dimensional slices are folded up into compact surfaces, the torus T^2 and the two-holed torus $T^2 \# T^2$. While they remain flat, removing the geometric degrees of freedom, the topology of these higher genus surfaces becomes important. The Einstein action becomes a functional not of the geometric quantities $g_{\mu\nu}$ but the topological quantities $[W|\Pi]$, the ISO(2,1) holonomies over the surface. The flatness of the genus g surface removes all but $2g$ distinct holonomies, and the action is written entirely in terms of these Poincaré transformations.

From the form of the reduced action, we can speculate of the dynamical variables and the phase space. It appears that the configuration space is the space of Lorentz transformations while the conjugate momenta lie in the collection of spacetime translations. We make this interpretation because the translations Π^a are related to infinitesimal Lorentz transformations, just as classically, momenta generate translation in configuration space.

While the action we have produced does not truly reveal the dynamical variables of this spacetime, and quantization of the phase space is not obvious, the method employed to reduce the action is quite revealing. It is apparent that the correct phase space of the classical conjugate variables is not simply $\mathbf{R}^{2n}$. It is most likely curved, and quantum mechanics on curved space is a problem that will not be tackled here. The result raised interesting questions about the gauge structure of spacetime, and also, therefore, about (2+1)-dimensional gravity over compact surfaces.

Bibliography

[1] Witten, E. 1988, "2+1 Dimensional Gravity as an Exactly Soluble System," *Nucl. Phys.*, **B311**, 46-78.

[2] Misner, C.W., Thorne, K.S., and Wheeler, J.A. 1973, *Gravitation* (San Francisco: Freeman).

[3] Wald, R.M. 1984, *General Relativity* (Chicago: The University of Chicago Press).

[4] Carlip, S. 1990, "Observables, Gauge Invariance, and Time in (2+1)-Dimensional Quantum Gravity," *Phys. Rev.* **D42**, 2647-2654.

[5] Carlip, S. 1989, "Exact Quantum Scattering in 2+1 Dimensional Gravity," *Nucl. Phys.*, **B324**, 106-122.

[6] Steenrod, N. 1951, *The Topology of Fibre Bundles* (Princeton: Princeton University Press).

[7] Arnowitt, R., Deser, S., and Misner, C.W. 1962, "The Dynamics of General Relativity" in *Gravitation: An Introduction to Current Research*, ed. L. Witten (New York: Wiley).

[8] Kelly, P., and Matthews, G. 1981, *The Non-Euclidean, Hyperbolic Plane* (New York: Springer-Verlag).

[9] Massey, W.S. 1967, *Algebraic Topology: An Introduction* (New York: Harcourt, Brace & World).

[10] Ryder, L.H. 1985, *Quantum Field Theory* (Cambridge: Cambridge University Press).

[11] Cornwell, J.F. 1984, *Group Theory in Physics*, Vol.**2** (London: Academic Press).

[12] Green, M.B., Schwarz, J.H., and Witten, E. 1987, *Superstring Theory*, Vol. **2**:*Loop Amplitudes, Anomolies and Phenomenology* (Cambridge: Cambridge University Press). Chapters 4 and 12 contain an introduction to the *dreibein* formalism.

Appendix A

Properties of Lorentz Transformations $\mathcal{U}^a{}_b$

By definition, a SO(2,1) Lorentz transformation $\mathcal{U}^a{}_b$ must keep the metric η^{ab} invariant:

$$\mathcal{U}^a{}_c \mathcal{U}^b{}_d \eta^{cd} = \eta^{ab} \ .$$

From this we see

$$\mathcal{U}^a{}_c \mathcal{U}^{bc} = \eta^{ab}$$

or

$$\boxed{\mathcal{U}^a{}_c \mathcal{U}_b{}^c = \delta^a_b} \tag{A.1}$$

A similar relation can be derived can be derived from this one. Re-write (A.1) as

$$\mathcal{U}_d{}^c \eta^{da} \mathcal{U}^b{}_c = \eta^{ab}$$

so that

$$\mathcal{U}_d{}^c \eta^{da} \mathcal{U}^b{}_c \eta_{ab} = 1 \ . \tag{A.2}$$

Suppose, in all generality, that the transformation $\mathcal{U}$ has both different left and right inverses:

$$\mathcal{U}(\mathcal{U}^{-1})_R = \mathbf{1} \qquad (\mathcal{U}^{-1})_L \mathcal{U} = \mathbf{1} \ .$$

Together, these give

$$(\mathcal{U}^{-1})_L \underbrace{\mathcal{U}(\mathcal{U}^{-1})_R} = (\mathcal{U}^{-1})_L$$

$$(\mathcal{U}^{-1})_L = (\mathcal{U}^{-1})_R$$

so that the left and right inverses are the same. From (A.2) we see $\eta^{da}\mathcal{U}^b_{\ c}\eta_{ab}$ is the right inverse of $\mathcal{U}_d^{\ c}$. Left and right inverses coincide, so

$$\eta^{da}\mathcal{U}^b_{\ f}\eta_{ab}\mathcal{U}_d^{\ c} = \delta^c_f$$

which is equivalent to

$$\boxed{\mathcal{U}^c_{\ a}\mathcal{U}_c^{\ b} = \delta^b_a} \tag{A.3}$$

A third property of Lorentz transformations is also used in producing the results of Chapter 2. We come across terms of the form

$$\epsilon_{abc}\mathcal{U}^a_{\ x}\mathcal{U}^b_{\ y}\mathcal{U}^c_{\ z} \ .$$

The indices x, y, z are still Lorentz indices, named from the end of the alphabet to clarify the index maniupulations. Relabel the dummy indices b, c:

$$\begin{aligned}\epsilon_{abc}\mathcal{U}^a_{\ x}\mathcal{U}^b_{\ y}\mathcal{U}^c_{\ z} &= \epsilon_{acb}\mathcal{U}^a_{\ x}\mathcal{U}^c_{\ y}\mathcal{U}^b_{\ z} \\ &= -\epsilon_{abc}\mathcal{U}^a_{\ x}\mathcal{U}^c_{\ y}\mathcal{U}^b_{\ z} \ .\end{aligned}$$

Comparing the first and last expression, we see $\epsilon_{abc}\mathcal{U}^a_{\ x}\mathcal{U}^b_{\ y}\mathcal{U}^c_{\ z}$ is antisymmetric in b and c. Analogous relabelling shows that $\epsilon_{abc}\mathcal{U}^a_{\ x}\mathcal{U}^b_{\ y}\mathcal{U}^c_{\ z}$ is totally antisymmetric, and therefore must be proportional to the *only* totally antisymmetric 3-tensor, ϵ_{xyz}: Write

$$\epsilon_{abc}\mathcal{U}^a_{\ x}\mathcal{U}^b_{\ y}\mathcal{U}^c_{\ z} = \lambda\epsilon_{xyz} \ .$$

Solve for λ:

$$\begin{aligned}\lambda &= \epsilon^{xyz}\epsilon_{abc}\mathcal{U}^a_{\ x}\mathcal{U}^b_{\ y}\mathcal{U}^c_{\ z} \\ &= \det(\mathcal{U}^a_{\ b}) \ .\end{aligned}$$

These Special Orthogonal matrices have unit determinant, so $\lambda = 1$, and

$$\boxed{\epsilon_{abc}\mathcal{U}^a_{\ x}\mathcal{U}^b_{\ y}\mathcal{U}^c_{\ z} = \epsilon_{xyz}} \tag{A.4}$$

Appendix B

Setting $\tilde{\omega}_i{}^a{}_b = 0$ and $\hat{e}_i{}^a = 0$

We have written

$$\omega_\mu{}^a{}_b = \tilde{\omega}_\mu{}^c{}_d \mathcal{U}_c{}^a \mathcal{U}^d_b + \mathcal{U}_c{}^a (\mathcal{U}^c{}_{b,\mu})$$

and claimed that because $R_{ij}{}^a{}_b = 0$, we can find a $\mathcal{U}^a_b$ such that $\tilde{\omega}_\imath{}^a_b = 0$. That is, there is a $\mathcal{U}^a_b$ for which

$$\omega_\imath{}^a_b = \mathcal{U}_c{}^a (\mathcal{U}^c{}_{b,i})$$

or

$$\mathcal{U}^a{}_{b,i} = \mathcal{U}^a{}_c \omega_i{}^c{}_b \ . \tag{B.1}$$

Formally we write the solution of this matrix equation

$$\mathcal{U} = e^{\int \omega_\mu dx^\mu} \ .$$

This is not well-defined as $\mathcal{U}(x^\mu)$ will in general have a different value for each path $\gamma(\lambda)$ integrated over, as

$$dx^\mu = \frac{d\gamma^\mu(\lambda)}{d\lambda} d\lambda$$

will be different for each path. Instead we write

$$\mathcal{U} = \mathcal{P} e^{\int_\gamma \omega_\mu dx^\mu} \tag{B.2}$$

where $\mathcal{P}$ stands for path-ordering the integral along the path γ. The meaning of a path-ordered integral can be seen as follows. Discretize the path $\gamma(\lambda)$ over the surface under

$0 = \lambda_0 < \lambda_1 < \cdots < \lambda_n = 1$ with $\gamma(\lambda_k) = x_k$. Then the integral (B.2) is the limiting value

$$\lim_{n\to\infty} e^{\omega_\mu(x_n)\Delta x^\mu_n} \ldots e^{\omega_\mu(x_1)\Delta x^\mu_1} e^{\omega_\mu(x_0)\Delta x^\mu_0}$$

Each term $e^{\omega_\mu(x_k)\Delta x^\mu_k}$ is a transformation acting on the term to the right, so that (B.2) is an infinite succession of infinitesimal transformations in a direction tangent to the path, parallel transporting along $\gamma(\lambda)$. If Δx_k is tangent to the path at x_k then

$$dx(x_k) = \frac{d\gamma(\lambda_k)}{d\lambda} d\lambda \simeq \Delta x_k$$

and

$$\frac{d}{dx^\nu(x_k)} \mathcal{P} e^{\int_\gamma \omega_\mu dx^\mu} = \mathcal{P} e^{\int_\gamma \omega_\mu dx^\mu} \omega_\nu(x_k)$$

or,

$$\mathcal{U}_{,\nu} = \mathcal{U}\omega_\nu$$

as required in (B.1).

There is no reason, however, for this result to hold for nearby paths. Integrated over a different path $\tilde{\gamma}$ between x_o and x_1, so that $\tilde{\gamma}(\tilde{\lambda}_m) = y_m = x_n = \gamma(\lambda_n)$, $\mathcal{U}(y_m)$ and $\mathcal{U}(x_n)$ may have different values. While both

$$\frac{d}{dx^\nu} \mathcal{P} e^{\int_\gamma \omega_\mu dx^\mu} = \mathcal{P} e^{\int_\gamma \omega_\mu dx^\mu} \omega_\nu \ ,$$

$$\frac{d}{dy^\nu} \mathcal{P} e^{\int_{\tilde{\gamma}} \omega_\mu dy^\mu} = \mathcal{P} e^{\int_{\tilde{\gamma}} \omega_\mu dy^\mu} \omega_\nu \ ,$$

it is not true in general that, say,

$$\frac{d}{dy^\nu} \mathcal{P} e^{\int_\gamma \omega_\mu dx^\mu} = \mathcal{P} e^{\int_\gamma \omega_\mu dx^\mu} \omega_\nu(x_k) \ ,$$

for in general

$$dx(x_n) = \frac{d\gamma(\lambda_n)}{d\lambda} d\lambda \neq \frac{d\tilde{\gamma}(\tilde{\lambda}_m)}{d\tilde{\lambda}} d\tilde{\lambda} = dy(y_m).$$

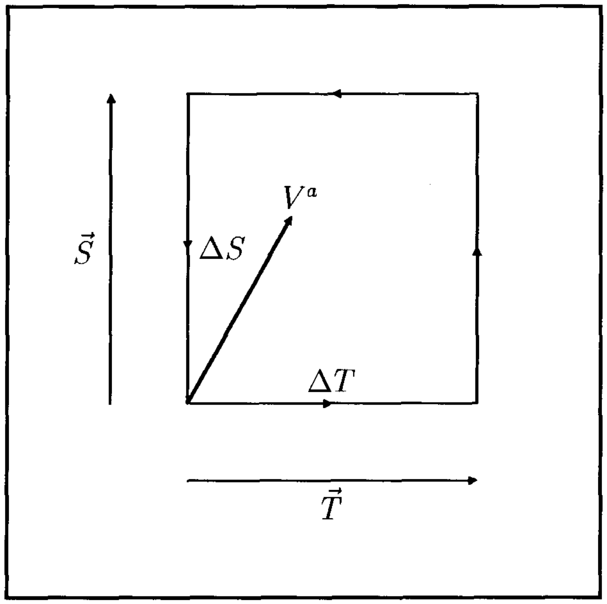

Figure B.1: An infinitesimal loop of generated by vectors T^μ, S^ν with area $\Delta T \Delta S$.

The constraints, however, tell us that the spatial components of the Reimann curvature tensor vanish, $R_{ij}{}^a{}_b = 0$. Recall what the Reimann tensor represents: Parallel transporting a vector V^a about an infinitesimal loop (See Figure B.1) changes V^a infinitesimally by

$$\delta V^a = \Delta T \Delta S V^b T^\mu S^\nu R_{\mu\nu b}{}^a \ .$$

The change in the vector V^a is proportional to $R_{\mu\nu b}{}^a$ and to the loop's area $\Delta T \Delta S$. Notice that the indices μ, ν link the components of the tensor to the paths in the directions T^μ, S^ν.

Returning to our problem, suppose $\mathcal{U}(x^\mu)$ is calculated along the path $\gamma(\lambda)$ as in (B.2). When $dx^\mu = (d\gamma^\mu/d\lambda)d\lambda$, equation (B.1) holds. We want to show that because $R_{ij}{}^a{}_b = 0$, the result holds for all paths on the surface:

$$\frac{d}{dy^i(y_n)}\mathcal{U}(x_n) = \mathcal{U}(x_n)\omega_i(x_n)$$

for some other $y^\mu = \tilde{\gamma}^\mu(\tilde{\lambda})$ passing through x_n.

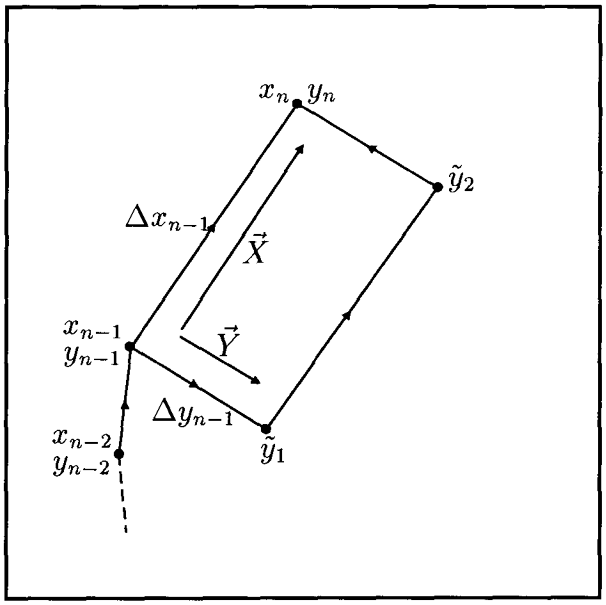

Figure B.2: Two infinitesimally different paths to $x_n = y_n$.

Consider an infinitesimal change in the path γ (See Figure B.2). Along coincident paths, $\mathcal{U}(x_0) = \mathcal{U}(y_0)$,$\mathcal{U}(x_1) = \mathcal{U}(y_1)$,...,$\mathcal{U}(x_{n-1}) = \mathcal{U}(y_{n-1})$. Whereas $\mathcal{U}(x_n)$ is calculated in the limit

$$\lim_{n\to\infty} e^{\omega_\mu(x_{n-1})\Delta x^\mu_{n-1}}\mathcal{U}(x_{n-1}) \ ,$$

$\mathcal{U}(y_n)$ is calculated by traversing the loop of area $\Delta x_{n-1}\Delta y_{n-1}$. The two paths from $x_0 = y_0$ to $x_n = y_n$ are

$$x_0,\ x_1\ \cdots\ x_{n-2},\ x_{n-1},\ x_n\ ,$$

$$y_0,\ y_1\ \cdots\ y_{n-2},\ y_{n-1},\ \tilde{y}_1,\ \tilde{y}_2,\ y_n\ .$$

Their difference is the *loop*

$$x_n,\ y_{n-1},\ \tilde{y}_1,\ \tilde{y}_2,\ y_n\ .$$

If $\mathcal{U}(x_n)$ had been calculated by path-ordered integration along the varied path, then the difference in its value would be

$$\delta\mathcal{U}^a{}_b = \Delta x_{n-1}\Delta y_{n-1} X^i Y^j (\mathcal{U}^c{}_b R_{ijc}{}^a - \mathcal{U}^a{}_c R_{ijb}{}^c) \ .$$

Only the spatial components are present as the paths lie in a surface of constant time. Since $R_{ij}{}^a{}_b = 0$, there is no change in $\mathcal{U}^a{}_b$ because of this infinitesimal change in the path.

Macroscopic changes in γ can be built from many infinitesimal ones, the result holding at each stage. The result tells us that the value of $\mathcal{U}^a{}_b$ at x^μ is independent of the path along which the integration occurs. Therefore, in checking that (B.1) is true at any point, *we can assume that the path used to calculate* $\mathcal{U}^a{}_b$ *in* (B.2) *is the one for which* $dx^i = (d\gamma^i/d\lambda)d\lambda$, so that (B.1) holds.

Finally, with $\mathcal{U}_{,i} = \mathcal{U}\omega_i$ at all points on the tile,$\tilde{\omega}_i{}^a{}_b = 0$ must vanish, leaving only the time components undetermined.

In the coordinate system where $\tilde{\omega}_i{}^a{}_b = 0$, we transformed the *dreibein* under

$$\tilde{e}_\mu{}^a = \hat{e}_\mu{}^a + \tilde{D}_\mu \rho^a \ .$$

In this coordinate sytem, $\tilde{D}_i\rho^a = \rho^a_{,i}$ as the spatial $\tilde{\omega}_i{}^a{}_b = 0$. We claim that because $D_{[i}e_{j]}^a = 0$, or $\tilde{e}^a_{[i,j]} = 0$, we can find a particular $\rho^a(x^\mu)$ such that $\hat{e}_\mu{}^a = 0$. The solution to the PDE

$$\tilde{e}_\mu{}^a = \rho^a{}_{,i} \tag{B.3}$$

can formally be written as

$$\rho^a = \mathcal{P}\int_\gamma \tilde{e}_\mu{}^a dx^\mu$$

where again $\mathcal{P}$ indicates path-ordering along the path γ. The result (B.3) holds along the path where $dx^\mu = (d\gamma^\mu/d\lambda)d\lambda$, but the result need not be true for different paths.

Like the $\omega_\mu{}^a{}_b$ case above, we interpret the path-ordered integral as a limit. Discretize the path $\gamma(\lambda)$ so that $x_k = \gamma(\lambda_k)$ and Δx_k is tangent to the path at, say, $x_k + \Delta x_k/2$.

The point where we evaluate does not matter in the limit. The path-ordered integral is the limiting value

$$\lim_{n\to\infty} \tilde{e}_\mu{}^a(x_o + \frac{\Delta x_o}{2})\Delta x_o^\mu + \tilde{e}_\mu{}^a(x_1 + \frac{\Delta x_1}{2})\Delta x_1^\mu + \cdots + \tilde{e}_\mu{}^a(x_n + \frac{\Delta x_n}{2})\Delta x_n^\mu \; .$$

To show that the value of $\rho^a(x^\mu)$ is in fact independent of the path to x^μ, again consider the two paths between $x_o = y_o$ and $x_n = y_n$ used above (See Figure B.2). The difference between the two paths, recall, is the loop

$$x_n, \; y_{n-1}, \; \tilde{y}_1, \; \tilde{y}_2, \; y_n \; .$$

This loop will contribute to the sum (or 'integral') an amount

$$\begin{aligned}\tilde{e}_\mu{}^a(x_n - \frac{\Delta x_{n-1}}{2})\Delta x_{n-1}^\mu + \tilde{e}_\mu{}^a(x_{n-1} + \frac{\Delta y_{n-1}}{2})\Delta y_{n-1}^\mu \\ + \tilde{e}_\mu{}^a(\tilde{y}_1 + \frac{\Delta x_{n-1}}{2})\Delta x_{n-1}^\mu - \tilde{e}_\mu{}^a(\tilde{y}_2 - \frac{\Delta y_{n-1}}{2})\Delta y_{n-1}^\mu \; .\end{aligned}$$

These paths, and this loop, lie in a surface of constant time, so the displacements Δx_k and Δy_k have no time-, or 0-, components. Furthermore, we can assume that these displacements are in the spatial 1- and 2-directions, for any loop can be approximated to an arbitrary degree by a tiling of parallelograms whose sides *are* in the 1- and 2-directions. Thus with $\Delta x_{n-1} = (0, \Delta x, 0)$ and $\Delta y_{n-1} = (0, 0, \Delta y)$, the loop contributes

$$\begin{aligned}\Bigg(&\frac{\tilde{e}_1{}^a(x_n - \frac{\Delta x}{2} + \Delta y) - \tilde{e}_1{}^a(x_n - \frac{\Delta x}{2})}{\Delta y} \\ &- \frac{\tilde{e}_2{}^a(x_{n-1} + \frac{\Delta y}{2} + \Delta x) - \tilde{e}_2{}^a(x_{n-1} + \frac{\Delta y}{2})}{\Delta x}\Bigg) \; \Delta x \Delta y \; .\end{aligned}$$

In the limit where $n \to \infty$, this is simply

$$(\tilde{e}_{1,2}^{\,a} - \tilde{e}_{2,1}^{\,a})\Delta x \Delta y \; .$$

The constraint tells us $\tilde{e}_{[i,j]}^{\,a} = 0$, so variations in the path do not affect the value of the path-ordered integral. Because of this path independence, we can always assume $dx^i = (d\gamma^i/d\lambda)d\lambda$ so that $\tilde{e}_i{}^a = \tilde{D}_i\rho^a$. Therefore, $\hat{e}_i{}^a$ must vanish.

Appendix C

Properties of $\{\dot{W}\}$

We define the symbol

$$\{\dot{W}\}_a = \dot{W}^b{}_x W^{cx} \epsilon_{abc}$$

so that

$$\Pi^a \dot{W}^b{}_x W^{cx} \epsilon_{abc} = \Pi\{\dot{W}\} \ .$$

There are three properties of this symbol which will simply the algebra of reducing the action to a more canonical form.

By definition,

$$\begin{aligned}
(\mathcal{U}\Pi)\,\{\mathcal{U}\dot{W}\} &= (\mathcal{U}\Pi)^a \mathcal{U}^b{}_y\, \dot{W}^y{}_x \mathcal{U}^c{}_z\, W^{zx} \epsilon_{abc} \\
&= \Pi^a \mathcal{U}^a{}_d \mathcal{U}^b{}_y \mathcal{U}^c{}_z\, \dot{W}^y{}_x\, W^{zx} \epsilon_{abc} \ .
\end{aligned}$$

Recalling property (A.4) of Lorentz transformations $\mathcal{U}$,

$$= \Pi^d \dot{W}^y{}_x W^{zx} \epsilon_{abc} \ .$$

Relabelling $dyz \to abc$ gives

$$= \Pi^a \dot{W}^b{}_x W^{cx} \epsilon_{abc} \ .$$

That is,

$$\boxed{(\mathcal{U}\Pi)\,\{\mathcal{U}\dot{W}\} = \Pi\{\dot{W}\}} \tag{C.1}$$

Next, in full component form,

$$\begin{aligned}
\Pi\{\overset{\cdot}{\overline{W_1 W_2}}\} &= \Pi^a (\overset{\cdot}{\overline{W_1{}^b{}_d W_2{}^d{}_x}})(W_1{}^b{}_d W_2{}^d{}_x)\epsilon_{abc} \\
&= \Pi^a \dot{W}_1{}^b{}_d W_2{}^d{}_x W_1{}^c{}_e W_2{}^{ex}\epsilon_{abc} \\
&\quad + \Pi^a W_1{}^b{}_d \dot{W}_2{}^d{}_x W_1{}^c{}_e W_2{}^{ex}\epsilon_{abc} \,.
\end{aligned}$$

Now the first term can be rewritten by first raising and lowering the e:

$$\begin{aligned}
\Pi^a \dot{W}_1{}^b{}_d W_1{}^{ce} \underbrace{W_2{}^d{}_x W_{2e}{}^x}_{\delta^d_e}\epsilon_{abc} &= \Pi^a \dot{W}_1{}^b{}_d W_1{}^{cd}\epsilon_{abc} \\
&= \Pi\{\dot{W}_1\} \,.
\end{aligned}$$

The second term is simply $\Pi\{W_1 \dot{W}_2\}$ so that

$$\boxed{\Pi\{\overset{\cdot}{\overline{W_1 W_2}}\} = \Pi\{\dot{W}_1\} + \Pi\{W_1 \dot{W}_2\}} \tag{C.2}$$

Lastly, using (C.2) just derived, we can write

$$\left(W_2^{-1}\Pi\right)\{\overset{\cdot}{\overline{W_2^{-1} W_1}}\} = \left(W_2^{-1}\Pi\right)\{W_2^{-1}\dot{W}_1\} + \left(W_2^{-1}\Pi\right)\{\dot{W}_2\} \,.$$

While $\left(W_2^{-1}\Pi\right)\{W_2^{-1}\dot{W}_1\} = \Pi\{\dot{W}_1\}$ by (C.1),

$$\begin{aligned}
\left(W_2^{-1}\Pi\right)\{\dot{W}_2\} &= \left(W_2 W_2^{-1}\Pi\right)\{W_2 \dot{W}_2^{-1}\} \\
&= \Pi\left(\{\overset{\cdot}{\overline{W_2 W_2^{-1}}}\} - \{\dot{W}_2\}\right) \\
&= -\Pi\{\dot{W}_2\}
\end{aligned}$$

as $\{\dot{1}\} \propto \dot{\delta}^b_x = 0$. Therefore

$$\boxed{\left(W_2^{-1}\Pi\right)\{\overset{\cdot}{\overline{W_2^{-1} W_1}}\} = \Pi(\{\dot{W}_1\} - \{\dot{W}_2\})} \tag{C.3}$$