TRIANGLES IN THE HEISENBERG GROUP

TRIANGLES IN THE HEISENBERG GROUP

By

CHRISTOPHER CAPPADOCIA, B.MATH.

A Thesis
Submitted to the School of Graduate Studies
in Partial Fulfilment of the Requirements
for the Degree
Master of Science

McMaster University

MASTER OF SCIENCE (2010)
(Mathematics)

McMaster University
Hamilton, Ontario

TITLE: Triangles in the Heisenberg Group
AUTHOR: Christopher Cappadocia, B.Math. (University of Waterloo)
SUPERVISOR: Professor Andrew Nicas
NUMBER OF PAGES: x, 148

Abstract

In the Heisenberg Lie group with the Carnot-Caratheodory metric, we classify geodesic triangles up to isometry in terms of side-length and geodesic parameters. We obtain an angle deficit formula for Heisenberg triangles. We construct classical moduli spaces T and $S_3\backslash T$ for ordered and for unordered Heisenberg triangles respectively, computing homotopy type and manifold properties of the spaces, and producing a compactification of T up to similarity under the non-isotropic dilation.

Acknowledgements

Foremost I would like to thank my thesis advisor, Dr. Andrew Nicas, whose enthusiasm for mathematics inspired and motivated me at every stage of this project; for his patient, expert guidance I am truly grateful.

I would also like to thank the mathematics department at McMaster, in particular Dr. Ian Hambleton and Dr. Maung Min-Oo, whose courses and seminars provided a rich environment in which to learn and do mathematics, and Dr. McKenzie Wang, who, along with Dr. Nicas and Dr. Min-Oo, formed my supervisory committee.

Last my fellow students, for always delivering on the impromptu lessons I so often requested, and all my family, with special mention to Joe, Christie, Hunter, and Stuart, who gave me a great place to live while I worked on this.

All figures were created using Maple, and funding for the author was provided in part by the Ontario Graduate Scholarship program.

Contents

List of Figures

0.1 Introduction

Both spherical and hyperbolic trigonometry admit close analogues of the Euclidean laws of sines and cosines, as well as geometric interpretations of the sum of the angles in a triangle. These three classical settings for trigonometry – the plane, the sphere, and the hyperbolic disk – are 2-dimensional surfaces with a high degree of symmetry[1].

The Heisenberg group $\mathbb{H}$ is not a surface but a 3-dimensional manifold, in fact a Lie group. Nevertheless the 2-step nilpotent group structure of $\mathbb{H}$ distinguishes two of the three dimensions by way of the horizontal distribution $\mathcal{H}$. The sub-Riemannian geometry that results is intimately connected with the Carnot-Caratheodory distance on $\mathbb{H}$ and produces a space with a great deal of symmetry.

Thus it is natural to ask what form trigonometry might take in the Heisenberg group. Such an investigation is aided greatly by the fact that there is already a complete explicit description of geodesics in $\mathbb{H}$.

This thesis proceeds in two directions: one, we attempt trigonometry in the Heisenberg group; and two, we describe the space of all triangles in the Heisenberg group, in which each point of the space corresponds to a unique isometry class of triangles. When we construct such a space we obtain an example of a *moduli space*[2].

Chapter 1 describes the construction of the moduli space for Euclidean triangles. It is hoped that the discussion of the Euclidean case will motivate by analogy the constructions for the Heisenberg case.

Chapter 2 provides the necessary background material. We cover the horizontal distribution $\mathcal{H}$ and the induced sub-Riemannian geometry on the Heisenberg group. The main purpose of the chapter is to record in full detail the complete classification of geodesics in $\mathbb{H}$.

Chapter 3 presents our efforts towards trigonometry. In particular, we obtain

[1]Each is *homogeneous* and *isotropic*; in words, the geometry is the same at every point and the same in all directions.

[2]To be precise, for ordered triangles we obtain a *fine* moduli space, and for unordered triangles we obtain a *coarse* moduli space.

a formula relating the angle deficit of a Heisenberg triangle to the total curvature around the triangle. Also, we obtain a trigonometric identity that is in some ways analogous to a law of sines.

Chapter 4 constructs the moduli space of Heisenberg triangles. We show that the moduli space is a 5-dimensional smooth manifold with the homotopy type of a thrice-punctured sphere. The moduli space embeds naturally in $\mathbb{R}^6$, and in this context we describe its boundary. There is a notion of similar triangles provided by the *non-isotropic dilation* of the Heisenberg group. When we take this action into account on the moduli space and add in the boundary, we obtain a compactification of the moduli space of Heisenberg triangles.

Chapter 1

The moduli space of Euclidean triangles

1.1 Triangles in the plane

In the plane $\mathbb{R}^2$, three distinct and non-colinear points determine a unique triangle. To describe all possible triangles in the plane, we begin with the set

$$\widetilde{T} := \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2 - \Sigma,$$

where $\Sigma \subset \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2$ is the set of triples (A, B, C) such that the points A, B, C are colinear. Note that if two points from the triple (A, B, C) coincide, then necessarily A, B, C are colinear. Thus $\widetilde{T}$ is the set of all ordered triples of distinct, non-colinear points in $\mathbb{R}^2$.

Proposition 1.1.1. *The set $\widetilde{T} \subset \mathbb{R}^6$ is open.*

Proof. Notice that Σ can be characterized by the condition

$$\Sigma = \left\{(A, B, C) \in \mathbb{R}^6 : (B - A) \times (C - A) = \vec{0}\right\}$$

where $\times$ denotes the cross product. Thus Σ is closed, so that $\widetilde{T} = \mathbb{R}^6 - \Sigma$ is open. $\square$

The space $\widetilde{T}$ can be viewed as a "configuration space" for the set of all Euclidean triangles, in which each point in $\widetilde{T}$ corresponds to a unique triangle. However $\widetilde{T}$ has two conceptual drawbacks: (1) it describes triangles with a definite location, and thus distinguishes between congruent triangles, unless they coincide; (2) it describes *ordered* triangles, and thus distinguishes, for example, between the triangles corresponding to the triples (A, B, C) and (B, C, A).

To remedy (1), we introduce on $\widetilde{T}$ an action of the group $\mathbb{R}^2 \rtimes O(2)$ of isometries of the plane. Taking the quotient space of this group action will have the effect of identifying congruent ordered triangles. To remedy (2), we introduce an action of the symmetric group S_3 on the quotient space obtained in the previous step. This will have the effect of ignoring the ordering imposed on the vertices of our triangles.

In the process of applying these two group actions to $\widetilde{T}$, we will see that the space obtained is naturally identified with an alternative way to characterize the "configuration space" of Euclidean triangles, which we now describe. In this view, we think of a triangle as being completely determined by the lengths of its three sides. This approach implicitly identifies congruent triangles and ignores any notion of ordering of the sides of a triangle. To construct the space of triangles corresponding to this view, we start with the set

$$T := \left\{(a, b, c) \in \mathbb{R}^3 | a + b > c, b + c > a, c + a > b\right\},$$

an open subset of $\mathbb{R}^3$.

Now T is the space of *ordered* side lengths a, b, c, so we introduce a group action of S_3 on T whereby a permutation $\sigma \in S_3$ permutes the triple (a, b, c). When we identify triples in the same orbit of S_3 we obtain the quotient space $S_3 \backslash T$ which describes unordered triangles in the plane up to isometry.

We have seen that both $\widetilde{T} \subset \mathbb{R}^6$ and $T \subset \mathbb{R}^3$ are open sets. Thus $\widetilde{T}$ and T are open submanifolds of $\mathbb{R}^6$ and $\mathbb{R}^3$ respectively, and so inherit differentiable structures and become smooth manifolds in their own right. There is a natural map $p : \widetilde{T} \to T$ defined by

$$(A, B, C) \mapsto (d(B, C), d(C, A), d(A, B)).$$

The continuity of the metric $d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ implies the continuity of the map p. Also, it is intuitively clear that for every triple of side lengths (a, b, c) satisfying the triangle inequalities, there are many triples of vertices (A, B, C) forming triangles with the specified side lengths, so that p will be surjective. We show below how such a triangle may be constructed.

Proposition 1.1.2. *The map $p : \widetilde{T} \to T$ is surjective.*

Proof. Place the vertex A at the origin $(0,0)$ and the vertex B on the y-axis at $(0,c)$. Draw the circle of radius b centered at A, draw the circle of radius a centered at B, and notice that the triangle inequality $a+b>c$ ensures that the two circles intersect.

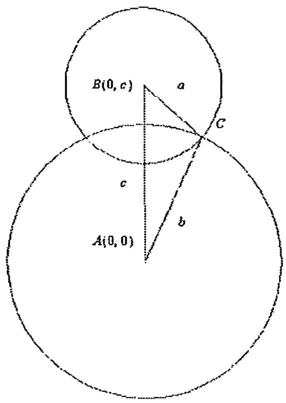


Figure 1.1: constructing a triangle with side-lengths a, b, c

There are thus two choices for the placement of the vertex C. □

We claim in addition that $p : \widetilde{T} \to T$ is an open map, that is, the image of an open subset of $\widetilde{T}$ is open in T. A clean way to see this is to write out the components of $p = (p^1, p^2, p^3)$ explicitly:

$$\begin{aligned}
p^1(x_A, y_A; x_B, y_B; x_C, y_C) &= \sqrt{(x_B - x_C)^2 + (y_B - y_C)^2}, \\
p^2(x_A, y_A; x_B, y_B; x_C, y_C) &= \sqrt{(x_C - x_A)^2 + (y_C - y_A)^2}, \\
p^3(x_A, y_A; x_B, y_B; x_C, y_C) &= \sqrt{(x_A - x_B)^2 + (y_A - y_B)^2}.
\end{aligned}$$

In this form it is apparent that p is a smooth map, since the expressions under each radical are non-zero on $\widetilde{T}$ and hence avoid the critical point $x = 0$ of the function $x \mapsto \sqrt{x}$. To see that the rank of the differential matrix $Dp(A, B, C)$ is 3 on $\widetilde{T}$, we write p as a composition of successive maps $p = q \circ r$, where $r : \mathbb{R}^6 \to \mathbb{R}^3$ is defined by

$$(x_A, y_A; x_B, y_B; x_C, y_C) \mapsto \begin{pmatrix} (x_B - x_C)^2 + (y_B - y_C)^2, \\ (x_C - x_A)^2 + (y_C - y_A)^2, \\ (x_A - x_B)^2 + (y_A - y_B)^2 \end{pmatrix}$$

and $q : \mathbb{R}^3 \to \mathbb{R}^3$ is defined by

$$(d_1, d_2, d_3) \mapsto \left(\sqrt{d_1}, \sqrt{d_2}, \sqrt{d_3}\right).$$

The differential $Dr(A, B, C)$ is

$$\begin{bmatrix} 0 & 0 & 2(x_B - x_C) & 2(y_B - y_C) & -2(x_B - x_C) & -2(y_B - y_C) \\ -2(x_C - x_A) & -2(y_C - y_A) & 0 & 0 & 2(x_C - x_A) & 2(y_C - y_A) \\ 2(x_A - x_B) & 2(y_A - y_B) & -2(x_A - x_B) & -2(y_A - y_B) & 0 & 0 \end{bmatrix}$$

and the differential $Dq(d_1, d_2, d_3)$ is

$$\begin{bmatrix} \frac{1}{2\sqrt{d_1}} & 0 & 0 \\ 0 & \frac{1}{2\sqrt{d_2}} & 0 \\ 0 & 0 & \frac{1}{2\sqrt{d_3}} \end{bmatrix}.$$

Proposition 1.1.3. *The rank of the differential matrix $Dr(A, B, C)$ is 3 if A, B, C are not colinear, i.e., the rank of $Dr(A, B, C)$ is 3 on $\widetilde{T}$.*

Proof. We examine the conditions in which the rows $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ of the 3×6 matrix $Dr(A, B, C)$ can be dependent. So assume

$$\alpha \mathbf{v}_1 + \beta \mathbf{v}_2 + \gamma \mathbf{v}_3 = 0$$

for real scalars α, β, γ not all zero. From the six components of this vector equation we obtain the system of equations

$$\begin{array}{llcl}
\text{(i)} & \beta(x_C - x_A) & = & \gamma(x_A - x_B) \\
\text{(ii)} & \beta(y_C - y_A) & = & \gamma(y_A - y_B) \\
\text{(iii)} & \alpha(x_B - x_C) & = & \gamma(x_A - x_B) \\
\text{(iv)} & \alpha(y_B - y_C) & = & \gamma(y_A - y_B) \\
\text{(v)} & \alpha(x_B - x_C) & = & \beta(x_C - x_A) \\
\text{(vi)} & \alpha(y_B - y_C) & = & \beta(y_C - y_A).
\end{array}$$

Suppose it is β that is non-zero. Then the fact that the points C and A differ means that at least one of the equations (i) and (ii) has non-zero left-hand side. From this it follows that γ is also non-zero. Equations (v) and (vi) likewise show that α is non-zero. Thus α, β, and γ are all non-zero. There are now three cases. Case (1), if any two of the x's are equal then the equations force all three of the x's to be equal (for example, $x_A = x_B$ and equations (i) and (iii) force $x_C = x_A$ and $x_B = x_C$ respectively). Then all three y's must be distinct since the points A, B, C are distinct. In this case A, B, C are colinear, lying on a vertical line. Case (2), similarly, if any two of the y's are equal then in fact all three y's are equal, and all three x's must be distinct. The points A, B, C are again colinear, this time lying on a horizontal line. Case (3), all three x's are distinct, and all three y's are distinct. Here we can divide, say, equation (ii) by equation (i) to get

$$\frac{y_A - y_B}{x_A - x_B} = \frac{y_C - y_A}{x_C - x_A},$$

which shows that any solution A, B, C must be colinear. □

Proposition 1.1.4. *The rank of $Dp(A,B,C)$ is 3 on $\widetilde{T}$.*

Proof. The matrix

$$Dq(d_1, d_2, d_3) = \begin{bmatrix} \frac{1}{2\sqrt{d_1}} & 0 & 0 \\ 0 & \frac{1}{2\sqrt{d_2}} & 0 \\ 0 & 0 & \frac{1}{2\sqrt{d_3}} \end{bmatrix}$$

is invertible on $r(\widetilde{T})$ since that is precisely where none of the d_i's are 0. Since the rank of $Dr(A,B,C)$ is 3 on $\widetilde{T}$, it follows that the rank of $Dp = Dq \circ Dr$ is also 3 on $\widetilde{T}$.

□

We now appeal to a version of the Rank theorem from the study of smooth manifolds. For a proof see [7].

Theorem 1.1.5. *Let M^n and N^m be smooth manifolds of dimension n and m respectively, with $n \geq m$, and suppose the smooth map $f : M^n \to N^m$ has rank m at the point $p \in M$. Then there exist coordinate systems (x, U) and (y, V) around p and $f(p)$ respectively such that*

$$y \circ f \circ x^{-1}(a^1, \dots, a^n) = (a^1, \dots, a^m).$$

Proposition 1.1.6. *The map $p : \widetilde{T} \to T$ is an open map.*

Proof. Given $(A, B, C) \in G$, where $G \subset \widetilde{T}$ is open. The rank of p is 3 at all points of $\widetilde{T}$, which equals the dimension, 3, of T. By the Rank Theorem, G contains an open neighborhood of (A, B, C) which p maps surjectively to an open neighborhood of $p(A, B, C)$. Therefore the image $p(G) \subset T$ is open.

□

We have thus constructed the continuous, open surjection $p : \widetilde{T} \to T$ defined by

$$(A, B, C) \mapsto (d(B, C), d(C, A), d(A, B)).$$

Proposition 1.1.7. *The points (A, B, C) and $(A', B', C') \in \widetilde{T}$ have the same image under p if and only if there exists an isometry g of $\mathbb{R}^2$ such that $g(A) = A'$, $g(B) = B'$, and $g(C) = C'$.*

Proof. One direction is immediate since an isometry preserves distance between points.

Suppose (A, B, C) and (A', B', C') have the same image (a, b, c) under p. We will describe a sequence of isometries on (A, B, C) and a sequence of isometries on (A', B', C') that bring these triangles to the same intermediate triangle. First, we translate the triangle (A, B, C) by $-A$, in effect sliding A to the origin. Second, we rotate the resulting triangle about the origin until B lies in the positive y-axis. A thus remains fixed at the origin throughout this rotation and, since each of these isometries preserves distance, B must end up at the point $(0, c) \in \mathbb{R}^2$. The point C

must end up at distance b from $A = O$ and at distance a from $B = (0, c)$. But a circle of radius b centered at A and a circle of radius a centered at B must intersect in precisely two points, giving two possibilities for C, and these two possibilities are related by a reflection in the y-axis.

In the same way we bring A' to the origin and B' to $(0, c)$, and see that C' can also lie only at one of the two possible points for C. Thus (A, B, C) and (A', B', C') either coincide at this stage, or differ by reflection in the y-axis.

□

Let $G := \text{Isom}(\mathbb{R}^2) = \mathbb{R}^2 \rtimes O(2)$, the group of isometries of $\mathbb{R}^2$. We define a group action of G on $\widetilde{T}$ by

$$g(A, B, C) = (g(A), g(B), g(C)), \text{ for } g \in G.$$

By the above proposition (A, B, C) and $(A', B', C') \in \widetilde{T}$ have the same image under p if and only if they are in the same orbit under the action of G. Thus p factors to the map

$$p : G\backslash\widetilde{T} \to T.$$

Since $G\backslash\widetilde{T}$ is given the identification topology, the map p remains continuous, open, and surjective, and now in addition is injective, since points in $\widetilde{T}$ with identical images under p have been identified. Thus the continuous bijection p is an open map, i.e., p is a homeomorphism. Thus

$$G\backslash\widetilde{T} \cong T.$$

To remove the notion of an ordering for our triangles, we define a group action on $G\backslash\widetilde{T}$ of the symmetric group S_3 of permutations of the set $\{1, 2, 3\}$. The action is defined by

$$\sigma \circ [(V_1, V_2, V_3)] = \left[\left(V_{\sigma(1)}, V_{\sigma(2)}, V_{\sigma(3)}\right)\right]$$

for $\sigma \in S_3$ and $(V_1, V_2, V_3) \in \widetilde{T}$. It is clear that this group action is well-defined. We also define a group action of S_3 on T, where $\sigma \in S_3$ likewise acts to permute the triple $(a, b, c) \in T$. The map p factors through these actions of S_3 producing the commutative diagram

$$\begin{array}{ccc} G\backslash\widetilde{T} & \longrightarrow & T \\ \downarrow & & \downarrow \\ S_3\backslash(G\backslash\widetilde{T}) & \longrightarrow & S_3\backslash T \end{array}$$

showing $S_3\backslash(G\backslash\widetilde{T}) \cong S_3\backslash T$.

The space

$$T := \left\{(a,b,c) \in \mathbb{R}^3 | a+b > c, b+c > a, c+a > b\right\}$$

is an open subset of $\mathbb{R}^3$ and can be readily visualized with the help of the following considerations. Observe that the inequalities $a+b > c$ and $b+c > a$ when added together imply $a+2b+c > a+c$, or $b > 0$; and likewise the triangle inequalities imply $a > 0$ and $c > 0$. Thus T lies in the octant of $\mathbb{R}^3$ in which all coordinates (a,b,c) are strictly positive.

Now T is the intersection of the three "half spaces"

$$\begin{aligned} T_1 &= \left\{(a,b,c) \in \mathbb{R}^3 | a+b > c\right\} \\ T_2 &= \left\{(a,b,c) \in \mathbb{R}^3 | b+c > a\right\} \\ T_3 &= \left\{(a,b,c) \in \mathbb{R}^3 | c+a > b\right\}. \end{aligned}$$

T_1 consists of all points below the plane P_1 defined by $a+b=c$, and indeed has T_1 has P_1 as topological boundary (or frontier). Furthermore, we can view $T_1 \cup P_1$ as a 3-dimensional manifold with boundary, the manifold boundary in this case coinciding with the frontier P_1. Likewise:

T_2 consists of all points lying to one-side of the plane P_2 defined by $b+c=a$; T_2 has frontier P_2, and $T_2 \cup P_2$ is a 3-dimensional manifold with boundary P_2.

T_3 consists of all points lying to one-side of the plane P_3 defined by $c + a = b$; T_3 has frontier P_3, and $T_3 \cup P_3$ is a 3-dimensional manifold with boundary P_3.

The planes P_1, P_2, P_3 intersect pairwise in lines

$$P_1 \cap P_2 : b = 0, c = a$$

$$P_2 \cap P_3 : c = 0, a = b$$

$$P_3 \cap P_1 : a = 0, b = c$$

and these three lines intersect at the same point $(0, 0, 0)$.

Thus we can view the space

$$T := \{(a, b, c) \in \mathbb{R}^3 | a + b > c, b + c > a, c + a > b\}$$

as an open set with a boundary $\bar{T} - T$ decomposing into pieces

$$P_1 = \{(a, b, c) \in \mathbb{R}^3 | a + b = c; a, b, c > 0\}$$

$$P_2 = \{(a, b, c) \in \mathbb{R}^3 | b + c = a; a, b, c > 0\}$$

$$P_3 = \{(a, b, c) \in \mathbb{R}^3 | c + a = b; a, b, c > 0\}$$

of dimension 2,

$$P_1 \cap P_2 = \{(a, b, c) \in \mathbb{R}^3 | b = 0, c = a; c, a > 0\}$$

$$P_2 \cap P_3 = \{(a, b, c) \in \mathbb{R}^3 | c = 0, a = b; a, b > 0\}$$

$$P_3 \cap P_1 = \{(a, b, c) \in \mathbb{R}^3 | a = 0, b = c; b, c > 0\}$$

of dimension 1, and the origin $(0, 0, 0)$ of dimension 0.

Recall the map $p : \widetilde{T} \longrightarrow T$ which sends a triangle ABC to its side lengths. The space $\widetilde{T}$ was obtained as triples of vertices $\mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2$ with vertices $\Sigma \subset \mathbb{R}^6$ giving degenerate triangles removed. But notice that the map p is still defined on Σ.

We can decompose the degenerate triangles Σ as

- Σ_0, those triangles for which all three points coincide
- Σ_1, those triangles for which exactly two points coincide
- Σ_2, those triangles determined by three distinct but colinear vertices.

1.2 Moduli spaces

We now say a few words about the concept of a *moduli space*. The precise definition requires the language of category theory and involves objects such as stacks and families. But we can indicate the main ideas in a classical setting, and in particular explain why the space T gives a *fine moduli space* for ordered Euclidean triangles and why the space $S_3\backslash T$ gives a *coarse moduli space* for unordered Euclidean triangles. We follow the discussion of Fulton found in [1].

For a topological space S, a *family* of unordered triangles over S is

(1) a topological space X and a map[1] $X \xrightarrow{p} S$

(2) a metric on each fibre[2] $X_s = p^{-1}(s)$ such that X_s is isometric to some Euclidean triangle (viewed as a 2-dimensional subspace of $\mathbb{R}^2$).

A family of ordered triangles must additionally include

(3) an ordered triple of sections $A, B, C : S \to X$ which specify the vertices $A(s)$, $B(s)$, $C(s)$ of each fibre X_s.

A morphism between families $X \to S$ and $X' \to S'$ is a pair of continuous maps f

[1]The map p must be continuous and proper, and make $X \xrightarrow{p} S$ into a fibre bundle.

[2]Each metric must vary continuously, i.e., it must come from the restriction $d|_{X_s \times X_s}$ of a continuous map $d : X \times_S X \to \mathbb{R}_{\geq 0}$, where $X \times_S X$ is the *fibered product*.

and g making the following diagram commute

$$\begin{array}{ccc} X & \xrightarrow{f} & X' \\ \downarrow & & \downarrow \\ S & \xrightarrow{g} & S' \end{array}$$

and such that, for each $s \in S$, f restricts to an isometry between the fibres X_s and ${X'}_{g(s)}$; for families of ordered triangles, a morphism must additionally respect the sections: $f\big(A(s)\big) = A\big(g(s)\big)$ for all $s \in S$, and likewise for B and C.

What makes T a fine moduli space for ordered triangles is that there is a *universal family* U. It can be constructed as a subspace $U \subset T \times \mathbb{R}^2$ where the fibre over the triple $(a, b, c) \in T$ is the triangle in the plane $\mathbb{R}^2$ with side-lengths a, b, c and first vertex at the origin, second vertex on the positive y-axis, and third vertex having strictly positive x-coordinate.

Given an ordered family of triangles $X \to S$ there is then a canonical map $S \xrightarrow{g} T$ sending $s \in S$ to the unique ordered triple $(a, b, c) \in T$ of side-lengths giving a triangle isometric to X_s, and moreover there is a unique isomorphism of X with the pullback via g of $U \to T$

$$\begin{array}{ccc} X \cong g^*U & & U \\ \downarrow & & \downarrow \\ S & \xrightarrow{g} & T \end{array}$$

Due to the existence of such a universal family, we say that T is a fine moduli space for ordered Euclidean triangles.

Now in the case of unordered Euclidean triangles, there does not exist a universal family over the space $S_3 \backslash T$. The key reason that such a universal family fails to exist is that some pairs of congruent triangles have multiple isometries between them, with no way to distinguish a canonical isometry. Between a congruent pair of isosceles triangles there are two isometries, and between a congruent pair of equilateral triangles, there are six (and notice that imposing an order on the vertices solves this problem). We say that $S_3 \backslash T$ is a coarse moduli space for unordered Euclidean triangles.

Chapter 2

The sub-Riemannian geometry of the Heisenberg group

In this chapter we present background material relating to the sub-Riemannian geometry of the Heisenberg group. The material presented in this chapter is taken from [4]. The goal is to completely describe geodesics between arbitrary distinct points in the Heisenberg group.

2.1 The Heisenberg group $\mathbb{H}$

On $\mathbb{R}^3$ with the usual differentiable structure we introduce the group multiplication law

$$(x_1, y_1, t) \circ (x_2, y_2, s) = \big(x_1 + x_2, y_1 + y_2, t + s - 2(x_1 y_2 - y_1 x_2)\big).$$

The identity element is the origin $O(0,0,0)$ and the inverse of an element is given by $(x, y, t)^{-1} = (-x, -y, -t)$. We frequently use letters P, Q, A, B, C, etc., to refer to elements in the group $(\mathbb{R}^3, \circ)$ and omit the symbol $\circ$ for multiplication. With these explicit formulas it is apparent that the maps

$$(P, Q) \mapsto PQ \text{ from } \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$$

$$P \mapsto P^{-1} \text{ from } \mathbb{R}^3 \to \mathbb{R}^3$$

are smooth, so that $(\mathbb{R}^3, \circ)$ is a Lie group. The Lie group $(\mathbb{R}^3, \circ)$ is called the (symmetric) Heisenberg group with 3 parameters and will be denoted here by $\mathbb{H}$. The word "symmetric" refers to the symmetry apparent in the formula for the inverse of a group element, and is meant to distinguish this characterization of the Heisenberg group from another (Lie group-isomorphic) characterization in which the group is realized as real upper-triangular 3×3 matrices with diagonal entries equal to 1.

At a fixed point $P \in \mathbb{R}^3$ we have the standard basis

$$\left.\frac{\partial}{\partial x}\right|_P, \left.\frac{\partial}{\partial y}\right|_P, \left.\frac{\partial}{\partial t}\right|_P$$

for $T_P\mathbb{R}^3$ which we shorten to $\partial_x|_P$, $\partial_y|_P$, $\partial_t|_P$. Because of the Lie group structure that $\mathbb{H}$ places on $\mathbb{R}^3$ it is more natural to use the basis

$$X|_P := \partial_x|_P + 2y(P)\,\partial_t|_P$$

$$Y|_P := \partial_y|_P - 2x(P)\,\partial_t|_P$$

$$T|_P := \partial_t|_P$$

which coincides with the standard basis at $P = (0,0,0)$ and is in fact the extension of the standard basis at the origin to a left-invariant vector field on $\mathbb{H}$. Thus we have the vector fields

$$X := \partial_x + 2y\partial_t$$

$$Y := \partial_y - 2x\partial_t$$

$$T := \partial_t$$

restricting at each point $P \in \mathbb{H}$ to a basis for $T_P\mathbb{H}$.

Proposition 2.1.1. *The vector fields X, Y, and T are left-invariant vector fields on $\mathbb{H}$.*

Proof. Let $A = (a_1, a_2, a_3) \in \mathbb{H}$. The map left-translation by A

$$L_A : \mathbb{H} \to \mathbb{H}$$

is defined by $L_A(P) = AP$ or in coordinates

$$(x_P, y_P, t_P) \mapsto \big(a_1 + x_P, a_2 + y_P, a_3 + t_P - 2(a_1 y_P - a_2 x_P)\big).$$

The differential can then be written directly

$$DL_A(P) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2a_2 & -2a_1 & 1 \end{bmatrix}.$$

This matrix was calculated using the coordinate system (x, y, t) on $\mathbb{H}$. When we express the vector $X_P \in T_P\mathbb{H}$ in these coordinates we get

$$X_P = \begin{bmatrix} 1 \\ 0 \\ 2y_P \end{bmatrix}$$

and therefore $DL_A(P)$ takes X_P to the vector

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2a_2 & -2a_1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 2y_P \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 2(a_2 + y_P) \end{bmatrix} \in T_{L_A(P)}\mathbb{H}$$

which is precisely $X|_{L_A(P)}$. Likewise

$$Y_P = \begin{bmatrix} 0 \\ 1 \\ -2x_P \end{bmatrix} \mapsto \begin{bmatrix} 0 \\ 1 \\ -2(a_1 + x_P) \end{bmatrix} = Y|_{L_A(P)}$$

and

$$T_P = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = T|_{L_A(P)}.$$

□

The left-invariant vector fields X, Y, T form a basis for the Lie algebra of left-invariant vector fields on $\mathbb{H}$ and restrict to a basis for $T_P\mathbb{H}$ at each point P. We calculate the Lie bracket $[X,Y]$ using the test function $f:\mathbb{R}^3\to\mathbb{R}$

$$\begin{aligned}
[X,Y]f &= X(Yf)-Y(Xf)\\
&= \left(\frac{\partial}{\partial x}+2y\frac{\partial}{\partial t}\right)\left(\frac{\partial f}{\partial y}-2x\frac{\partial f}{\partial t}\right)-\left(\frac{\partial}{\partial y}-2x\frac{\partial}{\partial t}\right)\left(\frac{\partial f}{\partial x}+2y\frac{\partial f}{\partial t}\right)\\
&= \left(\frac{\partial^2 f}{\partial x\partial y}-2\frac{\partial f}{\partial t}-2x\frac{\partial^2 f}{\partial x\partial t}+2y\frac{\partial^2 f}{\partial t\partial y}-4xy\frac{\partial^2 f}{\partial t^2}\right)\\
&\quad-\left(\frac{\partial^2 f}{\partial y\partial x}+2\frac{\partial f}{\partial t}-2x\frac{\partial^2 f}{\partial t\partial x}+2y\frac{\partial^2 f}{\partial y\partial t}-4xy\frac{\partial^2 f}{\partial t^2}\right)\\
&= -4\frac{\partial f}{\partial t}
\end{aligned}$$

so that $[X,Y]=-4T$. Thus X, Y, and $[X,Y]$ are a basis for the Lie algebra of $\mathbb{H}$. This property of X and Y is key to the sub-Riemannian geometry of $\mathbb{H}$.

At each point $P\in\mathbb{H}$ we single out the subspace

$$\mathcal{H}_P=\operatorname{span}\{X|_P\,,Y|_P\}\subset T_P\mathbb{H}.$$

The distribution $P\mapsto\mathcal{H}_P$ is called the horizontal distribution $\mathcal{H}$ on $\mathbb{H}$.

Definition 1. *A differentiable curve $c:[a,b]\to\mathbb{H}$ is called horizontal if*

$$\dot{c}(s)=\left(\frac{dc_1}{ds},\frac{dc_2}{ds},\frac{dc_3}{ds}\right)$$

is contained in $\mathcal{H}_{c(s)}$ for each $s\in[a,b]$, i.e., if the tangent vectors of the curve always lie within the horizontal distribution.

We have the following extremely useful criterion for a curve to be horizontal.

Proposition 2.1.2. *A curve $c(s)=\big(x(s),y(s),t(s)\big)$ in $\mathbb{H}$ is horizontal if and only if*

$$\dot{t}=2(\dot{x}y-x\dot{y}).$$

Proof. Expand

$$\dot{c} = \dot{x}\partial_x + \dot{y}\partial_y + \dot{t}\partial_t$$

$$= \dot{x}\left(\partial_x + 2y\partial_t\right) - 2\dot{x}y\partial_t$$

$$+ \dot{y}\left(\partial y - 2x\partial_t\right) + 2x\dot{y}\partial t$$

$$+ \dot{t}\partial_t$$

$$= \dot{x}X + \dot{y}Y + \left(\dot{t} - 2(\dot{x}y - x\dot{y})\right)\partial_t$$

and the result follows.

□

From this we have the following characterization of tangent vectors for horizontal curves.

Corollary 2.1.3. *A differentiable curve c is horizontal if and only if*

$$\dot{c} = \dot{x}X + \dot{y}Y.$$

Proposition 2.1.4. *For $c(s)$ a horizontal curve and $A = (a_1, a_2, a_3) \in \mathbb{H}$ the curve $\bar{c}(s) = L_A c(s)$ is also horizontal.*

Proof. Write $c = (c_1, c_2, c_3)$ and $\bar{c} = (\bar{c}_1, \bar{c}_2, \bar{c}_3)$. Then

$$\bar{c}_1(s) = a_1 + c_1(s)$$

$$\bar{c}_2(s) = a_2 + c_2(s)$$

so that

$$\dot{\bar{c}}_1(s) = \dot{c}_1(s)$$

$$\dot{\bar{c}}_2(s) = \dot{c}_2(s).$$

Then we compute

$$\dot{\bar{c}}_3 = \frac{d}{ds}[a_3 + c_3(s) - 2\big(a_1c_2(s) - a_2c_1(s)\big)]$$

$$= \dot{c}_3 - 2\left(a_1\dot{c}_2 - a_2\dot{c}_1\right)$$

$$= 2(\dot{c}_1c_2 - c_1\dot{c}_2) - 2\left(a_1\dot{c}_2 - a_2\dot{c}_1\right)$$

$$= 2\big(\dot{c}_1(a_2 + c_2) - (a_1 + c_1)\dot{c}_2\big)$$

$$= 2(\dot{\bar{c}}_1\bar{c}_2 - \bar{c}_1\dot{\bar{c}}_2).$$

□

Proposition 2.1.5. *For $c(s) = (x(s), y(s), t(s))$ a horizontal curve and $0 \le \varphi \le 2\pi$ the curve*

$$\bar{c}(s) = \left(\begin{bmatrix}\cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi\end{bmatrix}\begin{bmatrix}x(s) \\ y(s)\end{bmatrix}, t(s)\right)$$

is also horizontal.

Proof. The form of the equation for $\bar{c}$ above emphasises that we are taking c and rotating it by a fixed angle φ about the t-axis while keeping the height of the curve fixed at each point. We will write $R_\varphi c$ to represent this transformation of c. Explicitly the transformation is

$$\bar{x}(s) = x(s)\cos\varphi - y(s)\sin\varphi$$

$$\bar{y}(s) = x(s)\sin\varphi + y(s)\cos\varphi$$

so that

$$\dot{\bar{x}}\bar{y} - \bar{x}\dot{\bar{y}} = \left(\dot{x}\cos\varphi - \dot{y}\sin\varphi\right)\left(x\sin\varphi + y\cos\varphi\right)$$

$$- \left(x\cos\varphi - y\sin\varphi\right)\left(\dot{x}\sin\varphi + \dot{y}\cos\varphi\right)$$

$$= \dot{x}y - x\dot{y}$$

and then since $\bar{t} = t$ the horizontality condition $\dot{t} = 2(\dot{x}y - x\dot{y})$ on the original curve implies the horizontality condition $\dot{\bar{t}} = 2(\dot{\bar{x}}\bar{y} - \bar{x}\dot{\bar{y}})$ on the transformed curve.

□

Now at each point $P \in \mathbb{H}$ we put an inner product g on the subspace $\mathcal{H}_P$ by making the vectors X and Y orthonormal

$$g(X,X) = 1 \qquad g(X,Y) = 0 \qquad g(Y,Y) = 1.$$

Thus two vectors $\vec{v} = v_1 X + v_2 Y$ and $\vec{w} = w_1 X + w_2 Y$ in $\mathcal{H}_P$ have inner product

$$\begin{aligned} g(\vec{v},\vec{w}) &= g(v_1 X + v_2 Y, w_1 X + w_2 Y) \\ &= v_1 w_1 g(X,X) + v_1 w_2 g(X,Y) + v_2 w_1 g(Y,X) + v_2 w_2 g(Y,Y) \\ &= v_1 w_1 + v_2 w_2. \end{aligned}$$

These inner products on each $\mathcal{H}_P$ provide what is known as a sub-Riemannian metric for the distribution $\mathcal{H}$ on $\mathbb{H}$. The sub-Riemannian metric defines a length for horizontal vectors

$$|\vec{v}| = \sqrt{g(\vec{v},\vec{v})} = \sqrt{v_1^2 + v_2^2}$$

which induces a length for horizontal curves.

Definition 2. *The length of a horizontal curve $c : [a,b] \to \mathbb{H}$ is*

$$\int_a^b |\dot{c}(s)| ds$$

where $|\dot{c}(s)|$ is computed using the sub-Riemannian metric.

Proposition 2.1.6. *For horizontal $c : [a,b] \to \mathbb{H}$ and $A = (a_1, a_2, a_3) \in \mathbb{H}$ the curves $c(s)$ and $\bar{c}(s) = L_A c(s)$ have the same length.*

Proof. We have seen that $\dot{\bar{c}}_1 = \dot{c}_1$ and $\dot{\bar{c}}_2 = \dot{c}_2$. Thus

$$\dot{c}(s) = \dot{c}_1(s) X + \dot{c}_2(s) Y$$

$$\dot{\bar{c}}(s) = \dot{\bar{c}}_1(s) X + \dot{\bar{c}}_2(s) Y$$

have identical components so that

$$|\dot{\bar{c}}(s)| = |\dot{c}(s)|$$

for each $s \in [a, b]$. It follows that the length of $\bar{c}$ equals the length of c.

□

Proposition 2.1.7. *For horizontal $c : [a, b] \to \mathbb{H}$ and $0 \le \varphi \le 2\pi$ the curves $c(s)$ and $\bar{c}(s) = R_\varphi c(s)$ have the same length.*

Proof. Writing $c(s) = \big(x(s), y(s), t(s)\big)$ we have seen that

$$\dot{\bar{x}}(s) = \dot{x}(s)\cos\varphi - \dot{y}(s)\sin\varphi$$

$$\dot{\bar{y}}(s) = \dot{x}(s)\sin\varphi + \dot{y}(s)\cos\varphi$$

so that

$$\begin{aligned}
|\dot{\bar{c}}|^2 &= \dot{\bar{x}}^2 + \dot{\bar{y}}^2 \\
&= \dot{x}^2\cos^2\varphi - 2\dot{x}\dot{y}\sin\varphi\cos\varphi + \dot{y}^2\sin^2\varphi \\
&\quad + \dot{x}^2\sin^2\varphi + 2\dot{x}\dot{y}\sin\varphi\cos\varphi + \dot{y}^2\cos^2\varphi \\
&= \dot{x}^2 + \dot{y}^2 \\
&= |\dot{c}|^2
\end{aligned}$$

and therefore the length of $\bar{c}$ equals the length of c.

□

2.2 Geodesics in $\mathbb{H}$

Geodesics in $\mathbb{H}$ are defined in terms of a Hamiltonian function, that is, a real-valued function on the cotangent bundle of $\mathbb{H}$. We will see that the geodesics produced in this way have the natural properties associated to geodesics, such as being length-minimizing with respect to the sub-Riemannian metric.

Viewing $\mathbb{H}$ as $\mathbb{R}^3$ we introduce the function

$$H : T^*\mathbb{H} \to \mathbb{R}$$

defined by

$$(x_1, x_2, t, \xi_1, \xi_2, \theta) \mapsto \frac{1}{2}(\xi_1 + 2x_2\theta)^2 + \frac{1}{2}(\xi_2 - 2x_1\theta)^2$$

where $(x_1, x_2, t, \xi_1, \xi_2, \theta)$ are just Euclidean coordinates on $\mathbb{R}^6 \cong T^*\mathbb{H}$. We recall Hamilton's equations

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}$$

where the coordinates q and p are generalized position and momentum, respectively. For $T^*\mathbb{H} \cong \mathbb{R}^6$ the coordinates yield six specific equations

$$\dot{x}_i = \frac{\partial H}{\partial \xi_i} \qquad \dot{\xi}_i = -\frac{\partial H}{\partial x_i}$$

$$\dot{t} = \frac{\partial H}{\partial \theta} \qquad \dot{\theta} = -\frac{\partial H}{\partial t}$$

(where $i = 1, 2$) which we proceed to write explicitly. Now

$$\dot{x}_1 = \frac{\partial H}{\partial \xi_1} = \xi_1 + 2x_2\theta$$

$$\dot{x}_2 = \frac{\partial H}{\partial \xi_2} = \xi_2 - 2x_1\theta$$

then

$$\dot{t} = \frac{\partial H}{\partial \theta} = (\xi_1 + 2x_2\theta)2x_2 - (\xi_2 - 2x_1\theta)2x_1$$

$$= 2(\dot{x}_1 x_2 - x_1\dot{x}_2)$$

then

$$\dot{\xi}_1 = -\frac{\partial H}{\partial x_1} = -[(\xi_2 - 2x_1\theta)(-2\theta)] = 2\theta\dot{x}_2$$

$$\dot{\xi}_2 = -\frac{\partial H}{\partial x_2} = -[(\xi_1 + 2x_2\theta)(2\theta)] = -2\theta\dot{x}_1$$

and

$$\dot{\theta} = -\frac{\partial H}{\partial t} = 0.$$

Altogether we have obtained the system of equations

$$\begin{aligned} \dot{x}_1 &= \xi_1 + 2x_2\theta \\ \dot{x}_2 &= \xi_2 - 2x_1\theta \\ \dot{t} &= 2(\dot{x}_1 x_2 - x_1 \dot{x}_2) \\ \dot{\xi}_1 &= 2\theta\dot{x}_2 \\ \dot{\xi}_2 &= -2\theta\dot{x}_1 \\ \dot{\theta} &= 0. \end{aligned} \tag{2.2.1}$$

Definition 3. *A curve in the cotangent bundle* $T^*\mathbb{H}$

$$\bar{c}(s) = \big(x_1(s), x_2(s), t(s), \xi_1(s), \xi_2(s), \theta(s)\big)$$

that satisfies the system 2.2.1 is called a bicharacteristic.

Definition 4. *A curve* $c(s)$ *in* $\mathbb{H}$ *is called a geodesic if it is the projection of a bicharacteristic curve, i.e., if there exists a bicharacteristic curve* $\bar{c}(s)$ *in* $T^*\mathbb{H}$ *such that* $c(s) = \pi \circ \bar{c}(s)$ *where* $\pi : T^*\mathbb{H} \longrightarrow \mathbb{H}$ *is the projection.*

From the definition of a geodesic and from the system 2.2.1 we immediately have the following.

Proposition 2.2.1. *A geodesic is a horizontal curve.*

We proceed to determine further properties of geodesics. We first observe that

$\dot{\theta} = 0$ implies θ is constant. Then from 2.2.1 we have

$$\dot{x}_1 = \xi_1 + 2x_2\theta$$

$$\ddot{x}_1 = \dot{\xi}_1 + 2\theta\dot{x}_2$$

$$= 2\theta\dot{x}_2 + 2\theta\dot{x}_2$$

$$= 4\theta\dot{x}_2$$

and similarly

$$\ddot{x}_2 = -4\theta\dot{x}_1.$$

We can write this as

$$\begin{bmatrix} \ddot{x}_1(s) \\ \ddot{x}_2(s) \end{bmatrix} = 4\theta \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} \dot{x}_1(s) \\ \dot{x}_2(s) \end{bmatrix}$$

or more succinctly as

$$\ddot{x}(s) = 4\theta J\dot{x}(s)$$

where $J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ is the matrix corresponding to clockwise rotation by $\pi/2$.

Next suppose that $\dot{x}(s) = e^{4\theta Js}\dot{x}(0)$. Then

$$\ddot{x}(s) = 4\theta Je^{4\theta Js}\dot{x}(0)$$

$$= 4\theta J\dot{x}(s).$$

What this shows is that $\dot{x}(s) = e^{4\theta Js}\dot{x}(0)$ is a solution to the equation

$$\ddot{x}(s) = 4\theta J\dot{x}(s).$$

We solve for $x(s)$

$$x(s) = x(0) + \int_0^s e^{4\theta J u}\dot{x}(0)du$$

$$= x(0) + \left(\frac{1}{4\theta J}e^{4\theta J u}\dot{x}(0)\Big|_0^s\right)$$

$$= x(0) + \left(\frac{1}{4\theta}J^{-1}e^{4\theta J u}\dot{x}(0)\Big|_0^s\right)$$

$$= x(0) + \left(-\frac{1}{4\theta}Je^{4\theta J u}\dot{x}(0)\Big|_0^s\right)$$

$$= x(0) + \frac{1}{4\theta}J\dot{x}(0) - \frac{1}{4\theta}Je^{4\theta J s}\dot{x}(0).$$

Recalling that J and $e^{4\theta J s}$ commute, we can write this as .

$$x(s) = e^{4\theta J s}\left(-\frac{1}{4\theta}J\dot{x}(0)\right) + x(0) + \frac{1}{4\theta}J\dot{x}(0)$$

$$= e^{4\theta J s}K + C$$

where

$$K = -\frac{1}{4\theta}J\dot{x}(0)$$

$$C = x(0) - K.$$

We will need the following computation:

Proposition 2.2.2. *For $0 \leq \theta \leq 2\pi$, we have*

$$e^{4\theta J s} = R_{4\theta s}$$

where $R_{4\theta s}$ is the matrix corresponding to clockwise rotation by an angle of $4\theta s$.

Proof. With $J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ notice the pattern

$$J^1 = J,$$

$$J^2 = -I,$$

$$J^3 = -J,$$

$$J^4 = I, \text{ etc.}$$

We compute

$$\begin{aligned}
e^{4\theta Js} &= \sum_{n=0}^{\infty} \frac{(4\theta Js)^n}{n!} \\
&= \sum_{n=0}^{\infty} \frac{(4\theta s)^n J^n}{n!} \\
&= \sum_{k=0}^{\infty} \frac{(4\theta s)^{4k} I}{(4k)!} + \sum_{k=0}^{\infty} \frac{(4\theta s)^{4k+1} J}{(4k+1)!} + \sum_{k=0}^{\infty} \frac{(4\theta s)^{4k+2}(-I)}{(4k+2)!} + \sum_{k=0}^{\infty} \frac{(4\theta s)^{4k+3}(-J)}{(4k+3)!} \\
&= \sum_{k=0}^{\infty} \begin{bmatrix} \frac{(4\theta s)^{4k}}{(4k)!} - \frac{(4\theta s)^{4k+2}}{(4k+2)!} & \frac{(4\theta s)^{4k+1}}{(4k+1)!} - \frac{(4\theta s)^{4k+3}}{(4k+3)!} \\ -\frac{(4\theta s)^{4k+1}}{(4k+1)!} + \frac{(4\theta s)^{4k+3}}{(4k+3)!} & \frac{(4\theta s)^{4k}}{(4k)!} - \frac{(4\theta s)^{4k+2}}{(4k+2)!} \end{bmatrix} \\
&= \begin{bmatrix} \cos(4\theta s) & \sin(4\theta s) \\ -\sin(4\theta s) & \cos(4\theta s) \end{bmatrix} \\
&= R_{4\theta s}.
\end{aligned}$$

□

Now that we have determined the solution $x(s) = R_{4\theta s}K + C$ we compute $t(s)$

using the horizontality condition

$$\begin{aligned}
\dot{t}(s) &= 2\left(\dot{x}_1(s)x_2(s) - x_1(s)\dot{x}_2(s)\right) \\
&= 2\left\langle \begin{bmatrix} x_1(s) \\ x_2(s) \end{bmatrix}, \begin{bmatrix} -\dot{x}_2(s) \\ \dot{x}_1(s) \end{bmatrix} \right\rangle \\
&= 2\left\langle x(s), -J\dot{x}(s) \right\rangle \\
&= 2\left\langle e^{4\theta Js}K + C, -J\left(4\theta J e^{4\theta Js}K\right) \right\rangle \\
&= 2\left\langle e^{4\theta Js}K + C, 4\theta e^{4\theta Js}K \right\rangle \\
&= 8\theta \left\langle e^{4\theta Js}K, e^{4\theta Js}K \right\rangle + 8\theta\left\langle C, e^{4\theta Js}K \right\rangle \\
&= 8\theta\left|K\right|^2 + 8\theta\left\langle C, e^{4\theta Js}K \right\rangle .
\end{aligned}$$

In order to integrate $\dot{t}(s)$ notice that

$$\begin{aligned}
\frac{d}{ds}\left\langle JC, e^{4\theta Js}K \right\rangle &= \left\langle JC, \frac{d}{ds} e^{4\theta Js}K \right\rangle \\
&= \left\langle JC, 4\theta J e^{4\theta Js}K \right\rangle \\
&= 4\theta\left\langle J^T JC, e^{4\theta Js}K \right\rangle \\
&= 4\theta\left\langle C, e^{4\theta Js}K \right\rangle .
\end{aligned}$$

Then

$$\begin{aligned}
t(s) &= \int 8\theta\left|K\right|^2 + 8\theta\left\langle C, e^{4\theta Js}K \right\rangle ds \\
&= 8\theta\left|K\right|^2 s + 2\left\langle JC, e^{4\theta Js}K \right\rangle + C_1
\end{aligned}$$

where

$$C_1 = t(0) - 2\left\langle JC, K \right\rangle .$$

We summarize these results.

Proposition 2.2.3. *For a geodesic $(x(s), t(s))$ with constant parameter θ and initial conditions*

$$(x(0), t(0)) \qquad\qquad \dot{x}(0)$$

given, the equation is given by

$$x(s) = R_{4\theta s}K + C$$

$$t(s) = 8\theta\,|K|^2\,s + 2\left\langle JC, e^{4\theta Js}K\right\rangle + C_1$$

where

$$K = -\frac{1}{4\theta}J\dot{x}(0)$$

$$C = x(0) - K$$

$$C_1 = t(0) - 2\,\langle JC, K\rangle\,.$$

Proposition 2.2.4. *The image of a geodesic $(x_1(s), x_2(s), t(s))$ projected onto the x_1x_2-plane is a circle.*

Proof. The equation $x(s) = R_{4\theta s}K + C$ is the equation for a circle centered at C of radius $|K|$.

□

Note that when such a geodesic is parametrized by arc-length we have

$$|K| = \left|-\frac{1}{4\theta}J\dot{x}(0)\right| = \frac{1}{4\theta}$$

because $|\dot{x}(0)| = 1$.

We have defined the matrix $J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ and seen that a geodesic $(x(s), t(s))$

satisfies the equation $\ddot{x}(s) = 4\theta J\dot{x}(s)$. Now geodesics are horizontal so that

$$\dot{c} = \dot{x}_1 X + \dot{x}_2 Y$$

$$\ddot{c} = \ddot{x}_1 X + \ddot{x}_2 Y.$$

If we think of J as the operator

$$X \mapsto -Y$$

$$Y \mapsto X$$

then a geodesic satisfies $\ddot{c} = 4\theta J\dot{c}$. In fact this condition characterizes geodesics.

Proposition 2.2.5. *A curve c is a geodesic if and only if (1) c is horizontal and (2) $\ddot{c} = 4\theta J\dot{c}$.*

Proof. One direction is clear. We have seen already that a geodesic is horizontal and satisfies $\ddot{c} = 4\theta J\dot{c}$.

Now suppose we are given a horizontal curve $c(s) = (x_1(s), x_2(s), t(s))$ such that $\ddot{c} = 4\theta J\dot{c}$. Define the curve $\bar{c}$ in $T^*\mathbb{H}$ by

$$\bar{c}(s) = (x_1(s), x_2(s), t(s), \xi_1(s), \xi_2(s), \theta)$$

where

$$\xi_1(s) = \dot{x}_1(s) - 2x_2(s)\theta$$

$$\xi_2(s) = \dot{x}_2(s) + 2x_1(s)\theta.$$

Then $\bar{c}$ is a bicharacteristic curve in $T^*\mathbb{H}$ and projects to c in $\mathbb{H}$. Therefore c is a geodesic.

□

This criterion for a geodesic is very useful.

Proposition 2.2.6. *If $c = (c_1, c_2, c_3)$ is a geodesic with parameter θ then for $a \in \mathbb{H}$ arbitrary the left-translation $\bar{c} = L_a c$ is a geodesic with parameter θ.*

Proof. The left-translation of a horizontal curve is horizontal. Thus $L_a c$ is horizontal. Writing $a = (a_1, a_2, a_3)$ we have

$$\bar{c}_1 = a_1 + c_1 \Rightarrow \dot{\bar{c}}_1 = \dot{c}_1$$

$$\bar{c}_2 = a_2 + c_2 \Rightarrow \dot{\bar{c}}_2 = \dot{c}_2$$

and since for horizontal curves $\dot{c} = \dot{c}_1 X + \dot{c}_2 Y$ we see that the tangent vectors for c and for $\bar{c}$ have identical components. Thus $\bar{c}$ satisfies the same geodesic criterion $\ddot{\bar{c}} = 4\theta J \dot{\bar{c}}$ as c.

□

Proposition 2.2.7. *If $c = (x, y, t)$ is a geodesic with parameter θ then for $0 \leq \varphi \leq 2\pi$ the rotated curve $\bar{c} = R_\varphi c$ is a geodesic with parameter θ.*

Proof. From

$$\dot{\bar{x}}(s) = \dot{x}(s) \cos\varphi - \dot{y}(s) \sin\varphi$$

$$\dot{\bar{y}}(s) = \dot{x}(s) \sin\varphi + \dot{y}(s) \cos\varphi$$

and from the geodesic criterion on c

$$\ddot{x} = 4\theta \dot{y}$$

$$\ddot{y} = -4\theta \dot{x}$$

we have

$$\ddot{\bar{x}} = \ddot{x} \cos\varphi - \ddot{y} \sin\varphi$$

$$= (4\theta\dot{y}) \cos\varphi - (-4\theta\dot{x}) \sin\varphi$$

$$= 4\theta\dot{\bar{y}}$$

and likewise $\ddot{y} = -4\theta\dot{x}$. Therefore $\bar{c}$ satisfies the same geodesic criterion as c.

□

Proposition 2.2.8. *The speed of a geodesic is constant.*

Proof. For a geodesic $c = (x, y, t)$ we have $|\dot{c}|^2 = \dot{x}^2 + \dot{y}^2$. But

$$\frac{d}{ds}\left(\dot{x}^2 + \dot{y}^2\right) = 2\dot{x}\ddot{x} + 2\dot{y}\ddot{y}$$

$$= 2\dot{x}(4\theta\dot{y}) + 2\dot{y}(-4\theta\dot{x}) = 0$$

and it follows that $|\dot{c}|$ is constant.

□

2.3 The Lagrangian picture

There is also a Lagrangian picture of geodesics that is useful. The Lagrangian $L : T\mathbb{H} \to \mathbb{R}$ is defined by

$$L(x_1, x_2, t, \dot{x}_1, \dot{x}_2, \dot{t}) = \frac{1}{2}(\dot{x}_1^2 + \dot{x}_2^2) + \theta(\dot{t} - 2\dot{x}_1 x_2 + 2x_1\dot{x}_2).$$

Normally the Lagrangian is a function only of the coordinates of the tangent bundle $T\mathbb{H}$ but in our case the θ-coordinate in $T^*\mathbb{H}$ shows up in the Lagrangian. Thus when one wants to apply the Lagrangian to a curve in $\mathbb{H}$ (which with its velocity vectors can be treated as a curve in $T\mathbb{H}$) one must carry along the θ parameter as extra data for that curve. Since our curves will all be geodesics with constant θ we do not have to address the subtleties surrounding this issue. We merely treat the Hamiltonian as fundamental and use the Lagrangian picture for its usefulness.

Recall the Euler-Lagrange equations

$$\frac{d}{ds}\frac{\partial L}{\partial \dot{c}} = \frac{\partial L}{\partial c}.$$

We compute

$$\frac{d}{ds}\frac{\partial L}{\partial \dot{x}_1} = \ddot{x}_1 - 2\theta\dot{x}_2 \qquad \frac{\partial L}{\partial x_1} = 2\theta\dot{x}_2$$

$$\frac{d}{ds}\frac{\partial L}{\partial \dot{x}_2} = \ddot{x}_2 + 2\theta\dot{x}_1 \qquad \frac{\partial L}{\partial x_2} = -2\theta\dot{x}_1$$

$$\frac{d}{ds}\frac{\partial L}{\partial \dot{t}} = \dot{\theta} \qquad \frac{\partial L}{\partial t} = 0.$$

In calculating the first two rows we have used the fact that $\dot{\theta} = 0$, which comes from the third line. Altogether we have the system

$$\ddot{x}_1 = 4\theta\dot{x}_2$$

$$\ddot{x}_2 = -4\theta\dot{x}_1$$

$$\dot{\theta} = 0.$$

We recognize this condition as the geodesic criterion:

Proposition 2.3.1. *A geodesic satisfies the Euler-Lagrange equations.*

2.4 Connectivity by geodesics

We want to connect arbitrary distinct points $P, Q \in \mathbb{H}$ by a length-minimizing geodesic. We will show that, up to parametrization, there exists a unique such geodesic connecting P to Q. First, we apply the left translation $L_{P^{-1}}$ taking P to O and taking Q to $P^{-1}Q$. Under this translation, a geodesic c connecting P to Q becomes a geodesic $\bar{c} = L_{P^{-1}}c$ connecting O to $P^{-1}Q$, and c and $\bar{c}$ have the same length. Hence we need study only geodesics with initial point O.

It is best to work in cylindrical coordinates

$$x_1 = r\cos\phi$$

$$x_2 = r\sin\phi$$

$$t = t.$$

We compute the Lagrangian in these coordinates

$$\dot{x}_1 = \dot{r}\cos\phi + r(-\sin\phi)\dot{\phi}$$

$$\dot{x}_1^2 = \dot{r}^2\cos^2\phi - 2r\dot{r}(\sin\phi)(\cos\phi)\dot{\phi} + r^2\dot{\phi}^2\sin^2\phi$$

$$\dot{x}_2 = \dot{r}\sin\phi + r(\cos\phi)\dot{\phi}$$

$$\dot{x}_2^2 = \dot{r}^2\sin^2\phi + 2r\dot{r}(\sin\phi)(\cos\phi)\dot{\phi} + r^2\dot{\phi}^2\cos^2\phi$$

$$-2\dot{x}_1x_2 + 2x_1\dot{x}_2 = -2(r\dot{r}\sin\phi\cos\phi - r^2(\sin^2\phi)\dot{\phi})$$

$$+ 2(r\dot{r}\cos\phi\sin\phi + r^2(\cos^2\phi)\dot{\phi}) = 2r^2\dot{\phi}$$

so that L is given by

$$L = \frac{1}{2}(\dot{x}_1^2 + \dot{x}_2^2) + \theta(\dot{t} - 2\dot{x}_1x_2 + 2x_1\dot{x}_2)$$

$$= \frac{1}{2}(\dot{r}^2 + r^2\dot{\phi}^2) + \theta(\dot{t} + 2r^2\dot{\phi}).$$

Next we calculate the Euler-Lagrange equations in polar coordinates

$$\frac{d}{ds}\frac{\partial L}{\partial \dot{r}} = \ddot{r} \qquad \frac{\partial L}{\partial r} = r\dot{\phi}^2 + 4r\theta\dot{\phi} = r\dot{\phi}(\dot{\phi} + 4\theta)$$

$$\frac{d}{ds}\frac{\partial L}{\partial \dot{\phi}} = \frac{d}{ds}(r^2\dot{\phi} + 2\theta r^2) \qquad \frac{\partial L}{\partial \phi} = 0$$

$$\frac{d}{ds}\frac{\partial L}{\partial \dot{t}} = \dot{\theta} \qquad \frac{\partial L}{\partial t} = 0$$

to obtain the system

$$\ddot{r} = r\dot{\phi}(\dot{\phi} + 4\theta)$$

$$r^2(\dot{\phi} + 2\theta) \equiv C_0 \text{ constant} \tag{2.4.1}$$

$$\theta \equiv \theta_0 \text{ constant.}$$

For our geodesics, which start at O, we get $r(0) = 0 \Rightarrow C_0 = 0$. Next if r remains identically 0 on any non-trivial interval around $s = 0$ we must have $\dot{x}_1 = 0 = \dot{x}_2$ so that $\dot{t} = 2(\dot{x}_1 x_2 - x_1 \dot{x}_2) = 0$. A geodesic with this property starting at O will thus remain stuck at O. Since we want to classify geodesics connecting distinct points we can thus assume that r cannnot stay at 0 over any non-trivial interval whereupon we have

$$r^2(\dot{\phi} + 2\theta) \equiv 0 \Rightarrow \dot{\phi} + 2\theta \equiv 0.$$

Thus, for geodesics starting at the origin, the system 2.4.1 sharpens to

$$\ddot{r} = -4\theta^2 r$$

$$\dot{\phi} = -2\theta \tag{2.4.2}$$

$$\theta \equiv \theta_0 \text{ constant.}$$

There is not a uniform description of geodesics from $O(0,0,0)$ to distinct points $P(x,y,t)$. Rather the description depends on the nature of the point $P(x,y,t)$. The possibilities are

- $P(x,y,t)$ is on the t-axis so that $x^2 + y^2 = 0$ and $t \neq 0$
- $P(x,y,t)$ is in the xy-plane so that $x^2 + y^2 \neq 0$ and $t = 0$
- $P(x,y,t)$ is not in the xy-plane or on the t-axis so that $x^2 + y^2 \neq 0$ and $t \neq 0$.

2.4.1 Geodesics from O to $P(0,0,t)$ with $t \neq 0$

We repeat the Euler-Lagrange equations for a geodesic with initial point O ...

$$\ddot{r} = -4\theta^2 r$$

$$\dot{\phi} = -2\theta$$

$$\theta \equiv \theta_0 \text{ constant.}$$

The general solution to

$$\ddot{r}(s) = -4\theta^2 r(s)$$

is

$$r(s) = A\sin(2\theta s) + B\cos(2\theta s).$$

Since $r(0) = 0$ we have $B = 0$ so that

$$r(s) = A\sin(2\theta s).$$

We do not parametrize by arc-length but instead assume our geodesic is such that $\gamma(0) = O$ and $\gamma(1) = P$. Then we have $r(1) = 0$ so that $2\theta = m\pi$ for some integer m. We will say more about the sign of the integer m shortly.

Next from $\dot{\phi} = -2\theta$ we get $\phi(s) - \phi(0) = -2\theta s$ so that

$$\phi_1 = \phi_0 - 2\theta = \phi_0 - m\pi.$$

Recall $\dot{t} = -2r^2\dot{\phi}$. Thus

$$
\begin{aligned}
t &= t(1) - t(0) \\
&= \int_{\phi_0}^{\phi_0 - m\pi} -2r^2 d\phi \\
&= -2\int_{\phi_0}^{\phi_0 - m\pi} A^2 \sin^2(\phi - \phi_0) d\phi \\
&= -2\int_{0}^{-m\pi} \sin^2 u du \\
&= A^2 m\pi.
\end{aligned}
$$

This gives $A^2 = \frac{t}{m\pi} = \frac{t}{2\theta}$. This means that when $t > 0$ the integer m must be positive and when $t < 0$ the integer m must be negative. Similarly $\theta > 0$ when $t > 0$ and $\theta < 0$ when $t < 0$.

Recall that for geodesics the quantity

$$|\dot{\gamma}|^2 = \dot{x}_1^2 + \dot{x}_2^2 = \dot{r}^2 + 4r^2\theta^2$$

is constant. Now when $r(s) = A\sin(2\theta s)$ hits its maximum value of A we will have $\dot{r} = 0$. This gives

$$
\begin{aligned}
|\dot{\gamma}|^2 = 4A^2\theta^2 &= 4\left(\frac{t}{2\theta}\right)\theta^2 \\
&= 2\theta t = m\pi t.
\end{aligned}
$$

Therefore $|\dot{\gamma}| = \sqrt{m\pi t}$ so that the arc-length $\tau = \sqrt{m\pi t}$ because the parameter s is on $[0, 1]$.

For the t-coordinate we have

$$t(\phi) = \int_{\phi_0}^{\phi} -2r^2 d\phi$$

$$= \int_{\phi_0}^{\phi} -2A^2 \sin^2(\phi - \phi_0) d\phi$$

$$= -2A^2 \int_{0}^{\phi-\phi_0} \sin^2 u du$$

$$= -\frac{2t}{m\pi}\left[\frac{1}{2}u - \frac{1}{4}\sin 2u\right]_0^{\phi-\phi_0}$$

$$= \frac{t}{2m\pi}\left[\sin 2(\phi - \phi_0) - 2(\phi - \phi_0)\right].$$

Theorem 2.4.1. *Between the points $O(0,0,0)$ and $P(0,0,t)$ with $t > 0$ the equation for a geodesic from O to P leaving the origin at an angle of ϕ_0 with the positive x-axis is given by*

$$r(s) = \sqrt{\frac{t}{m\pi}}\sin(2\theta s)$$

$$\phi(s) = \phi_0 - 2\theta s$$

$$t(s) = \frac{t}{2m\pi}\left(4\theta s - \sin(4\theta s)\right)$$

where $\theta = \frac{m\pi}{2}$. There is one such geodesic for each $m = 1, 2, \ldots$. The arc-length of each geodesic is $\sqrt{m\pi t}$.

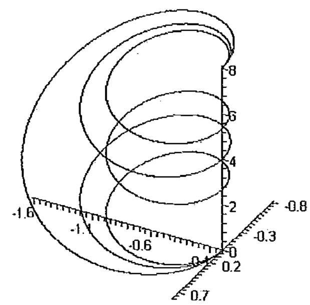


Figure 2.1: geodesics to $(0,0,t)$ for $m = 1,2,3$

Theorem 2.4.2. *A length-minimizing geodesic from $O(0,0,0)$ to $P(0,0,t)$ with $t > 0$ has arc-length $\sqrt{\pi t}$.*

There are of course analogous statements to describe the case when $t < 0$.

Observe that if γ is any geodesic connecting $O(0,0,0)$ to $P(0,0,t)$ then $R_\varphi c$ is another geodesic connecting $O(0,0,0)$ to $P(0,0,t)$ of equal length for any $0 \leq \varphi \leq 2\pi$. Thus for each $m = 1, 2, \ldots$, we obtain a family of geodesics from O to P of equal length $\sqrt{m\pi t}$ parametrized by $0 \leq \varphi \leq 2\pi$.

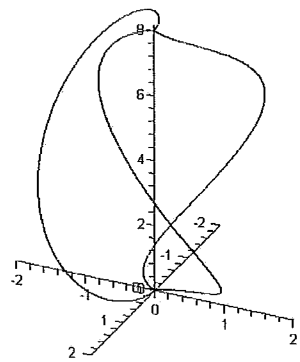


Figure 2.2: a one-parameter family of geodesics with $\varphi = 0, 2\pi/3, 4\pi/3$ shown

2.4.2 Geodesics from O to $P(x, y, 0)$ with $x^2 + y^2 \neq 0$

From the horizontality condition we have

$$\dot{t} = 2(\dot{x}_1 x_2 - x_1 \dot{x}_2) = -2r^2 \dot{\phi}.$$

But $\dot{\phi} = -2\theta$ so $\dot{t} = 4\theta r^2$.

Since θ is constant we see $\theta > 0$ implies $\dot{t} > 0$ which implies $t(s)$ is always increasing. Likewise $\theta < 0$ implies $\dot{t} < 0$ which implies $t(s)$ is always decreasing. Therefore to end up at a point $P(x, y, 0)$ the geodesic must have $\theta = 0$. The Euler-

Lagrange equations reduce to

$$\ddot{r} = 0$$

$$\dot{\phi} = 0$$

$$\theta = 0.$$

Thus we have constant angle ϕ_0 and constant $\dot{r} = \lambda \neq 0$. Since our geodesic starts at the origin the equation is

$$r(s) = \lambda s$$

$$\phi(s) = \phi_0$$

$$t(s) = 0.$$

In order for $P(x, y, 0)$ to be the terminal point we must take ϕ_0 to be a solution to

$$\phi_0 = \arctan\left(\frac{y}{x}\right).$$

The sign of λ depends on the solution we choose above; the magnitude of λ determines the rate at which we move along the geodesic. To parametrize by arc-length we choose $\lambda = \pm 1$ and either way the result is the unique arc-length parametrized geodesic from O to P.

Theorem 2.4.3. *Between the origin $O(0,0,0)$ and a point $P(x, y, 0)$ in the xy-plane with $x^2 + y^2 \neq 0$ there exists a unique arc-length parametrized geodesic with equation*

$$x(s) = \frac{x}{\sqrt{x^2 + y^2}} s$$

$$y(s) = \frac{y}{\sqrt{x^2 + y^2}} s$$

$$t(s) = 0$$

where $0 \leq s \leq \sqrt{x^2 + y^2}$ and the arc-length is $\tau = \sqrt{x^2 + y^2}$.

2.4.3 Gaveau's function

Next we will describe geodesics connecting $O(0,0,0)$ to points $P(x,y,t)$ with $x^2 + y^2 \neq 0$ and $t \neq 0$. The classification of such geodesics will depend crucially on the properties of Gaveau's function, which we now discuss.

The function

$$\mu(x) = \frac{x - \sin x \cos x}{\sin^2 x}$$

is called Gaveau's function.

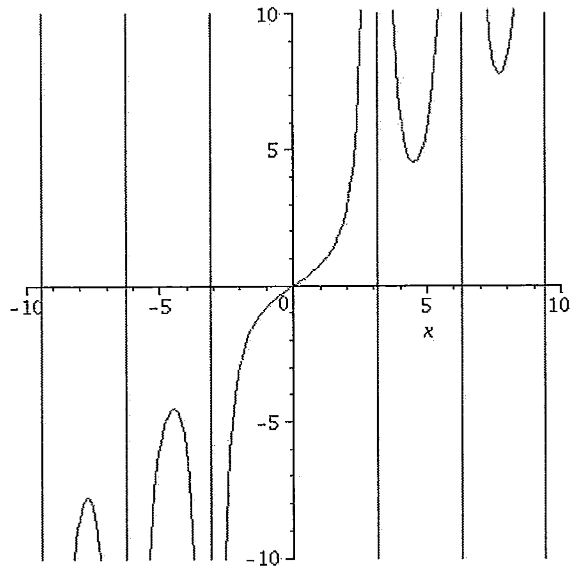


Figure 2.3: Gaveau's function $\mu(x)$

As written μ is defined for all $x \in \mathbb{R}$ except for integer multiples of π where $\sin x = 0$. Now μ has a removable singularity at $x = 0$ so that we can define

$$\mu(0) = \lim_{x \to 0} \mu(x) = 0.$$

Note that the Gaveau function is odd: $\mu(-x) = -\mu(x)$. We record the following properties of the Gaveau function. These results can all be proven using elementary calculus.

Proposition 2.4.4. *μ is a strictly increasing diffeomorphism from $(-\pi, \pi)$ to $\mathbb{R}$.*

Proposition 2.4.5. *For each integer $m = 1, 2, \ldots,$ there is exactly one critical point x_m for μ on the interval $\big(m\pi, (m+1)\pi\big)$. μ is strictly decreasing on $(m\pi, x_m)$ and strictly increasing on $\big(x_m, (m+1)\pi\big)$. We also have the limits*

$$\lim_{x \to m\pi^+} \mu(x) = +\infty = \lim_{x \to (m+1)\pi^-} \mu(x).$$

We can be more specific about where x_m lies in the interval $(m\pi, (m+1)\pi)$.

Proposition 2.4.6. *The critical point $x_m \in \big(m\pi, (m+1)\pi\big)$ is less than the midpoint*

$$x_m < \left(m + \frac{1}{2}\right)\pi$$

but within $\frac{1}{m\pi}$ of this midpoint

$$\left(m + \frac{1}{2}\right)\pi - x_m < \frac{1}{m\pi}.$$

And we can likewise say more about the local minimum values $\mu(x_m)$.

Proposition 2.4.7. *The local minimum values $\mu(x_m)$ satisfy the inequalities*

$$\mu(x_{m+1}) > \mu(x_m) + \pi.$$

In particular we have the crucial result that for a fixed constant $\lambda > 0$ there are only finitely many solutions to the equation

$$\mu(x) = \lambda.$$

There is always exactly one solution coming from a unique x in the interval $(-\pi, \pi)$ which solution is in fact in the interval $(0, \pi)$ because $\lambda > 0$. There are no solutions for $x < 0$ since $\mu(x) < 0$ when $x < 0$. For $x > 0$ eventually the local minimum values $\mu(x_m)$ on the intervals $\big(m\pi, (m+1)\pi\big)$ will be larger than λ and therefore solutions exist in only finitely many of the intervals. The graph of $\mu(x)$ makes these remarks apparent.

Likewise we can count the strictly finite number of solutions to $\mu(x) = \lambda \leq 0$.

2.4.4 Geodesics from O to $P(x, y, t)$ with $x^2 + y^2 \neq 0$ and $t \neq 0$

Lemma 2.4.8. *An arc-length parametrized geodesic starting at O with $\theta \neq 0$ has equation*

$$r^2(\phi) = \frac{1}{4\theta^2}\sin^2(\phi - \phi_0)$$

$$t(\phi) = \frac{1}{4\theta^2}\frac{\sin 2(\phi - \phi_0) - 2(\phi - \phi_0)}{2}.$$

Proof. For an arc-length parametrized geodesic we have

$$1 = \dot{x}_1^2(s) + \dot{x}_2^2(s)$$

$$= \dot{r}^2 + 4\theta^2 r^2.$$

We thus compute

$$\left(\frac{dr}{ds}\right)^2 = 1 - 4\theta^2 r^2$$

$$\frac{dr}{\sqrt{1 - 4\theta^2 r^2}} = \pm ds$$

$$\frac{1}{2\theta}\arcsin(2\theta r) = \pm s$$

$$2\theta r = \sin(\pm 2\theta s)$$

therefore

$$r = \pm\frac{1}{2\theta}\sin(2\theta s).$$

Next from the horizontality condition

$$\dot{t} = 2(\dot{x}_1 x_2 - x_1 \dot{x}_2) = -2r^2\dot{\phi}$$

we can compute t in terms of ϕ

$$
\begin{aligned}
t(\phi) &= \int_{\phi_0}^{\phi} -2r^2 d\phi \\
&= -2\int_{\phi_0}^{\phi} \left(\frac{1}{2\theta}\right)^2 \sin^2(\phi-\phi_0)d\phi \\
&= -2\left(\frac{1}{4\theta^2}\right)\int_0^{\phi-\phi_0} \sin^2 u\, du \\
&= -2\left(\frac{1}{4\theta^2}\right)\int_0^{\phi-\phi_0} \frac{1}{2}-\frac{1}{2}\cos 2u\, du \\
&= -2\left(\frac{1}{4\theta^2}\right)\left[\frac{1}{2}u-\frac{1}{4}\sin 2u\right]_0^{\phi-\phi_0} \\
&= \frac{1}{4\theta^2}\frac{\sin 2(\phi-\phi_0)-2(\phi-\phi_0)}{2}.
\end{aligned}
$$

□

We further restrict our geodesics to start at O and terminate at some point $P(x, y, t)$ with $x^2 + y^2 = R^2 \neq 0$ and $t \neq 0$. We do not specify (x, y) but merely the distance R from the origin. If we insist that the geodesic leaves the origin at an angle of $\phi_0 = 0$, then once we fix $R \neq 0$ and $t \neq 0$, we will see that the terminal point of the geodesic is determined.

A way to visualize this is to draw a circle of radius R with fixed height $t \neq 0$ and centered around the t-axis. We want to determine all possible geodesics leaving the origin O at an angle of $\phi_0 = 0$, i.e., leaving along the positive x-axis, and ending up somewhere on this elevated ring.

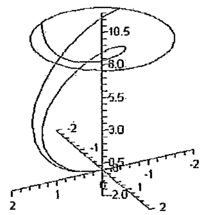


Figure 2.4: geodesics to an elevated ring

Lemma 2.4.9. *For an arc-length parametrized geodesic* $c : [0,\tau] \longrightarrow \mathbb{H}$ *with* $\theta \neq 0$ *and boundary conditions*

$$x(0) = 0 \qquad t(0) = 0 \qquad \phi(0) = 0$$

$$x^2(\tau) + y^2(\tau) = R^2 \neq 0 \qquad t(\tau) = t \neq 0 \qquad \phi(\tau) = \phi_1$$

the following relations hold

$$\phi_1 = -2\theta\tau$$

$$\sin^2 \phi_1 = 4\theta^2 R^2$$

$$t = \frac{1}{4\theta^2}\frac{\sin 2\phi_1 - 2\phi_1}{2}$$

$$\frac{t}{R^2} = -\mu(\phi_1) = \mu(2\theta\tau).$$

Proof. From $\dot{\phi} = -2\theta$ we get

$$\phi_1 = \phi(\tau) = \phi(\tau) - \phi(0) = -2\theta\tau.$$

From the previous lemma we get

$$R^2 = \frac{1}{4\theta^2}\sin^2\phi_1$$

and

$$t = \frac{1}{4\theta^2}\frac{\sin 2\phi_1 - 2\phi_1}{2}.$$

Using these we have

$$\begin{aligned} \frac{t}{R^2} &= \frac{\sin 2\phi_1 - 2\phi_1}{2\sin^2\phi_1} \\ &= -\left[\frac{\phi_1}{\sin^2\phi_1} - \cot\phi_1\right] \\ &= -\mu(\phi_1) \\ &= \mu(2\theta\tau). \end{aligned}$$

□

Lemma 2.4.10. *Consider arc-length parametrized geodesics* $c : [0,\tau] \longrightarrow \mathbb{H}$ *with* $\theta \neq 0$ *and boundary conditions*

$$x(0) = 0 \qquad t(0) = 0 \qquad \phi(0) = 0$$

$$x^2(\tau) + y^2(\tau) = R^2 \neq 0 \qquad t(\tau) = t \neq 0 \qquad \phi(\tau) = \phi_1.$$

There is precisely one such geodesic for each solution

$$\zeta_1 < \zeta_2 < \ldots < \zeta_N$$

to the equation

$$\mu(\zeta) = \frac{t}{R^2}$$

(note that the solutions ζ_m are all positive when $t > 0$ and all negative when $t < 0$). If $\tau = s_m$ is the arc-length of the geodesic associated to the solution ζ_m, then

$$s_m^2 = \left(\frac{\zeta_m}{\sin\zeta_m}\right)^2 R^2.$$

Proof. For such a geodesic we know that

$$\mu(2\theta\tau) = \frac{t}{R^2}.$$

The equation $\mu(\zeta) = t/R^2$ has finitely many solutions $\zeta_1, \ldots, \zeta_N$. Any geodesic must correspond to some such solution. Furthermore each such solution corresponds to a geodesic, as we can check by the explicit equation of such a geodesic, which we now have enough information to compute.

Let the value $\zeta_m = 2\theta\tau$ be the given solution to $\mu(\zeta) = t/R^2$. Now $\phi_1 = -2\theta\tau$ shows that the terminal angle ϕ_1 is determined. In particular the terminal point $(R\cos\phi_1, R\sin\phi_1, t)$ of the geodesic is also determined. The value θ is determined from the equation

$$R^2 = \frac{1}{4\theta^2}\sin^2\zeta_m.$$

This equation tells us θ^2. In fact, since $\phi_1 = -2\theta\tau$ and the arc-length τ is non-negative, the sign of θ is determined, so that the value of θ is completely determined. Note that $\theta > 0$ when $t > 0$ and $\theta < 0$ when $t < 0$.

Thus all the parameters θ, ϕ_1, τ are determined by the fact that the terminal point must be on the elevated ring $x^2 + y^2 = R^2 \neq 0$ with constant height $t \neq 0$.

Now for the arc-length parameter $s \in [0, \tau]$, we have $\dot{\phi} = -2\theta$ which implies $\phi = -2\theta s$. We plug $\phi = -2\theta s$ into our previous relations to get

$$r(s) = \frac{1}{2\theta}\sin(2\theta s)$$

$$\phi(s) = -2\theta s$$

$$t(s) = \frac{1}{4\theta^2}\frac{4\theta s - \sin 4\theta s}{2}.$$

Notice that we have chosen a positive sign in the equation for $r(s)$. This ensures that the geodesic leaves the origin at an angle of $\phi_0 = 0$ with the positive x-axis. (A negative sign would produce an angle of $\phi_0 = 0$ with the *negative* x-axis.)

Since we now have the equation of the geodesic c, it is straightforward to check directly that it is indeed horizontal and satisfies the criterion $\ddot{c} = 4\theta J\dot{c}$. Thus c is a geodesic.

Finally we can solve for $\tau = s_m$ using the two relations

$$\phi_1 = -2\theta\tau \Rightarrow \phi_1^2 = 4\theta^2\tau^2$$

$$\sin^2\phi_1 = 4\theta^2 R^2$$

and dividing to obtain

$$\tau^2 = \left(\frac{\phi_1}{\sin\phi_1}\right)^2 R^2 = \left(\frac{\zeta_m}{\sin\zeta_m}\right)^2 R^2.$$

□

The geodesics we have analyzed leave the origin O at an angle of $\phi_0 = 0$ and rise to the elevated ring $x^2 + y^2 = R^2 \neq 0$, $t \neq 0$. From the previous theorem, the geodesic corresponding to the solution $\zeta \in (-\pi, \pi)$ to the equation $\mu(\zeta) = t/R^2$ gives the geodesic of strictly minimum length.

Given a specific point now $P(x, y, t)$ where not only $\|(x, y)\| = R \neq 0$ is specified but also (x, y), we can find a geodesic terminating at this point by rotating one of the above $\phi_0 = 0$ geodesics. If we rotate the geodesic of strictly minimum length, we obtain a geodesic from O to $P(x, y, t)$ of strictly minimum length. This is because any geodesic from O to P can be rotated to one of the above $\phi_0 = 0$ geodesics and vice versa, preserving length.

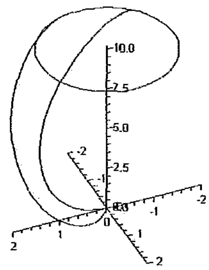


Figure 2.5: rotating a geodesic

Below we record the equation of the unique arc-length parametrized length-minimizing geodesic from O to such an elevated ring.

Theorem 2.4.11. *For fixed $R > 0$ and fixed $t \neq 0$, there is precisely one arc-length parametrized and length-minimizing geodesic leaving the origin O at an angle of ϕ_0 with the positive x-axis and terminating on the ring*

$$\left\{(x, y, t) | x^2 + y^2 = R^2\right\}.$$

This geodesic corresponds to the unique $\zeta \in (-\pi, \pi)$ which solves $\mu(\zeta) = t/R^2$.

The equation of this geodesic is

$$x(s) = \frac{1}{2\theta} \sin(2\theta s) \cos(\phi_0 - 2\theta s)$$

$$y(s) = \frac{1}{2\theta} \sin(2\theta s) \sin(\phi_0 - 2\theta s)$$

$$t(s) = \frac{1}{2\theta} s - \frac{1}{4\theta^2} \sin(2\theta s) \cos(2\theta s)$$

for $0 \leq s \leq \tau$*, where*

$$\tau = \frac{\zeta}{\sin\zeta} R$$

is the length of the geodesic, and

$$\theta = \frac{\zeta}{2\tau}.$$

Proof. This follows from the work done so far, just recalling that $\phi(s) - \phi(0) = -2\theta s$, that is, $\phi(s) = \phi_0 - 2\theta s$.

□

By an appropriate choice of the initial angle ϕ_0, we can make such a geodesic terminate at any specified point (x, y, t) where $x^2 + y^2 \neq 0$ and $t \neq 0$.

Theorem 2.4.12. *The equation of the unique arc-length parametrized length-minimizing geodesic from* O *to the fixed point* (x, y, t) *where* $x^2 + y^2 \neq 0$ *and* $t \neq 0$ *is*

$$x(s) = (x\cot\zeta - y)\sin(2\theta s)\cos(2\theta s) + (x + y\cot\zeta)\sin^2(2\theta s)$$

$$y(s) = (x + y\cot\zeta)\sin(2\theta s)\cos(2\theta s) - (x\cot\zeta - y)\sin^2(2\theta s)$$

$$t(s) = \frac{1}{2\theta}s - \frac{1}{4\theta^2}\sin(2\theta s)\cos(2\theta s)$$

for $0 \leq s \leq \tau$ *where*

$$\zeta = \mu^{-1}\left(\frac{t}{x^2 + y^2}\right) \qquad \tau = \frac{\zeta}{\sin\zeta}\sqrt{x^2 + y^2} \qquad \theta = \frac{\sin\zeta}{2\sqrt{x^2 + y^2}}.$$

Note that we write $\mu^{-1}(\alpha)$ *for* $\alpha \in \mathbb{R}$ *to denote the unique solution on the interval* $(-\pi, \pi)$.

Proof. We start with the equation for the unique arc-length parametrized and length minimizing geodesic leaving the origin at an angle of $\phi_0 = 0$ and terminating on the ring of height t and radius $R^2 = x^2 + y^2$

$$x(s) = \frac{1}{2\theta}\sin(2\theta s)\cos(2\theta s)$$

$$y(s) = -\frac{1}{2\theta}\sin(2\theta s)\sin(2\theta s)$$

$$t(s) = \frac{1}{2\theta}s - \frac{1}{4\theta^2}\sin(2\theta s)\cos(2\theta s)$$

with $0 \leq s \leq \tau$ where

$$\zeta = \mu^{-1}\left(\frac{t}{x^2+y^2}\right) \qquad \tau = \frac{\zeta}{\sin\zeta}\sqrt{x^2+y^2} \qquad \theta = \frac{\sin\zeta}{2\sqrt{x^2+y^2}}.$$

The terminal point of this geodesic can be directly computed

$$\left(\sqrt{x^2+y^2}\cos\zeta, -\sqrt{x^2+y^2}\sin\zeta, t\right).$$

An appropriate rotation will then rotate this terminal point into the terminal point (x, y, t). The rotation matrix turns out to be

$$\frac{1}{\sqrt{x^2+y^2}}\begin{bmatrix} x\cos\zeta - y\sin\zeta & -(x\sin\zeta + y\cos\zeta) \\ x\sin\zeta + y\cos\zeta & x\cos\zeta - y\sin\zeta \end{bmatrix}.$$

Rotating the original geodesic by this matrix results in the stated equation.

□

2.4.5 Summary of connectivity

Between the origin $O(0,0,0)$ and a point $P(0,0,t)$ with $t \neq 0$ there is a one-parameter family of length-minimizing geodesics from O to P of length $\sqrt{\pi|t|}$. In particular there do not exist unique length-minimizing geodesics between O and such points P.

Between the origin $O(0,0,0)$ and a point $P(x,y,0)$ in the xy-plane there is a unique arc-length parametrized geodesic from O to P of length

$$\sqrt{x^2+y^2}.$$

Between the origin $O(0,0,0)$ and a point $P(x,y,t)$ with $x^2+y^2 \neq 0$ and $t \neq 0$ there is a unique arc-length parametrized length-minimizing geodesic from O to P with ζ-value

$$\zeta = \mu^{-1}\left(\frac{t}{x^2+y^2}\right) \tag{2.4.3}$$

and with length

$$\frac{\zeta}{\sin\zeta}\sqrt{x^2+y^2} = \sigma(\zeta)\sqrt{x^2+y^2}. \tag{2.4.4}$$

We introduce the function

$$\sigma : (-\pi, \pi) \to \mathbb{R}$$

$$x \mapsto \frac{x}{\sin x}$$

to help unify the properties of these two types of geodesics.

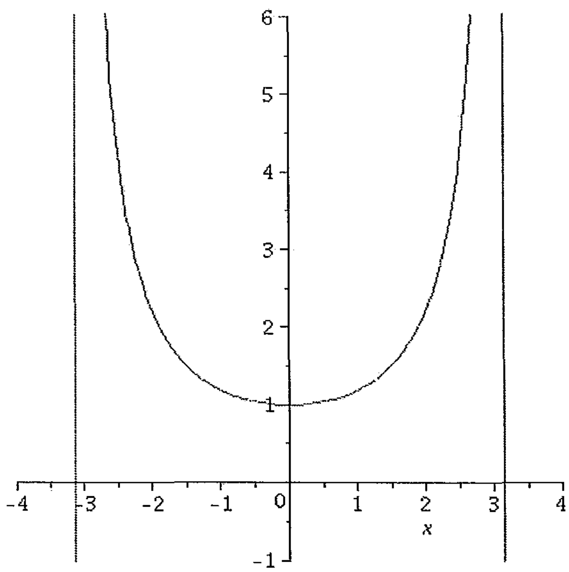


Figure 2.6: $\sigma(x)$

Now σ has a removable singularity at $x = 0$ so that we can define $\sigma(0)$ by the limit as $x \to 0$

$$\sigma(0) = \lim_{x\to 0} \frac{x}{\sin x} = 1.$$

Note that σ is an even function. The behaviour of σ is simple to describe. σ decreases from $+\infty$ to 1 on the interval $(-\pi, 0]$ and increases from 1 to $+\infty$ on the interval $[0, \pi)$.

Now observe that if we assign a ζ-value of 0 to the first group of geodesics from O to $P(x, y, 0)$ then formula 2.4.3 for ζ-value and formula 2.4.4 for length apply to all geodesics from O to points P not on the t-axis, that is, to all geodesics from O to points P between which there is a *unique* length-minimizing geodesic.

This classification of geodesics from O to P actually classifies geodesics between arbitrary points A and B because any geodesic between A and B left-translates to a geodesic between O and $A^{-1}B$ and vice versa.

In what follows we work primarily with points A and B between which a unique length-minimizing geodesic exists. The condition for this to happen between O and P is merely that $(x, y) \neq (0, 0)$. Thus the criteria for this to happen between A and B is merely that $(x_A, y_A) \neq (x_B, y_B)$.

For such points A and B there exists a unique arc-length parametrized length-minimizing geodesic from A to B with ζ-value

$$\zeta = \mu^{-1} \left(\frac{t_B - t_A - 2(x_B y_A - y_B x_A)}{(x_B - x_A)^2 + (y_B - y_A)^2} \right) \tag{2.4.5}$$

and length

$$\sigma(\zeta) \sqrt{(x_B - x_A)^2 + (y_B - y_A)^2}. \tag{2.4.6}$$

2.5 Isometries

Definition 5. *The distance*

$$d(A, B)$$

between two points A and B in $\mathbb{H}$ is defined to be the length of a length-minimizing geodesic between A and B.

It is a non-trivial theorem that this distance function defines a metric on $\mathbb{H}$. This metric is called the Carnot-Caratheodory metric and sometimes we write d_{CC} for this metric. We have two explicit formulas for $d(A, B)$

$$d\left(A(x_A, y_A, t_A), B(x_B, y_B, t_B)\right) = \sqrt{\pi |t_B - t_A|}$$

which applies when $(x_A, y_A) = (x_B, y_B)$ and

$$d\left(A(x_A, y_A, t_A), B(x_B, y_B, t_B)\right) = \sigma(\zeta)\sqrt{(x_B - x_A)^2 + (y_B - y_A)^2} \tag{2.5.1}$$

which applies when $(x_A, y_A) \neq (x_B, y_B)$.

To describe geodesic connectivity we needed to know that left-translation $L_A c$ and rotation $R_\varphi c$ of a geodesic c preserve length. These facts immediately give us two sorts of isometries.

Proposition 2.5.1. *For any $A \in \mathbb{H}$ the map left-translation by A*

$$L_A : \mathbb{H} \to \mathbb{H}$$

is an isometry.

Proposition 2.5.2. *For any $0 \leq \varphi \leq 2\pi$ the map rotation by φ*

$$R_\varphi : \mathbb{H} \to \mathbb{H}$$

is an isometry.

There is a transformation that provides another sort of isometry. We define the *involution* r by

$$r : \mathbb{H} \to \mathbb{H}$$

$$(x, y, t) \mapsto (x, -y, -t).$$

Proposition 2.5.3. *For $c(s) = (x(s), y(s), t(s))$ a horizontal curve the curve $\bar{c}(s) = r \circ c(s)$ is also horizontal.*

Proof. We have

$$\dot{\bar{t}} = -\dot{t} = -2(\dot{x}y - x\dot{y})$$

$$= 2\left[\dot{x}(-y) - x(-\dot{y})\right]$$

$$= 2\left(\dot{\bar{x}}\bar{y} - \bar{x}\dot{\bar{y}}\right).$$

□

Proposition 2.5.4. *For horizontal $c : [a, b] \to \mathbb{H}$ the curves $c(s)$ and $\bar{c}(s) = r \circ c(s)$ have the same length.*

Proof. We have $\dot{\bar{x}}^2 + \dot{\bar{y}}^2 = \dot{x}^2 + \dot{y}^2$ and the result follows.

□

Thus far we have encountered isometries in the form of the left-translations L_A, the rotations R_φ, and the involution r. It is a non-trivial result that these three types of isometries generate the full group of isometries of $\mathbb{H}$. We record this result below, as well as results about the structure of this group of isometries. For details see [8].

Theorem 2.5.5. *The group Isom($\mathbb{H}$) of isometries of $\mathbb{H}$ is generated by the group $\mathbb{H}$ of left-translations $L_A : \mathbb{H} \longrightarrow \mathbb{H}$ for $A \in \mathbb{H}$; the group $SO(2)$ of rotations $R_\varphi : \mathbb{H} \longrightarrow \mathbb{H}$ for $0 \leq \varphi \leq 2\pi$; and the involution $r : \mathbb{H} \longrightarrow \mathbb{H}$.*

The subgroup of Isom($\mathbb{H}$) generated by the left-translations $\mathbb{H}$ and the rotations $SO(2)$ is a normal subgroup of Isom($\mathbb{H}$) of index 2. This subgroup has the semi-direct product structure

$$\mathbb{H} \rtimes SO(2).$$

The full group of isometries can then be written as the disjoint union of the cosets

$$Isom\ (\mathbb{H}) = [\mathbb{H} \rtimes SO(2)] \cup r\,[\mathbb{H} \rtimes SO(2)]\,.$$

Note also that together the rotations $SO(2)$ and the involution r generate a subgroup of Isom($\mathbb{H}$) that is isomorphic to $O(2)$. Isom($\mathbb{H}$) can then be written as the semi-direct product $\mathbb{H} \rtimes O(2)$.

Chapter 3

Triangles in the Heisenberg group

Now that we have a complete description of geodesics between arbitrary distinct points in the Heisenberg group, we can construct triangles by connecting triples of vertices by geodesics. Our first result is an angle deficit formula relating the angles of a Heisenberg triangle to the ζ-values of the geodesic sides of the triangle.

3.1 An angle deficit formula

3.1.1 Tangent vectors to geodesics

Between the origin O and any point $P(x, y, t)$ such that $(x, y) \neq (0, 0)$ and $t \neq 0$, there exists a unique arc-length parametrized geodesic $\gamma(s) = (x(s), y(s), t(s))$ from O to P. We record in full the equation for this geodesic:

$$x(s) = (x \cot \zeta - y) \sin(2\theta s) \cos(2\theta s) + (x + y \cot \zeta) \sin^2(2\theta s)$$

$$y(s) = (x + y \cot \zeta) \sin(2\theta s) \cos(2\theta s) - (x \cot \zeta - y) \sin^2(2\theta s)$$

$$t(s) = \frac{1}{2\theta} s - \frac{1}{4\theta^2} \sin(2\theta s) \cos(2\theta s)$$

for $0 \leq s \leq \tau$, where

$$\zeta = \mu^{-1}\left(\frac{t}{x^2+y^2}\right)$$

$$\tau = \sigma(\zeta)\sqrt{x^2+y^2}$$

$$2\theta = \frac{\sin\zeta}{\sqrt{x^2+y^2}}.$$

From this expression we determine the tangent vectors to γ, which will in turn allow us to determine the tangent vectors to arc-length parametrized geodesics between arbitrary points. Since γ is horizontal, the fact that $\dot{\gamma} = \dot{x}X + \dot{y}Y$ simplifies the calculation, since we do not have to work with $t(s)$. We calculate separately that

$$\frac{d}{ds}\left[\sin(2\theta s)\cos(2\theta s)\right] = 2\theta\cos(4\theta s)$$

$$\frac{d}{ds}\left[\sin^2(2\theta s)\right] = 2\theta\sin(4\theta s).$$

Then

$$\dot{x}(s) = 2\theta(x\cot\zeta - y)\cos(4\theta s) + 2\theta(x + y\cot\zeta)\sin(4\theta s)$$

$$\dot{y}(s) = 2\theta(x + y\cot\zeta)\cos(4\theta s) - 2\theta(x\cot\zeta - y)\sin(4\theta s).$$

We are particularly interested in tangent vectors at O ($s = 0$) and at P ($s = \tau$), because we use these vectors to calculate angles between geodesics. The tangent vector to γ at O is

$$\begin{aligned}\dot{\gamma}(0) &= \dot{x}(0)X + \dot{y}(0)Y \\ &= 2\theta(x\cot\zeta - y)X + 2\theta(x + y\cot\zeta)Y.\end{aligned}$$

The tangent vector to γ at P is

$$\begin{aligned}\dot{\gamma}(\tau) &= \dot{x}(\tau)X + \dot{y}(\tau)Y \\ &= 2\theta\left[(x\cot\zeta - y)\cos 2\zeta + (x + y\cot\zeta)\sin 2\zeta\right]X \\ &\quad + 2\theta\left[(x + y\cot\zeta)\cos 2\zeta - (x\cot\zeta - y)\sin 2\zeta\right]Y\end{aligned}$$

where we have used the fact that $2\theta\tau = \zeta$. A calculation shows that the coefficients of X and Y above simplify:

$$(x\cot\zeta - y)\cos 2\zeta + (x + y\cot\zeta)\sin 2\zeta = x\cot\zeta + y$$

$$(x + y\cot\zeta)\cos 2\zeta - (x\cot\zeta - y)\sin 2\zeta = -x + y\cot\zeta.$$

Thus

$$\dot\gamma(\tau) = 2\theta(x\cot\zeta + y)X + 2\theta(-x + y\cot\zeta)Y.$$

Finally, using $2\theta = \frac{\sin\zeta}{\sqrt{x^2+y^2}}$, we obtain expressions in the most convenient form:

$$\dot\gamma(0) = \frac{1}{\|(x,y)\|}\left[(x\cos\zeta - y\sin\zeta)X + (x\sin\zeta + y\cos\zeta)Y\right]$$

$$\dot\gamma(\tau) = \frac{1}{\|(x,y)\|}\left[(x\cos\zeta + y\sin\zeta)X + (-x\sin\zeta + y\cos\zeta)Y\right].$$

This calculation was carried out for $P(x, y, t)$ with $t \neq 0$ and therefore $\zeta \neq 0$. For the case $t = 0$, we have $\zeta = 0$ whereupon $\cot\zeta$ is undefined and the above calculation does not apply. We check, however, that the final expressions obtained for $\dot\gamma(0)$ and $\dot\gamma(\tau)$ remain valid. Indeed, the unique arc-length parametrized geodesic from O to $P(x, y, 0)$ with $x^2 + y^2 > 0$ has the equation

$$\gamma(s) = \frac{1}{\sqrt{x^2+y^2}}(xs, ys, 0) \quad 0 \le s \le \sqrt{x^2+y^2}$$

with tangent vectors

$$\dot\gamma(s) = \frac{x}{\sqrt{x^2+y^2}}X + \frac{y}{\sqrt{x^2+y^2}}Y$$

independent of s. For such a geodesic $\zeta = 0$ and the formula for $\dot\gamma$ holds.

We have covered the case for geodesics with initial point O. In general, between distinct points $A(x, y, t)$ and $B(u, v, s)$ such that $(x, y) \neq (u, v)$, there exists a unique arc-length parametrized geodesic $\bar\gamma(s) = (\bar\gamma_1(s), \bar\gamma_2(s), \bar\gamma_3(s))$ from A to B. As we have seen, $\bar\gamma$ is the left-translation by A of the unique arc-length parametrized geodesic γ from O to $A^{-1}B$. Thus

$$\dot{\bar\gamma}_1(s) = \frac{d}{ds}(-x + \gamma_1(s)) = \dot\gamma_1(s)$$

$$\dot{\bar\gamma}_2(s) = \frac{d}{ds}(-y + \gamma_2(s)) = \dot\gamma_2(s)$$

so that the tangent vectors to $\bar{\gamma}$ at A and B have the same components as the tangent vectors to γ at O and $A^{-1}B$ respectively. Therefore for such a geodesic we have

$$\dot{\gamma}(0) = \frac{1}{\|(\Delta x, \Delta y)\|}\left[(\Delta x \cos\zeta - \Delta y \sin\zeta)X + (\Delta x \sin\zeta + \Delta y \cos\zeta)Y\right]$$

$$\dot{\gamma}(\tau) = \frac{1}{\|(\Delta x, \Delta y)\|}\left[(\Delta y \sin\zeta + \Delta x \cos\zeta)X + (-\Delta x \sin\zeta + \Delta y \cos\zeta)Y\right]$$

where $\Delta x = u - x$, $\Delta y = v - y$, and

$$\zeta = \mu^{-1}\left(\frac{s - t - 2(uy - vx)}{\Delta x^2 + \Delta y^2}\right).$$

3.1.2 Angles between geodesics

Consider a triangle in $\mathbb{H}$ determined by an arbitrary set of three distinct points

$$A(x_A, y_A, t_A)$$

$$B(x_B, y_B, t_B)$$

$$C(x_C, y_C, t_C)$$

subject to the condition that (x_A, y_A), (x_B, y_B), (x_C, y_C) are distinct as well.

By analogy with the Euclidean case, we define the cosine of the angle between horizontal vectors $\mathbf{a}$ and $\mathbf{b}$ based at the same point by the formula

$$\|\mathbf{a}\|\,\|\mathbf{b}\|\cos\phi = g(\mathbf{a}, \mathbf{b})$$

where g is the sub-Riemannian metric and the norm $\|\cdot\|$ is calculated using g.

To calculate the angle at A we use the two tangent vectors at A, the first on γ_C moving towards B, the second on γ_B moving towards C.

The tangent vector at A towards B:

$$\frac{1}{\|(x_B - x_A, y_B - y_A)\|}\begin{bmatrix} ((x_B - x_A)\cos\zeta_C - (y_B - y_A)\sin\zeta_C)\,X \\ + ((x_B - x_A)\sin\zeta_C + (y_B - y_A)\cos\zeta_C)\,Y \end{bmatrix}.$$

The tangent vector at A towards C:

$$\frac{-1}{\|(x_A - x_C, y_A - y_C)\|}\begin{bmatrix} ((x_A - x_C)\cos\zeta_B + (y_A - y_C)\sin\zeta_B)\,X \\ + (-(x_A - x_C)\sin\zeta_B + (y_A - y_C)\cos\zeta_B)\,Y \end{bmatrix}.$$

Since the geodesics γ_B and γ_C are parametrized by arc-length, the tangent vectors are of unit length. Thus $\cos\angle A$ is given by the sub-Riemmannian inner product of the tangent vectors at A. After simplifying and using the angle sum formulas for sin and cos, we get

$$\|(x_B - x_A, y_B - y_A)\|\,\|(x_C - x_A, y_C - y_A)\|\cos\angle A$$

$$= [(x_B - x_A)(x_C - x_A) + (y_B - y_A)(y_C - y_A)]\cos(\zeta_B + \zeta_C)$$

$$+ [(x_B - x_A)(y_C - y_A) - (y_B - y_A)(x_C - x_A)]\sin(\zeta_B + \zeta_C).$$

If we write

$$A(z_A, t_A) \qquad z_A = x_A + iy_A$$

$$B(z_B, t_B) \qquad z_B = x_B + iy_B$$

$$C(z_C, t_C) \qquad z_C = x_C + iy_C$$

then

$$(z_B - z_A)\overline{(z_C - z_A)} = [(x_B - x_A)(x_C - x_A) + (y_B - y_A)(y_C - y_A)]$$

$$- i\,[(x_B - x_A)(y_C - y_A) - (y_B - y_A)(x_C - x_A)].$$

Thus we can write

$$\cos\angle A = \operatorname{Re}\left(\frac{(z_B - z_A)\overline{(z_C - z_A)}}{|z_B - z_A|\,|z_C - z_A|}e^{i(\zeta_B + \zeta_C)}\right).$$

In the same way, we obtain analogous expressions for $\cos\angle B$ and $\cos\angle C$. We record all three:

$$\cos \angle A = \mathrm{Re}\left(\frac{(z_B - z_A)\overline{(z_C - z_A)}}{|z_B - z_A|\,|z_C - z_A|}e^{i(\zeta_B+\zeta_C)}\right)$$

$$\cos \angle B = \mathrm{Re}\left(\frac{(z_C - z_B)\overline{(z_A - z_B)}}{|z_C - z_B|\,|z_A - z_B|}e^{i(\zeta_C+\zeta_A)}\right)$$

$$\cos \angle C = \mathrm{Re}\left(\frac{(z_A - z_C)\overline{(z_B - z_C)}}{|z_A - z_C|\,|z_B - z_C|}e^{i(\zeta_A+\zeta_B)}\right).$$

Notice that the complex number

$$\frac{(z_B - z_A)\overline{(z_C - z_A)}}{|z_B - z_A|\,|z_C - z_A|}e^{i(\zeta_B+\zeta_C)}$$

has unit modulus, and real part equal to $\cos \angle A$. This motivates the following definition.

Definition 6. *In a triangle determined by vertices $(A, B, C) \in \mathbb{H}^3 - \Sigma$ the angle at A, denoted $\angle A$, is the unique angle in $[0, 2\pi)$ such that*

$$e^{i\angle A} = \frac{(z_B - z_A)\overline{(z_C - z_A)}}{|z_B - z_A|\,|z_C - z_A|}e^{i(\zeta_B+\zeta_C)}.$$

We similarly define the angles at B and C:

$$e^{i\angle B} = \frac{(z_C - z_B)\overline{(z_A - z_B)}}{|z_C - z_B|\,|z_A - z_B|}e^{i(\zeta_C+\zeta_A)}$$

$$e^{i\angle C} = \frac{(z_A - z_C)\overline{(z_B - z_C)}}{|z_A - z_C|\,|z_B - z_C|}e^{i(\zeta_A+\zeta_B)}.$$

From these definitions we can read off that

$$e^{i(\angle A+\angle B+\angle C)} = (-1)e^{2i(\zeta_A+\zeta_B+\zeta_C)}$$

which gives us an angle deficit formula for Heisenberg triangles.

Theorem 3.1.1. *For a triangle ABC we have*

$$\angle A + \angle B + \angle C - \pi \equiv 2\left(\zeta_A + \zeta_B + \zeta_C\right) \;\; mod\ 2\pi.$$

Proof. In

$$e^{i(\angle A + \angle B + \angle C)} = (-1)e^{2i(\zeta_A + \zeta_B + \zeta_C)}$$

we write $-1 = e^{i\pi}$ and the result follows.

□

3.2 Curvature of a geodesic

The quantity $2\left(\zeta_A + \zeta_B + \zeta_C\right)$ has a geometric interpretation which we proceed to demonstrate. For a horizontal curve $c(s)$ the curvature can be given by the expression

$$\kappa(s) = |\ddot{c}(s)| \,/\, |\dot{c}(s)|$$

where length $|\cdot|$ is computed using the sub-Riemannian metric g. When c is an arc-length parametrized geodesic, $\kappa(s) = |\ddot{c}(s)|$.

We work instead with a signed curvature, which for geodesics is easy to characterize: we define $\kappa(s)$ as above (at this stage a non-negative quantity) and give it a sign, + if the ζ-value for the geodesic is positive, and − if the ζ-value for the geodesic is negative.

For geodesics $\ddot{c} = \ddot{x}X + \ddot{y}Y$. Using the equation for a geodesic emanating from the origin, we compute

$$\begin{aligned}
\ddot{x}(s) &= \frac{d}{ds}\dot{x}(s) \\
&= \frac{d}{ds}\left[2\theta(x\cot\zeta - y)\cos(4\theta s) + 2\theta(x + y\cot\zeta)\sin(4\theta s)\right] \\
&= 8\theta^2\left[-(x\cot\zeta - y)\sin(4\theta s) + (x + y\cot\zeta)\cos(4\theta s)\right]
\end{aligned}$$

and

$$\ddot{y}(s) = 8\theta^2 \left[-(x + y\cot\zeta)\sin(4\theta s) - (x\cot\zeta - y)\cos(4\theta s) \right].$$

Then

$$|\ddot{c}| = (g(\ddot{c}, \ddot{c}))^{\frac{1}{2}} = \left(\ddot{x}^2 + \ddot{y}^2\right)^{\frac{1}{2}}$$

and after computing we obtain

$$\kappa(s) = \pm 8\theta^2 \sqrt{\frac{x^2+y^2}{\sin^2\zeta}} = \frac{8\theta^2}{\sin\zeta}\sqrt{x^2+y^2}.$$

Using the identity $2\theta = \frac{\sin\zeta}{\sqrt{x^2+y^2}}$ this simplifies further to

$$\kappa(s) = 2\frac{\sin\zeta}{\sqrt{x^2+y^2}}.$$

Notice the curvature is independent of s, i.e., curvature is constant for a geodesic. Thus

$$\int_c \kappa(s)ds = 2\frac{\sin\zeta}{\sqrt{x^2+y^2}} \times \text{length } c$$

$$= \left(2\frac{\sin\zeta}{\sqrt{x^2+y^2}}\right)\left(\frac{\zeta}{\sin\zeta}\sqrt{x^2+y^2}\right) = 2\zeta.$$

Observe that left-translation of a geodesic preserves the ζ-value, and as we have already noted preserves the components of the tangent vectors and therefore of the acceleration vectors. The above result therefore holds for arbitrary geodesics. We integrate the curvature around a triangle to obtain the following result.

Theorem 3.2.1. *For a triangle ABC, the integral of the curvature around the triangle is given by*

$$\int_\triangle \kappa(s)ds = 2\left(\zeta_A + \zeta_B + \zeta_C\right).$$

The angle deficit formula can now be written as

Theorem 3.2.2. *For a triangle ABC we have*

$$\angle A + \angle B + \angle C - \pi \equiv \int_\triangle \kappa(s)ds \ \mathit{mod}\ 2\pi.$$

3.3 A law of sines

We look more closely at the coefficient

$$\frac{(z_B - z_A)\overline{(z_C - z_A)}}{|z_B - z_A|\,|z_C - z_A|}.$$

for $e^{i(\zeta_B+\zeta_C)}$ in the expression for $e^{i\angle A}$. This coefficient is of unit modulus, so we can define α to be the unique angle in $[-\pi, \pi)$ such that

$$e^{i\alpha} = \frac{(z_B - z_A)\overline{(z_C - z_A)}}{|z_B - z_A|\,|z_C - z_A|}.$$

Analogously we obtain β, γ such that

$$e^{i\beta} = \frac{(z_C - z_B)\overline{(z_A - z_B)}}{|z_C - z_B|\,|z_A - z_B|}$$

$$e^{i\gamma} = \frac{(z_A - z_C)\overline{(z_B - z_C)}}{|z_A - z_C|\,|z_B - z_C|}.$$

Looking at $e^{i\alpha}$, we have already computed that

$$\mathrm{Re}\left(\frac{(z_B - z_A)\overline{(z_C - z_A)}}{|z_B - z_A|\,|z_C - z_A|}\right) = \frac{(x_B - x_A)(x_C - x_A) + (y_B - y_A)(y_C - y_A)}{|z_B - z_A|\,|z_C - z_A|}$$

$$\mathrm{Im}\left(\frac{(z_B - z_A)\overline{(z_C - z_A)}}{|z_B - z_A|\,|z_C - z_A|}\right) = -\frac{(x_B - x_A)(y_C - y_A) - (y_B - y_A)(x_C - x_A)}{|z_B - z_A|\,|z_C - z_A|}.$$

The expression for the real part is precisely the cosine of the angle between the vectors $\overrightarrow{z_A z_B}$ and $\overrightarrow{z_A z_C}$. Recalling our criterion for orientation of triangles in the plane, we see that if $\triangle z_A z_B z_C$ has clockwise orientation, that is, if

$$(x_B - x_A)(y_C - y_A) - (y_B - y_A)(x_C - x_A) < 0,$$

then $\mathrm{Im}\,(e^{i\alpha})$ is positive. Therefore in this case α is between 0 and π, so that α is in fact the angle at z_A in the triangle $z_A z_B z_C$. So α is the angle at z_A in the triangle $z_A z_B z_C$ to which the Heisenberg triangle ABC projects, when $\triangle z_A z_B z_C$ has clockwise orientation. When $\triangle z_A z_B z_C$ has counterclockwise orientation, we find that $-\alpha$, $-\beta$, and $-\gamma$ are the angles in the triangle $z_A z_B z_C$.

In both cases, the Euclidean law of sines tells us that

$$\lambda = \frac{\sin\alpha}{\tilde{a}} = \frac{\sin\beta}{\tilde{b}} = \frac{\sin\gamma}{\tilde{c}}.$$

Now by the choice of α, β, γ we have

$$e^{i\angle A} = e^{i\alpha}e^{i(\zeta_B+\zeta_C)}$$

$$e^{i\angle B} = e^{i\beta}e^{i(\zeta_C+\zeta_A)}$$

$$e^{i\angle C} = e^{i\gamma}e^{i(\zeta_A+\zeta_B)}$$

so that

$$\sin\alpha = \sin\left(\angle A - (\zeta_B + \zeta_C)\right) = \sin\left((\angle A + \zeta_A) - \Delta\right)$$

$$\sin\beta = \sin\left(\angle B - (\zeta_C + \zeta_A)\right) = \sin\left((\angle B + \zeta_B) - \Delta\right)$$

$$\sin\gamma = \sin\left(\angle C - (\zeta_A + \zeta_B)\right) = \sin\left((\angle C + \zeta_C) - \Delta\right)$$

where $\Delta = \zeta_A + \zeta_B + \zeta_C$. Expanding using the angle sum formula for sine we have

$$\sin\left((\angle A + \zeta_A) - \Delta\right) = \sin(\angle A + \zeta_A)\cos\Delta - \cos(\angle A + \zeta_A)\sin\Delta$$

with similar expressions for $\sin\beta$ and $\sin\gamma$. Hence we have

$$\lambda = \frac{1}{\tilde{a}}\sin(\angle A + \zeta_A)\cos\Delta - \frac{1}{\tilde{a}}\cos(\angle A + \zeta_A)\sin\Delta$$

$$= \frac{1}{\tilde{b}}\sin(\angle B + \zeta_B)\cos\Delta - \frac{1}{\tilde{b}}\cos(\angle B + \zeta_B)\sin\Delta$$

$$= \frac{1}{\tilde{c}}\sin(\angle C + \zeta_C)\cos\Delta - \frac{1}{\tilde{c}}\cos(\angle C + \zeta_C)\sin\Delta.$$

In other words the points

$$\left(\frac{\sin(\angle A+\zeta_A)}{\tilde{a}}, \frac{\cos(\angle A+\zeta_A)}{\tilde{a}}\right)$$

$$\left(\frac{\sin(\angle B+\zeta_B)}{\tilde{b}}, \frac{\cos(\angle B+\zeta_B)}{\tilde{b}}\right)$$

$$\left(\frac{\sin(\angle C+\zeta_C)}{\tilde{c}}, \frac{\cos(\angle C+\zeta_C)}{\tilde{c}}\right)$$

lie on the line $x\cos\Delta - y\sin\Delta = \lambda$. The slope of this line is $\cot\Delta$, so $\cot\Delta$ equals the slope between any two of the three points. In particular we have the following.

Theorem 3.3.1. *For a triangle ABC the triangle parameters (a,ζ_A), (b,ζ_B), (c,ζ_C) and the Heisenberg angles $\angle A$, $\angle B$, $\angle C$ satisfy the identity*

$$\frac{\cos(\angle A+\zeta_A)/\tilde{a} - \cos(\angle B+\zeta_B)/\tilde{b}}{\sin(\angle A+\zeta_A)/\tilde{a} - \sin(\angle B+\zeta_B)/\tilde{b}}$$

$$= \frac{\cos(\angle B+\zeta_B)/\tilde{b} - \cos(\angle C+\zeta_C)/\tilde{c}}{\sin(\angle B+\zeta_B)/\tilde{b} - \sin(\angle C+\zeta_C)/\tilde{c}}$$

$$= \frac{\cos(\angle C+\zeta_C)/\tilde{c} - \cos(\angle A+\zeta_A)/\tilde{a}}{\sin(\angle C+\zeta_C)/\tilde{c} - \sin(\angle A+\zeta_A)/\tilde{a}}$$

with each ratio equal to the common value $\cot\Delta$, where $\Delta = \zeta_A+\zeta_B+\zeta_C$.

Chapter 4

The moduli space of triangles in the Heisenberg group

4.1 The space of vertices

A triangle in $\mathbb{H}$ is determined by an arbitrary set of three distinct vertices

$$A(x_A, y_A, t_A) = A(z_A, t_A)$$

$$B(x_B, y_B, t_B) = B(z_B, t_B)$$

$$C(x_C, y_C, t_C) = C(z_C, t_C)$$

subject to the condition that z_A, z_B, z_C are distinct as well, that is, that no vertex A, B, C lies vertically above any other. For such triples of vertices, there exist unique length-minimizing geodesics between each pair of vertices, so that the sides of a well-defined triangle are formed by the geodesics γ_A from B to C, γ_B from C to A, and γ_C from A to B.

We can describe all such "good" triples of vertices by the set

$$\widetilde{T} := \mathbb{H} \times \mathbb{H} \times \mathbb{H} - \Sigma$$

where Σ removes all the "bad" triples, that is, those triples for which at least one vertex lies vertically above another. Notice that Σ includes all triples (A, B, C) with coinciding points. Thus $\mathbb{H}^3 - \Sigma$ contains only triples of distinct vertices. Moreover, the set Σ decomposes as a union

$$\Sigma = \Sigma_1 \cup \Sigma_2 \cup \Sigma_3$$

where Σ_1 contains those bad triples for which A lies vertically above or below B, Σ_2 contains those for which B lies vertically above or below C, and Σ_3 those for which C lies vertically above or below A. The advantage of decomposing Σ in this way is that each set Σ_i can be described explicitly. Indeed, Σ_1 contains exactly those points $(A, B, C) \in \mathbb{H}^3$ for which

$$x_A = x_B$$

$$y_A = y_B.$$

In full detail, Σ_1 consists of 9-tuples

$$(x_A, y_A, t_A; x_B, y_B, t_B; x_C, y_C, t_C) \in \mathbb{R}^9$$

subject to two independent linear constraints

$$x_A - x_B = 0$$

$$y_A - y_B = 0.$$

Therefore Σ_1 is a 7-dimensional subspace of $\mathbb{R}^9$, in other words, a subspace of codimension 2. Σ_2 and Σ_3 are likewise each 7-dimensional subspaces of $\mathbb{R}^9$. The point is that it will be possible to analyze the topological properties of the space $\widetilde{T} = \mathbb{H}^3 - \Sigma$, because it is really just $\mathbb{R}^9$ with three distinct codimension 2 subspaces removed. For instance, this remark already shows that $\widetilde{T}$ is an open subset of $\mathbb{R}^9$ and therefore inherits a smooth manifold structure as an open submanifold of Euclidean space $\mathbb{R}^9$.

4.2 The space of parameters

4.2.1 Description by defining equation

But the space $\widetilde{T} = \mathbb{H}^3 - \Sigma$ is too large to serve as the moduli space of Heisenberg triangles. For instance, if $g : \mathbb{H} \to \mathbb{H}$ is an isometry, then the triples (A, B, C)

and (gA, gB, gC) determine congruent triangles. Up to isometry the triangles are identical. To account for this, we need to look more closely at the map

$$p : \mathbb{H}^3 - \Sigma \longrightarrow \mathbb{R}^6$$

which sends the triangle ABC to its geodesic parameters

$$(A, B, C) \mapsto (a, \zeta_A), (b, \zeta_B), (c, \zeta_C)$$

where the parameters come from the geodesics γ_A, γ_B, γ_C respectively. The image of this map will give an alternative description of the moduli space of Heisenberg triangles, analogous to describing a Euclidean triangle by its side lengths rather than by the vertices that determine it.

Since the Carnot-Caratheodory distance gives a metric, the Heisenberg side lengths are strictly positive and satisfy the (non-strict) triangle inequalites

$$a + b \geq c$$

$$b + c \geq a$$

$$c + a \geq b$$

and the ζ-values are in the range

$$-\pi < \zeta_A, \zeta_B, \zeta_C < \pi.$$

Equivalently, we can work with the Euclidean lengths $\tilde{a}$, $\tilde{b}$, $\tilde{c}$ of the line segments connecting the projections z_A, z_B, z_C of A, B, C onto the xy-plane. The Euclidean lengths are related to the Heisenberg lengths by the expressions

$$\tilde{a} = \frac{a}{\sigma(\zeta_A)}, \tilde{b} = \frac{b}{\sigma(\zeta_B)}, \tilde{c} = \frac{c}{\sigma(\zeta_C)}.$$

Since the projections z_A, z_B, z_C are distinct, the projected side-lengths are strictly positive and satisfy the (non-strict) triangle inequalities

$$\tilde{a} + \tilde{b} \geq \tilde{c}$$

$$\tilde{b} + \tilde{c} \geq \tilde{a}$$

$$\tilde{c} + \tilde{a} \geq \tilde{b}.$$

We proceed to show that in addition to the properties just outlined, the parameters $(a, \zeta_A), (b, \zeta_B), (c, \zeta_C)$ satisfy a defining equation.

We compute the parameters. For γ_A

$$\mu(\zeta_A) = \frac{t_C - t_B - 2(x_C y_B - y_C x_B)}{(x_C - x_B)^2 + (y_C - y_B)^2}$$

$$\tilde{a} = \sqrt{(x_C - x_B)^2 + (y_C - y_B)^2};$$

for γ_B

$$\mu(\zeta_B) = \frac{t_A - t_C - 2(x_A y_C - y_A x_C)}{(x_A - x_C)^2 + (y_A - y_C)^2}$$

$$\tilde{b} = \sqrt{(x_A - x_C)^2 + (y_A - y_C)^2};$$

for γ_C

$$\mu(\zeta_C) = \frac{t_B - t_A - 2(x_B y_A - y_B x_A)}{(x_B - x_A)^2 + (y_B - y_A)^2}$$

$$\tilde{c} = \sqrt{(x_B - x_A)^2 + (y_B - y_A)^2}.$$

From this we obtain

$$\begin{aligned}
\tilde{a}^2\mu(\zeta_A) + \tilde{b}^2\mu(\zeta_B) + \tilde{c}^2\mu(\zeta_C) &= t_C - t_B - 2(x_C y_B - y_C x_B) \\
&\quad + t_A - t_C - 2(x_A y_C - y_A x_C) \\
&\quad + t_B - t_A - 2(x_B y_A - y_B x_A) \\
&= -2(x_C y_B - y_C x_B) \\
&\quad - 2(x_A y_C - y_A x_C) \\
&\quad - 2(x_B y_A - y_B x_A) \\
&= 2\left[(x_B - x_A)(y_C - y_A) - (x_C - x_A)(y_B - y_A)\right] \\
&= 2\begin{vmatrix} x_B - x_A & y_B - y_A \\ x_C - x_A & y_C - y_A \end{vmatrix} \\
&= \pm 4 \text{ area } \triangle z_A z_B z_C.
\end{aligned}$$

We have $-$ when $\triangle z_A z_B z_C$ is oriented clockwise, and $+$ when $\triangle z_A z_B z_C$ is oriented counter-clockwise.

Now $\triangle z_A z_B z_C$ is a triangle in the xy-plane with sides of length $\tilde{a}$, $\tilde{b}$, $\tilde{c}$. As we have seen

$$\text{area } \triangle z_A z_B z_C = \frac{1}{4}\sqrt{(\tilde{a}^2 + \tilde{b}^2 + \tilde{c}^2)^2 - 2(\tilde{a}^4 + \tilde{b}^4 + \tilde{c}^4)}.$$

Therefore the triangle parameters satisfy the equation

$$\tilde{a}^2\mu(\zeta_A) + \tilde{b}^2\mu(\zeta_B) + \tilde{c}^2\mu(\zeta_C) = \pm\sqrt{(\tilde{a}^2 + \tilde{b}^2 + \tilde{c}^2)^2 - 2(\tilde{a}^4 + \tilde{b}^4 + \tilde{c}^4)}.$$

We collect these facts in a proposition.

Proposition 4.2.1. *Given vertices $(A, B, C) \in \mathbb{H}^3 - \Sigma$, the triangle parameters*

(a, ζ_A), (b, ζ_B), (c, ζ_C) are in the range

$$a, b, c > 0$$

$$-\pi < \zeta_A, \zeta_B, \zeta_C < \pi$$

satisfy the non-strict triangle inequalities

$$a + b \geq c, b + c \geq a, c + a \geq b$$

and the equation

$$\tilde{a}^2\mu(\zeta_A) + \tilde{b}^2\mu(\zeta_B) + \tilde{c}^2\mu(\zeta_C) = \pm\sqrt{(\tilde{a}^2 + \tilde{b}^2 + \tilde{c}^2)^2 - 2(\tilde{a}^4 + \tilde{b}^4 + \tilde{c}^4)}. \qquad (4.2.1)$$

Points in $\mathbb{R}^6$ that satisfy (4.2.1) and the inequalities listed in the above proposition will turn out to give a description of Heisenberg triangles up to isometry, so we give a name to the set of such points.

Definition 7. *Let $T \subset \mathbb{R}^6$ be the space of parameters (a, ζ_A), (b, ζ_B), (c, ζ_C) in the range*

$$a, b, c > 0$$

$$-\pi < \zeta_A, \zeta_B, \zeta_C < \pi$$

satisfying the non-strict triangle inequalities

$$a + b \geq c, b + c \geq a, c + a \geq b$$

and equation (4.2.1)

$$\tilde{a}^2\mu(\zeta_A) + \tilde{b}^2\mu(\zeta_B) + \tilde{c}^2\mu(\zeta_C) = \pm\sqrt{(\tilde{a}^2 + \tilde{b}^2 + \tilde{c}^2)^2 - 2(\tilde{a}^4 + \tilde{b}^4 + \tilde{c}^4)}$$

where $\tilde{a}$, $\tilde{b}$, $\tilde{c}$ are defined as before in terms of (a, ζ_A), (b, ζ_B), (c, ζ_C).

We will show that in fact T provides a description of the moduli space of Heisenberg triangles up to isometry in $\mathbb{H} \rtimes SO(2)$. Suppose we are given parameters

$$(a, \zeta_A), (b, \zeta_B), (c, \zeta_C) \in T.$$

We claim that we can find distinct points $A, B, C \in \mathbb{H}$ no one of which lies vertically above another, such that $\triangle ABC$ has the given parameters $(a, \zeta_A), (b, \zeta_B), (c, \zeta_C)$.

For the construction of such a triangle, we will assume that the sign in equation (4.2.1) is $-$. This detail shows up below in the fact that the triangle constructed has clockwise orientation. If the equation had been the $+$ version, the triangle constructed would have been constructed with counter-clockwise orientation.

To start we observe that $a > 0$ implies $\tilde{a} = \frac{a}{\sigma(\zeta_A)} > 0$, and likewise $\tilde{b}, \tilde{c} > 0$. Also in order for the right-hand side of equation (4.2.1) to be well-defined, we must have

$$(\tilde{a}^2 + \tilde{b}^2 + \tilde{c}^2)^2 - 2(\tilde{a}^4 + \tilde{b}^4 + \tilde{c}^4) \geq 0.$$

Recall that this is equivalent to $\tilde{a}, \tilde{b}, \tilde{c}$ satisfying the non-strict triangle inequalites

$$\tilde{a} + \tilde{b} \geq \tilde{c}, \tilde{b} + \tilde{c} \geq \tilde{a}, \tilde{c} + \tilde{a} \geq \tilde{b}.$$

These preliminary remarks show that, if we could find vertices

$$A(z_A, t_A), B(z_B, t_B), C(z_C, t_C)$$

corresponding to the specified parameters, then $\triangle ABC$ would project to three distinct points z_A, z_B, z_C in the xy-plane forming a (possibly degenerate) triangle with side lengths $\tilde{a}$, $\tilde{b}$, $\tilde{c}$. We will thus use a triangle in the xy-plane with side lengths $\tilde{a}$, $\tilde{b}$, $\tilde{c}$ as a kind of frame over which to construct the desired triangle.

By choosing the vertex A to be at the origin and the vertex B to be in the yt-plane with $y > 0$ we have enough information to compute what the vertices should be and check that they give the desired parameters. The vertex B must be at $(0, \tilde{c}, t_B)$ with t_B such that $\mu(\zeta_C) = t_B/\tilde{c}^2$. Thus B is determined.

Now if α is the angle opposite the side of length $\tilde{a}$ in a triangle with side-lengths $\tilde{a}$, $\tilde{b}$, $\tilde{c}$, then C must project to the point $(\tilde{b}\sin\alpha, \tilde{b}\cos\alpha)$ (the counter-clockwise orientation of the triangle shows up in this choice for C; clockwise orientation would require the projection to be $(-\tilde{b}\sin\alpha, \tilde{b}\cos\alpha)$). The vertex C must have $\mu(-\zeta_B) = t_C/\tilde{b}^2$.

Thus our vertices are

$$A = (0,0,0)$$

$$B = \big(0, \tilde{c}, \tilde{c}^2\mu(\zeta_C)\big)$$

$$C = \big(\tilde{b}\sin\alpha, \tilde{b}\cos\alpha, -\tilde{b}^2\mu(\zeta_B)\big).$$

Then we note that the ζ-value ζ_{BC} for the geodesic from B to C satisfies precisely the same equation

$$\tilde{a}^2\mu(\zeta_{BC}) + \tilde{b}^2\mu(\zeta_B) + \tilde{c}^2\mu(\zeta_C) = \pm\sqrt{(\tilde{a}^2+\tilde{b}^2+\tilde{c}^2)^2 - 2(\tilde{a}^4+\tilde{b}^4+\tilde{c}^4)}$$

as ζ_A. It follows that $\zeta_A = \zeta_{BC}$ and therefore the triangle we have constructed has the required parameters (a,ζ_A), (b,ζ_B), (c,ζ_C).

Finally, note that if the specified parameters (a,ζ_A), (b,ζ_B), (c,ζ_C) had been such that the right hand side of equation (4.2.1) was 0, the above construction still goes through. In this case the triangle ABC projects to a degenerate triangle $z_Az_Bz_C$ where the vertices are co-linear (but distinct). We summarize these results in a theorem.

Theorem 4.2.2. *The parameters (a,ζ_A), (b,ζ_B), (c,ζ_C) associated to a triangle ABC are in the range*

$$a, b, c > 0$$

$$-\pi < \zeta_A, \zeta_B, \zeta_C < \pi$$

satisfy the (non-strict) triangle inequalities

$$a + b \geq c$$

$$b + c \geq a$$

$$c + a \geq b$$

and equation (4.2.1)

$$\tilde{a}^2\mu(\zeta_A) + \tilde{b}^2\mu(\zeta_B) + \tilde{c}^2\mu(\zeta_C) = \pm\sqrt{(\tilde{a}^2+\tilde{b}^2+\tilde{c}^2)^2 - 2(\tilde{a}^4+\tilde{b}^4+\tilde{c}^4)}.$$

Moreover, given parameters $(a, \zeta_A), (b, \zeta_B), (c, \zeta_C) \in \mathbb{R}_+ \times (-\pi, \pi)$ *satisfying equation* (4.2.1) *and the non-strict triangle inequalities on* a, b, c, *there exist vertices* A, B, C *such that* $\triangle ABC$ *has* (a, ζ_A), (b, ζ_B), (c, ζ_C) *as parameters.*

Next we show that if two triangles coming from vertices (A, B, C) and (X, Y, Z) have identical triangle parameters

$$(a, \zeta_A), (b, \zeta_B), (c, \zeta_C) = (x, \zeta_X), (y, \zeta_Y), (z, \zeta_Z)$$

then $\triangle ABC$ and $\triangle XYZ$ can be made to coincide by an isometry of $\mathbb{H}$.

We do this by showing that both triangles can be made to coincide with the same "standard" triangle. A standard triangle refers to a triangle with first vertex at O, second vertex in the yt-plane with $y > 0$, and third vertex arbitrary.

Any triangle ABC can be mapped by isometries to such a triangle. We first left-translate by A^{-1} to put the vertex A at the origin, and then rotate until the vertex B hits the yt-plane with $y > 0$. (Although $\triangle ABC$ is specified by three fixed vertices A, B, C sometimes we continue to refer to a translated vertex by the same label as the original vertex, and imagine the vertex moving through $\mathbb{H}$.)

By looking at the "frame triangle", the triangle $z_A z_B z_C$ in the xy-plane to which $\triangle ABC$ projects, we will be able to see that in fact the standard triangle with which $\triangle ABC$ is congruent is unique and depends only on the triangle parameters $(a, \zeta_A), (b, \zeta_B), (c, \zeta_C)$.

To see this, note that since the two sets of triangle parameters are identical, they satisfy the defining equation (4.2.1) with the same sign, $+$ or $-$. This means that their respective frame triangles have the same orientation, counter-clockwise or clockwise. The frame triangles clearly have the same side-lengths as well

$$\frac{a}{\sigma(\zeta_A)} = \frac{x}{\sigma(\zeta_X)} \qquad \frac{b}{\sigma(\zeta_B)} = \frac{y}{\sigma(\zeta_Y)} \qquad \frac{c}{\sigma(\zeta_C)} = \frac{z}{\sigma(\zeta_Z)}.$$

Therefore the frame triangles of the two triangles are congruent and have the same orientation. Mapping the triangles to their respective standard triangles then causes the frame triangles to coincide. Thus each triangle ABC, XYZ is congruent to a standard triangle built out of geodesics over a frame with side lengths $\tilde{a} = \tilde{x}$, $\tilde{b} = \tilde{y}$, $\tilde{c} = \tilde{z}$. Moreover, the geodesics above each frame side have identical parameters $\zeta_A = \zeta_X$, $\zeta_B = \zeta_Y$, $\zeta_C = \zeta_Z$.

This shows that once mapped to standard triangles, the projections of A, B, C must coincide with the projections of X, Y, Z respectively. Just as in the construction of a triangle given fixed parameters, the respective t coordinates are determined. We have

$$t_B = \tilde{c}^2\mu(\zeta_C) = \tilde{z}^2\mu(\zeta_Z) = t_Y$$

and similarly $t_C = t_Z$.

For the converse, if g is an isometry in the component $\mathbb{H} \rtimes SO(2)$ of $\mathrm{Isom}(\mathbb{H})$, then the vertices (A, B, C) and (gA, gB, gC) determine triangles with identical parameters, since arc-length and ζ-values are preserved by such isometries.

Theorem 4.2.3. *The triangles ABC and XYZ are congruent by an isometry in $\mathbb{H} \rtimes SO(2)$ if and only if their triangle parameters are identical*

$$(a, \zeta_A), (b, \zeta_B), (c, \zeta_C) = (x, \zeta_X), (y, \zeta_Y), (z, \zeta_Z).$$

4.2.2 Moduli spaces for ordered Heisenberg triangles

With these results in hand, we can describe the moduli space of all possible triangles in the Heisenberg group in terms of the space of parameters. We know that to each triangle $(A, B, C) \in \mathbb{H}^3 - \Sigma$ there correspond parameters $(a, \zeta_A), (b, \zeta_B), (c, \zeta_C)$ in the image T of p. Conversely we know that to each set of parameters $(a, \zeta_A), (b, \zeta_B), (c, \zeta_C)$ in the image T of p there exist vertices (A, B, C) forming triangles with those parameters. Moreover, the triangles determined by the vertices (A, B, C) and (X, Y, Z) have the same parameters if and only if they are congruent by an isometry in $\mathbb{H} \rtimes SO(2)$. These remarks give the following.

Theorem 4.2.4. *The map p factors through the quotient space obtained by identifying* *congruent triangles*

$$\begin{array}{ccc} \mathbb{H}^3 - \Sigma & \xrightarrow{\quad p \quad} & T \subset \mathbb{R}^6 \\ \big\downarrow & \nearrow_{\tilde{p}} & \\ \mathbb{H} \rtimes SO(2) \backslash \left(\mathbb{H}^3 - \Sigma\right) & & \end{array}$$

where the map $\tilde{p}$ is a continuous bijection.

Proof. Notice that by theorem 4.2.3 the effect of the $\mathbb{H} \rtimes SO(2)$-action on $\mathbb{H}^3 - \Sigma$ is precisely to identify those triples (A, B, C) with the same image under p. Thus $\tilde{p}$ is injective, and $\tilde{p}$ is continuous and surjective because p is continuous and surjective.

□

In section 4.2.3 we will compute the rank of p and conclude that $p : \mathbb{H}^3 - \Sigma \longrightarrow T$ is an open map. Therefore $\tilde{p}$ will actually give a homeomorphism between the two characterizations of the space of Heisenberg triangles up to congruence by isometry in $\mathbb{H} \rtimes SO(2)$

$$\mathbb{H} \rtimes SO(2) \backslash \left(\mathbb{H}^3 - \Sigma\right) \cong T \subset \mathbb{R}^6.$$

To discuss the space T in the context of moduli space theory, we now give one possible definition of a family of ordered, oriented Heisenberg triangles. A few remarks are necessary to motivate the definition.

First, the space T characterizes ordered Heisenberg triangles up to isometry in $\mathbb{H} \rtimes SO(2)$, so the space T distinguishes between congruent triangles with opposite orientation. The definition of a family will have to be compatible with this. Second, recall that in the Euclidean case we took fibres X_s to be isometric to a Euclidean triangle, where a Euclidean triangle was viewed as a 2-dimensional subspace of the plane $\mathbb{R}^2$. A Heisenberg triangle cannot be filled in the way a Euclidean triangle can.

So for this definition we view a Heisenberg triangle with vertices (A, B, C) as the union $\gamma_A \cup \gamma_B \cup \gamma_C$ of its geodesic sides. The triangle $\gamma_A \cup \gamma_B \cup \gamma_C \subset \mathbb{H}$ then becomes a metric space with the metric induced from the Carnot-Caratheodory metric on the ambient space $\mathbb{H}$.

Definition 8. *For a topological space S, a family of ordered Heisenberg triangles over S is a subspace $X \subset S \times \mathbb{H}$ such that*

(1) projection onto the first coordinate is a (continuous, proper) fibre bundle projection

(2) each fibre $X_s = p^{-1}(s)$ is (after projection onto the second coordinate) a Heisenberg triangle $\gamma_A \cup \gamma_B \cup \gamma_C \subset \mathbb{H}$

(3) there is an ordered triple of sections $A, B, C : S \to X$ which specify the vertices $A(s)$, $B(s)$, $C(s)$ of each fibre X_s.

Morphisms between families are defined just as for ordered Euclidean triangles, except that isometries between fibres must be in $\mathbb{H} \rtimes SO(2)$.

A universal family $U \to T$ is the subspace $U \subset T \times \mathbb{H}$ with fibre over the point $(a, \zeta_A, b, \zeta_B, c, \zeta_C) \in T$ equal to the Heisenberg triangle $\gamma_A \cup \gamma_B \cup \gamma_C$ determined by the vertices (A, B, C) of the unique standard triangle having the parameters $(a, \zeta_A, b, \zeta_B, c, \zeta_C)$. The space T is therefore a fine moduli space for ordered Heisenberg triangles up to isometry in $\mathbb{H} \rtimes SO(2)$.

To complete the classification up to isometry in the full group of isometries Isom $(\mathbb{H}) = \mathbb{H} \rtimes O(2)$ we use the decomposition

$$\mathrm{Isom}(\mathbb{H}) = \big(\mathbb{H} \rtimes SO(2)\big) \cup r\big(\mathbb{H} \rtimes SO(2)\big)$$

where r is the involution isometry

$$(x, y, t) \mapsto (x, -y, -t).$$

Now the transformation of a triangle (A, B, C) to the triangle (rA, rB, rC) has the effect on parameters

$$(a, \zeta_A), (b, \zeta_B), (c, \zeta_C) \mapsto (a, -\zeta_A), (b, -\zeta_B), (c, -\zeta_C).$$

Hence allowing the full group of isometries to act on the space of vertices $\widetilde{T} = \mathbb{H}^3 - \Sigma$ has the effect on T of taking the quotient by the identification

$$(a, \zeta_A), (b, \zeta_B), (c, \zeta_C) \sim (a, -\zeta_A), (b, -\zeta_B), (c, -\zeta_C).$$

The definition of a family of ordered Heisenberg triangles over a topological space S is unchanged, except that now morphisms are allowed to restrict to orientation reversing isometries on fibres, i.e., isometries can come from the full group $\mathbb{H} \rtimes O(2)$. The universal family U' is now over the quotient space of T by $\sim$, with fibre over the point $[(a, \zeta_A, b, \zeta_B, c, \zeta_C), (a, -\zeta_A, b, -\zeta_B, c, -\zeta_C)]$ coming from the unique clockwise oriented standard triangle having the specified parameters. Again we obtain a fine moduli space.

Since we have characterized the moduli space of Heisenberg triangles in terms of the image of the map p, we take a closer look at the properties of this map.

4.2.3 Manifold structure of the parameter space

For a triple

$$(x_A, y_A, t_A; x_B, y_B, t_B; x_C, y_C, t_C) \in \mathbb{H}^3 - \Sigma$$

we obtain the geodesic parameters by the functions

$$\zeta_A = \mu^{-1}\left(\frac{t_C - t_B - 2(x_C y_B - y_C x_B)}{(x_C - x_B)^2 + (y_C - y_B)^2}\right)$$

$$a = \sigma(\zeta_A)\sqrt{(x_C - x_B)^2 + (y_C - y_B)^2}$$

with analogous functions giving b, ζ_B, c, ζ_C. The first step will be to determine the rank of p. To do so, we decompose p into successive transformations.

The first transformation is $p_1 : \mathbb{R}^9 \longrightarrow \mathbb{R}^6$ which sends

$$\begin{pmatrix} x_A, y_A, t_A; \\ x_B, y_B, t_B; \\ x_C, y_C, t_C \end{pmatrix} \mapsto \begin{pmatrix} (x_C - x_B)^2 + (y_C - y_B)^2, & t_C - t_B - 2(x_C y_B - y_C x_B) \\ (x_A - x_C)^2 + (y_A - y_C)^2, & t_A - t_C - 2(x_A y_C - y_A x_C) \\ (x_B - x_A)^2 + (y_B - y_A)^2, & t_B - t_A - 2(x_B y_A - y_B x_A) \end{pmatrix}.$$

This is the key component of the transformation. The remaining maps p_2, p_3, p_4, p_5 will map $\mathbb{R}^6 \longrightarrow \mathbb{R}^6$ and have the full rank 6 on their domains. Below we express the maps in coordinates; the letters used for the coordinates are meant to be descriptive, but the reader should remember that they are really just Euclidean coordinates. The maps are:

$p_2 : \mathbb{R}^6 \longrightarrow \mathbb{R}^6$ which sends

$$(D_1, t_1, D_2, t_2, D_3, t_3) \mapsto \left(D_1, \frac{t_1}{D_1}, D_2, \frac{t_2}{D_2}, D_3, \frac{t_3}{D_3}\right);$$

$p_3 : \mathbb{R}^6 \longrightarrow \mathbb{R}^6$ which sends

$$(D_1, \lambda_1, D_2, \lambda_2, D_3, \lambda_3) \mapsto \left(D_1, \mu^{-1}(\lambda_1), D_2, \mu^{-1}(\lambda_2), D_3, \mu^{-1}(\lambda_3)\right);$$

$p_4 : \mathbb{R}^6 \longrightarrow \mathbb{R}^6$ which sends

$$(D_1, \zeta_1, D_2, \zeta_2, D_3, \zeta_3) \mapsto \left(\sqrt{D_1}, \zeta_1, \sqrt{D_2}, \zeta_2, \sqrt{D_3}, \zeta_3\right);$$

$p_5 : \mathbb{R}^6 \longrightarrow \mathbb{R}^6$ which sends

$$(d_1, \zeta_1, d_2, \zeta_2, d_3, \zeta_3) \mapsto (\sigma(\zeta_1)d_1, \zeta_1, \sigma(\zeta_2)d_2, \zeta_2, \sigma(\zeta_3)d_3, \zeta_3,).$$

Observe that indeed $p = p_5p_4p_3p_2p_1$. We look first at the differentials of the maps p_2, p_3, p_4, p_5 in turn, leaving p_1 for last because the discussion will take substantially longer.

For p_2 we have

$$Dp_2 = \begin{bmatrix} 1 & \frac{-t_1}{D_1^2} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{D_1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & \frac{-t_2}{D_2^2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{D_2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \frac{-t_3}{D_3^2} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{D_3} \end{bmatrix}.$$

Now Dp_2 has rank 6 on the image of $\mathbb{H}^3 - \Sigma$ under p_1, because that is precisely where each $D_i \neq 0$. (In fact $p_1(\mathbb{H}^3 - \Sigma) \subset \mathbb{R}^6$ is precisely the domain of p_2.)

For p_3 we have

$$Dp_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & (\mu^{-1})'(\lambda_1) & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & (\mu^{-1})'(\lambda_2) & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & (\mu^{-1})'(\lambda_3) \end{bmatrix}$$

where the rank is 6 because $\mu^{-1} : \mathbb{R} \longrightarrow (-\pi, \pi)$ has strictly non-zero derivative.

For p_4 we have

$$Dp_4 = \begin{bmatrix} \frac{1}{2\sqrt{D_1}} & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2\sqrt{D_2}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2\sqrt{D_3}} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

where the rank is 6 because, as we have seen, we have each $D_i \neq 0$ on our domain.

For p_5 we have

$$Dp_5 = \begin{bmatrix} \sigma(\zeta_1) & 0 & 0 & 0 & 0 & 0 \\ \sigma'(\zeta_1)d_1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma(\zeta_2) & 0 & 0 & 0 \\ 0 & 0 & \sigma'(\zeta_2)d_2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma(\zeta_3) & 0 \\ 0 & 0 & 0 & 0 & \sigma'(\zeta_3)d_3 & 1 \end{bmatrix}$$

where once again the rank is 6 because each $\sigma(\zeta_i) \geq 1$.

Thus we have $Dp = Dp_5 Dp_4 Dp_3 Dp_2 Dp_1$ where the matrices Dp_2, Dp_3, Dp_4, Dp_5 are each 6×6 matrices with rank 6 and consequently are invertible. Thus for the two linear transformations

$$Dp : \mathbb{R}^9 \longrightarrow \mathbb{R}^6$$

$$Dp_1 : \mathbb{R}^9 \longrightarrow \mathbb{R}^6$$

we find that the dimension of the image of Dp equals the dimension of the image of Dp_1, that is, the rank of Dp equals the rank of Dp_1.

At last we consider the differential of the map p_1. The matrix Dp_1 is

$$\begin{bmatrix} 0 & 0 & 0 & -2(x_C - x_B) & -2(y_C - y_B) & 0 & 2(x_C - x_B) & 2(y_C - y_B) & 0 \\ 0 & 0 & 0 & 2y_C & -2x_C & -1 & -2y_B & 2x_B & 1 \\ 2(x_A - x_C) & 2(y_A - y_C) & 0 & 0 & 0 & 0 & -2(x_A - x_C) & -2(y_A - y_C) & 0 \\ -2y_c & 2x_C & 1 & 0 & 0 & 0 & 2y_A & -2x_A & -1 \\ -2(x_B - x_A) & -2(y_B - y_A) & 0 & 2(x_B - x_A) & 2(y_B - y_A) & 0 & 0 & 0 & 0 \\ 2y_B & -2x_B & -1 & -2y_A & 2x_A & 1 & 0 & 0 & 0 \end{bmatrix}.$$

It is not as bad as it looks. We label the rows

$$Dp_1 = \begin{bmatrix} \vec{r}_1 \\ \vdots \\ \vec{r}_6 \end{bmatrix} \qquad \vec{r}_i \in \mathbb{R}^9$$

and show first that the rows are *not* independent, and therefore that the rank of the matrix is strictly less than 6. We do this by showing that the equation

$$\alpha\vec{r}_1 + \lambda_1\vec{r}_2 + \beta\vec{r}_3 + \lambda_2\vec{r}_4 + \gamma\vec{r}_5 + \lambda_3\vec{r}_6 = 0$$

has a non-trivial solution for real scalars α, λ_1, β, λ_2, γ, λ_3.

If such a solution was found, then by looking at the 3rd, 6th, and 9th components of this vector equation, it would be apparent that $\lambda_1 = \lambda_2 = \lambda_3$. Thus a non-trivial solution exists if and only a non-trivial solution exists for the equation

$$\alpha\vec{r}_1 + \lambda\vec{r}_2 + \beta\vec{r}_3 + \lambda\vec{r}_4 + \gamma\vec{r}_5 + \lambda\vec{r}_6 = 0.$$

In this form, the 3rd, 6th, and 9th components of the vector equation merely tell us $0 = 0$, but the remaining six components give us a system of six equations in four unknowns α, β, γ, λ:

$$\beta(x_A - x_C) - \gamma(x_B - x_A) - \lambda(y_C - y_B) = 0 \tag{4.2.2}$$

$$\beta(y_A - y_C) - \gamma(y_B - y_A) + \lambda(x_C - x_B) = 0 \tag{4.2.3}$$

$$-\alpha(x_C - x_B) + \gamma(x_B - x_A) - \lambda(y_A - y_C) = 0 \tag{4.2.4}$$

$$-\alpha(y_C - y_B) + \gamma(y_B - y_A) + \lambda(x_A - x_C) = 0 \tag{4.2.5}$$

$$\alpha(x_C - x_B) - \beta(x_A - x_C) - \lambda(y_B - y_A) = 0 \tag{4.2.6}$$

$$\alpha(y_C - y_B) - \beta(y_A - y_C) + \lambda(x_B - x_A) = 0. \tag{4.2.7}$$

But in this system

$$(3) + (5) = -(1)$$

$$(4) + (6) = -(2)$$

so that the system reduces to just equations (3), (4), (5), (6), four equations in four unknowns:

$$-\alpha(x_C - x_B) + \gamma(x_B - x_A) - \lambda(y_A - y_C) = 0$$

$$-\alpha(y_C - y_B) + \gamma(y_B - y_A) + \lambda(x_A - x_C) = 0$$

$$\alpha(x_C - x_B) - \beta(x_A - x_C) - \lambda(y_B - y_A) = 0$$

$$\alpha(y_C - y_B) - \beta(y_A - y_C) + \lambda(x_B - x_A) = 0.$$

We set up the 4×4 matrix for this system, and embark on a determinant calculation. We temporarily adopt the notation

$$x_{CB} = x_C - x_B, \text{ etc.}$$

The matrix is

$$M = \begin{bmatrix} -x_{CB} & 0 & x_{BA} & -y_{AC} \\ -y_{CB} & 0 & y_{BA} & x_{AC} \\ x_{CB} & -x_{AC} & 0 & -y_{BA} \\ y_{CB} & -y_{AC} & 0 & x_{BA} \end{bmatrix}.$$

But we have the relations

$$-x_{AC} = -x_A + x_C$$

$$= -x_A + x_B - x_B + x_C = x_{BA} + x_{CB}$$

$$-y_{AC} = -y_A + y_C$$

$$= -y_A + y_B - y_B + y_C = y_{BA} + y_{CB}$$

so the matrix becomes

$$M = \begin{bmatrix} -x_{CB} & 0 & x_{BA} & y_{BA}+y_{CB} \\ -y_{CB} & 0 & y_{BA} & -x_{BA}-x_{CB} \\ x_{CB} & x_{BA}+x_{CB} & 0 & -y_{BA} \\ y_{CB} & y_{BA}+y_{CB} & 0 & x_{BA} \end{bmatrix}.$$

For the determinant, we expand along the third column:

$$\det M = (x_{BA}) \begin{vmatrix} -y_{CB} & 0 & -x_{BA}-x_{CB} \\ x_{CB} & x_{BA}+x_{CB} & -y_{BA} \\ y_{CB} & y_{BA}+y_{CB} & x_{BA} \end{vmatrix}$$

$$- (y_{BA}) \begin{vmatrix} -x_{CB} & 0 & y_{BA}+y_{CB} \\ x_{CB} & x_{BA}+x_{CB} & -y_{BA} \\ y_{CB} & y_{BA}+y_{CB} & x_{BA} \end{vmatrix}$$

$$= (x_{BA}) \det M_1 - (y_{BA}) \det M_2.$$

Expanding along the top row in M_1

$$\det M_1 = (-y_{CB}) \begin{vmatrix} x_{BA}+x_{CB} & -y_{BA} \\ y_{BA}+y_{CB} & x_{BA} \end{vmatrix} + (-x_{BA}-x_{CB}) \begin{vmatrix} x_{CB} & x_{BA}+x_{CB} \\ y_{CB} & y_{BA}+y_{CB} \end{vmatrix}$$

$$= (-y_{CB}) \left[(x_{BA}+x_{CB})x_{BA} - (-y_{BA})(y_{BA}+y_{CB})\right]$$

$$- (x_{BA}+x_{CB}) \left[x_{CB}(y_{BA}+y_{CB}) - (x_{BA}+x_{CB})y_{CB}\right]$$

$$= (-y_{CB}) \left[x_{BA}^2 + x_{CB}x_{BA} + y_{BA}^2 + y_{CB}y_{BA}\right]$$

$$- (x_{BA}+x_{CB}) \left[x_{CB}y_{BA} - x_{BA}y_{CB}\right]$$

$$= -x_{BA}^2 y_{CB} - x_{CB}x_{BA}y_{CB} - y_{CB}y_{BA}^2 - y_{CB}^2 y_{BA}$$

$$- x_{CB}x_{BA}y_{BA} + x_{BA}^2 y_{CB} - x_{CB}^2 y_{BA} + x_{CB}x_{BA}y_{CB}$$

$$= -y_{CB}y_{BA}^2 - y_{CB}^2 y_{BA} - x_{CB}x_{BA}y_{BA} - x_{CB}^2 y_{BA}.$$

Expanding along the top row in M_2

$$\det M_2 = (-x_{CB}) \begin{vmatrix} x_{BA} + x_{CB} & -y_{BA} \\ y_{BA} + y_{CB} & x_{BA} \end{vmatrix} + (-y_{BA} - y_{CB}) \begin{vmatrix} x_{CB} & x_{BA} + x_{CB} \\ y_{CB} & y_{BA} + y_{CB} \end{vmatrix}$$

$$= (-x_{CB}) \left[(x_{BA} + x_{CB})x_{BA} - (-y_{BA})(y_{BA} + y_{CB})\right]$$

$$+ (y_{BA} + y_{CB}) \left[x_{CB}(y_{BA} + y_{CB}) - (x_{BA} + x_{CB})y_{CB}\right]$$

$$= (-x_{CB}) \left[x_{BA}^2 + x_{CB}x_{BA} + y_{BA}^2 + y_{CB}y_{BA}\right]$$

$$+ (y_{BA} + y_{CB}) \left[x_{CB}y_{BA} - x_{BA}y_{CB}\right]$$

$$= -x_{CB}x_{BA}^2 - x_{CB}^2x_{BA} - x_{CB}y_{BA}^2 - x_{CB}y_{CB}y_{BA}$$

$$+ x_{CB}y_{BA}^2 - x_{BA}y_{CB}y_{BA} + x_{CB}y_{CB}y_{BA} - x_{BA}y_{CB}^2$$

$$= -x_{CB}x_{BA}^2 - x_{CB}^2x_{BA} - x_{BA}y_{CB}y_{BA} - x_{BA}y_{CB}^2.$$

Then

$$\det M = (x_{BA}) \det M_1 - (y_{BA}) \det M_2$$

$$= (x_{BA}) \left[-y_{CB}y_{BA}^2 - y_{CB}^2y_{BA} - x_{CB}x_{BA}y_{BA} - x_{CB}^2y_{BA}\right]$$

$$- (y_{BA}) \left[-x_{CB}x_{BA}^2 - x_{CB}^2x_{BA} - x_{BA}y_{CB}y_{BA} - x_{BA}y_{CB}^2\right]$$

$$= -x_{BA}y_{CB}y_{BA}^2 - x_{BA}y_{CB}^2y_{BA} - x_{CB}x_{BA}^2y_{BA} - x_{CB}^2x_{BA}y_{BA}$$

$$x_{CB}x_{BA}^2y_{BA} + x_{CB}^2x_{BA}y_{BA} + x_{BA}y_{CB}y_{BA}^2 + x_{BA}y_{CB}^2y_{BA}$$

$$= 0.$$

Therefore there exist non-trivial solutions α, β, γ, λ to the equation

$$\alpha\vec{r}_1 + \lambda\vec{r}_2 + \beta\vec{r}_3 + \lambda\vec{r}_4 + \gamma\vec{r}_5 + \lambda\vec{r}_6 = 0.$$

We conclude that the rows of Dp_1 are dependent, and therefore that the rank of Dp_1 is strictly less than 6.

Now we show that the rank is 5. Indeed, the first four rows of Dp_1 are independent. If $\vec{r}_5$ cannot be expressed as a linear combination of the first four rows, the first five rows are independent and we are done. Otherwise suppose that $\vec{r}_5$ can be expressed as a linear combination of $\vec{r}_1$, $\vec{r}_2$, $\vec{r}_3$, $\vec{r}_4$. We will show that in this case, $\vec{r}_6$ cannot be expressed as a linear combination of the first four rows.

So now supposing that $\vec{r}_5$ can be expressed as a linear combination of $\vec{r}_1$, $\vec{r}_2$, $\vec{r}_3$, $\vec{r}_4$, then by looking at the 3rd and 6th columns of Dp_1, we can see that the coefficients of $\vec{r}_2$, $\vec{r}_4$ must be 0. Thus we would have

$$\alpha\vec{r}_1 + \beta\vec{r}_3 = \vec{r}_5.$$

The components of this vector equation yield six equations (nine actually, but three are $0 = 0$):

$$\beta(x_A - x_C) = -(x_B - x_A)$$

$$\beta(y_A - y_C) = -(y_B - y_A)$$

$$-\alpha(x_C - x_B) = (x_B - x_A)$$

$$-\alpha(y_C - y_B) = (y_B - y_A)$$

$$\alpha(x_C - x_B) = \beta(x_A - x_C)$$

$$\alpha(y_C - y_B) = \beta(y_A - y_C).$$

But notice that the first and third equations above imply the fifth, and the second and fourth equations imply the sixth, so that the system is equivalent to just the four equations

$$\beta(x_A - x_C) = -(x_B - x_A) \tag{4.2.8}$$

$$\beta(y_A - y_C) = -(y_B - y_A) \tag{4.2.9}$$

$$-\alpha(x_C - x_B) = (x_B - x_A) \tag{4.2.10}$$

$$-\alpha(y_C - y_B) = (y_B - y_A). \tag{4.2.11}$$

Recall that p_1 is defined on the domain $\mathbb{H}^3 - \Sigma$. On this domain $(x_A, y_A) \neq (x_B, y_B)$. This fact and equations (1) and (2) together imply that $\beta \neq 0$. Likewise $\alpha \neq 0$.

With $\alpha, \beta \neq 0$, we can see from equation (1) that $x_A - x_C = 0$ if and only if $x_B - x_A = 0$, and from equation (3) that $x_B - x_A = 0$ if and only if $x_C - x_B = 0$. Therefore x_A, x_B, x_C are either all equal or all distinct. Note that when the x's are equal the y's must be distinct in order for the points A, B, C to be distinct.

Likewise from (2) we have $y_A - y_C = 0$ if and only if $y_B - y_A = 0$, and from (4) we have $y_B - y_A = 0$ if and only if $y_C - y_B = 0$. Therefore y_A, y_B, y_C are either all equal or all distinct. Note that when the y's are equal the x's must be distinct in order for the points A, B, C to be distinct.

There are thus three cases:

(i) $x_A = x_B = x_C$

(ii) $y_A = y_B = y_C$

(iii) x_A, x_B, x_C are distinct, and y_A, y_B, y_C are distinct.

Case (i): $x_A = x_B = x_C$ and the y's are distinct. Then the rows $\vec{r}_1$, $\vec{r}_2$, $\vec{r}_3$, $\vec{r}_4$ are

$$\begin{array}{lccccccccr}
(& 0 & 0 & 0 & 0 & -2(y_C - y_B) & 0 & 0 & 2(y_C - y_C) & 0 &) \\
(& 0 & 0 & 0 & 2y_C & -2x_C & -1 & -2y_B & 2x_B & 1 &) \\
(& 0 & 2(y_A - y_C) & 0 & 0 & 0 & 0 & 0 & -2(y_A - y_C) & 0 &) \\
(& -2y_C & 2x_C & 1 & 0 & 0 & 0 & 2y_A & -2x_A & -1 &)
\end{array}$$

and the sixth row $\vec{r}_6$ is

$$\begin{pmatrix} 2y_B & -2x_B & -1 & -2y_A & 2x_A & 1 & 0 & 0 & 0 \end{pmatrix}.$$

If $\vec{r}_6$ were to be expressed as a linear combination of $\vec{r}_1$, $\vec{r}_2$, $\vec{r}_3$, $\vec{r}_4$, it is apparent from the 3rd and 6th columns that the coefficients of $\vec{r}_2$ and $\vec{r}_4$ must both be -1. But then the 1st column tells us that $2y_C = 2y_B$, contradicting that the y's are distinct. Thus $\vec{r}_6$ cannot be expressed as a linear combination of the four independent vectors $\vec{r}_1$, $\vec{r}_2$, $\vec{r}_3$, $\vec{r}_4$ and therefore the rank of the matrix Dp_1 is 5.

Case (ii): $y_A = y_B = y_C$ and the x's are distinct. Then the rows $\vec{r}_1$, $\vec{r}_2$, $\vec{r}_3$, $\vec{r}_4$ are

$$\begin{array}{cccccccccc} (& 0 & 0 & 0 & -2(x_C - x_B) & 0 & 0 & 2(x_C - x_B) & 0 & 0 &) \\ (& 0 & 0 & 0 & 2y_C & -2x_C & -1 & -2y_B & 2x_B & 1 &) \\ (& 2(x_A - x_C) & 0 & 0 & 0 & 0 & 0 & -2(x_A - x_C) & 0 & 0 &) \\ (& -2y_C & 2x_C & 1 & 0 & 0 & 0 & 2y_A & -2x_A & -1 &) \end{array}$$

and the sixth row $\vec{r}_6$ is

$$\begin{pmatrix} 2y_B & -2x_B & -1 & -2y_A & 2x_A & 1 & 0 & 0 & 0 \end{pmatrix}.$$

Once again the vectors $\vec{r}_1$, $\vec{r}_2$, $\vec{r}_3$, $\vec{r}_4$, $\vec{r}_6$ are independent and the rank of the matrix Dp_1 is 5.

Case (iii): the x's are distinct, and the y's are distinct. We can now solve for α in equations (3) and (4), and for β in equations (1) and (2). We have

$$\alpha = \frac{x_A - x_B}{x_C - x_B} = \frac{y_A - y_B}{y_C - y_B}$$

$$\beta = \frac{x_B - x_A}{x_C - x_A} = \frac{y_B - y_A}{y_C - y_A}.$$

This is equivalent to

$$\frac{y_C - y_B}{x_C - x_B} = \frac{y_A - y_B}{x_A - x_B} = k$$

$$\frac{y_C - y_A}{x_C - x_A} = \frac{y_B - y_A}{x_B - x_A} = k.$$

Note that $k \neq 0$ because the y's are distinct. Incidentally, these equations say that (x_A, y_A), (x_B, y_B), (x_C, y_C) are colinear. Using the above relations involving the constant k, we write out the rows $\vec{r}_1$, $\vec{r}_2$, $\vec{r}_3$, $\vec{r}_4$:

$$\begin{array}{ccccccccccc}
(& 0 & 0 & 0 & -2(x_C - x_B).. & -2k(x_C - x_B) & 0 & 2(x_C - x_B) & 2k(x_C - x_B) & 0 &) \\
(& 0 & 0 & 0 & 2y_C & -2x_C & -1 & -2y_B & 2x_B & 1 &) \\
(& 2(x_A - x_C) & 2k(x_A - x_C) & 0 & 0 & 0 & 0 & -2(x_A - x_C) & -2k(x_A - x_C) & 0 &) \\
(& -2y_C & 2x_C & 1 & 0 & 0 & 0 & 2y_A & -2x_A & -1 &)
\end{array}$$

and the sixth row $\vec{r}_6$ is

$$(\ 2y_B \quad -2x_B \quad -1 \quad -2y_A \quad 2x_A \quad 1 \quad 0 \quad 0 \quad 0 \ ).$$

As before, if $\vec{r}_6$ were to be expressed as a linear combination of $\vec{r}_1$, $\vec{r}_2$, $\vec{r}_3$, $\vec{r}_4$, it is apparent from the 3rd and 6th columns that the coefficients of $\vec{r}_2$ and $\vec{r}_4$ must both be -1. And then if

$$\mu\vec{r}_1 - \vec{r}_2 + \nu\vec{r}_3 - \vec{r}_4 = \vec{r}_6$$

the first column tells us

$$2\nu(x_A - x_C) + 2y_C = 2y_B \text{ or } \nu = \frac{y_B - y_C}{x_A - x_C}$$

while the second column tells us

$$2\nu k(x_A - x_C) - 2x_C = -2x_B \text{ or } \nu = \frac{1}{k}\frac{x_C - x_B}{x_A - x_C} = -\frac{x_C - x_B}{y_C - y_A}.$$

Equating the expressions for ν gives

$$\frac{y_B - y_C}{x_A - x_C} = -\frac{x_C - x_B}{y_C - y_A}$$

$$\text{or } \frac{y_C - y_B}{x_C - x_B} = -\frac{x_C - x_A}{y_C - y_A}$$

$$\text{or } k = -\frac{1}{k}$$

$$\text{or } k^2 = -1$$

a contradiction.

Thus we have:

Proposition 4.2.5. *The rank of Dp_1 is 5 everywhere on its domain.*

Corollary 4.2.6. *The rank of Dp is 5 everywhere on its domain.*

The Rank theorem (theorem 1.1.5) now gives the following:

Theorem 4.2.7. *The moduli space of Heisenberg triangles, represented as the space of parameters*

$$(a, \zeta_A), (b, \zeta_B), (c, \zeta_C) \in T \subset \mathbb{R}^6$$

is a 5-dimensional manifold.

Moreover

Theorem 4.2.8. *We have the homeomorphism*

$$\mathbb{H} \rtimes SO(2) \backslash \big(\mathbb{H}^3 - \Sigma\big) \cong T \subset \mathbb{R}^6.$$

Proof. By theorem 4.2.4 the map $\tilde{p} : \mathbb{H} \rtimes SO(2) \backslash \big(\mathbb{H}^3 - \Sigma\big) \longrightarrow T$ is a continuous bijection. In fact since p is a rank 5 mapping into the 5-dimensional manifold T, p is an open map. It follows that $\tilde{p}$ is an open map, i.e., that $\tilde{p}^{-1}$ is continuous.

□

4.2.4 Fibre bundle structure of the map p

We have the commutative diagram showing the factorization of the map p

$$\begin{array}{ccc} \mathbb{H}^3 - \Sigma & \xrightarrow{\quad p \quad} & T \subset \mathbb{R}^6 \\ \downarrow & \nearrow_{\tilde{p}} & \\ \mathbb{H} \rtimes SO(2)\backslash(\mathbb{H}^3 - \Sigma) & & \end{array}$$

where $\tilde{p}$ is a homeomorphism.

We will use facts about proper group actions to see that the quotient map

$$\mathbb{H}^3 - \Sigma \longrightarrow \mathbb{H} \rtimes SO(2)\backslash(\mathbb{H}^3 - \Sigma)$$

has a fibre bundle structure, and therefore that the map

$$p : \mathbb{H}^3 - \Sigma \longrightarrow T$$

likewise has a fibre bundle structure. Our first goal is to show that $\mathbb{H} \rtimes SO(2)$ acts properly on the space of vertices $\mathbb{H}^3 - \Sigma$. We decompose this action into two stages, first the action of the group $\mathbb{H}$ of left-translations on the vertices, and second the action of the group $SO(2)$ of rotations on the vertices. Once we have shown that the successive group actions of

$$\mathbb{H} \text{ on } \mathbb{H}^3 - \Sigma \qquad\qquad SO(2) \text{ on } \mathbb{H}\backslash(\mathbb{H}^3 - \Sigma)$$

are proper we can conclude that the composition of these two actions is proper (the composition of proper group actions is proper). But first we must check that the successive action of $\mathbb{H}$ and then of $SO(2)$ is the same as the action of $\mathbb{H} \rtimes SO(2)$, i.e., that

$$SO(2)\backslash\mathbb{H}\backslash(\mathbb{H}^3 - \Sigma) = \mathbb{H} \rtimes SO(2)\backslash(\mathbb{H}^3 - \Sigma).$$

In fact it is not even immediately apparent that the action of $SO(2)$ on $\mathbb{H}\backslash(\mathbb{H}^3 - \Sigma)$ is well-defined so we check this first, with the help of the next proposition.

Proposition 4.2.9. *Suppose the points $A, B \in \mathbb{H}$ are separated by the left-translation $V \in \mathbb{H}$, so that if $V = (u, v, s)$ then*

$$A = (x, y, t)$$

$$B = VA = \big(u + x, v + y, s + t - 2(uy - vx)\big).$$

Then for $0 \leq \varphi \leq 2\pi$ the rotated points $R_\varphi A$, $R_\varphi B$ are separated by the left-translation

$$R_\varphi V = (u\cos\varphi - v\sin\varphi, u\sin\varphi + v\cos\varphi, s).$$

In particular, the displacement between the rotated points $R_\varphi A$, $R_\varphi B$ depends only on the rotation R_φ and on the displacement V between the original points A, B and not on the precise location of A and B in $\mathbb{H}$.

Proof. The result follows from a routine calculation. The rotated points are

$$R_\varphi A = (x\cos\varphi - y\sin\varphi, x\sin\varphi + y\cos\varphi, t)$$

$$R_\varphi B = \big((u + x)\cos\varphi - (v + y)\sin\varphi, (u + x)\sin\varphi + (v + y)\cos\varphi, s + t - 2(uy - vx)\big).$$

To find the displacement from $R_\varphi A$ to $R_\varphi B$ one calculates $R_\varphi B\big(R_\varphi A\big)^{-1}$. The result is the displacement vector $R_\varphi V$ as stated.

□

We remark that by looking only at the first two coordinates, the above proposition proves an analogous result for points in $\mathbb{R}^2$ or $\mathbb{C}$ where displacement is just the usual vector addition. In this setting, however, the reader would likely accept a "proof" by visualization.

Proposition 4.2.10. *The action of $SO(2)$ on $\mathbb{H}\backslash\big(\mathbb{H}^3 - \Sigma\big)$ induced by*

$$R_\varphi(A, B, C) = \big(R_\varphi A, R_\varphi B, R_\varphi C\big)$$

for $0 \leq \varphi \leq 2\pi$ and $(A, B, C) \in \mathbb{H}^3 - \Sigma$ is well-defined.

Proof. A point in $\mathbb{H}\backslash(\mathbb{H}^3-\Sigma)$ is an equivalence class of triples mutually related by left-translation. Suppose then that we have two representatives of the same equivalence class

$$[(A,B,C)] = [(XA,XB,XC)]$$

where $(A,B,C) \in \mathbb{H}^3 - \Sigma$ and $X \in \mathbb{H}$. A group element $R_\varphi \in SO(2)$ has the effect

$$[(A,B,C)] \mapsto [(R_\varphi A, R_\varphi B, R_\varphi C)]$$

$$[(XA,XB,XC)] \mapsto \left[((R_\varphi(XA), R_\varphi(XB), R_\varphi(XC))\right].$$

But by proposition 4.2.9 the displacements between $R_\varphi A$ and $R_\varphi(XA)$, between $R_\varphi B$ and $R_\varphi(XB)$, and between $R_\varphi C$ and $R_\varphi(XC)$ are all equal to $R_\varphi X$. Therefore $[(R_\varphi A, R_\varphi B, R_\varphi C)]$ and $\left[(R_\varphi(XA), R_\varphi(XB), R_\varphi(XC))\right]$ are equal in $\mathbb{H}\backslash(\mathbb{H}^3-\Sigma)$ and the $SO(2)$-action is well-defined.

□

Proposition 4.2.11. *The composite action $SO(2)\circ\mathbb{H}$ on $\mathbb{H}^3-\Sigma$ and the action of $\mathbb{H}\rtimes SO(2)$ on $\mathbb{H}^3-\Sigma$ have identical orbits. Therefore*

$$SO(2)\backslash\mathbb{H}\backslash(\mathbb{H}^3-\Sigma) = \mathbb{H}\rtimes SO(2)\backslash(\mathbb{H}^3-\Sigma).$$

Proof. If the triples (A,B,C) and (P,Q,R) are identified under the composite group action of $SO(2)$ and $\mathbb{H}$ it means that some translate of (A,B,C) and some translate of (P,Q,R) are related by a rotation, for example

$$R\circ V\circ(A,B,C) = W\circ(P,Q,R)$$

for $V,W\in\mathbb{H}$ and $R\in SO(2)$, whereupon $(P,Q,R) = W^{-1}\circ R\circ V\circ(A,B,C)$ so that (A,B,C) and (P,Q,R) get identified under the $\mathbb{H}\rtimes SO(2)$-action on $\mathbb{H}^3-\Sigma$.

Conversely suppose the triples (A,B,C) and (P,Q,R) are related by an isometry $g\in\mathbb{H}\rtimes SO(2)$. The left-translations and the rotations generate $\mathbb{H}\rtimes SO(2)$ so that g can be written as a finite sequence $g = g_1g_2\ldots g_n$ where each g_i is in either $\mathbb{H}$ or $SO(2)$ so that

$$(A,B,C) = g_1g_2\ldots g_n(P,Q,R).$$

If $g_n\in\mathbb{H}$ then the points (P,Q,R) and (g_nP, g_nQ, g_nR) get identified by the $\mathbb{H}$-action and the problem is reduced to showing that (A,B,C) and (g_nP,g_nQ,g_nR) get identified, with strictly fewer isometries relating the triples. If $g_n \in SO(2)$ then the equivalence classes $[(P, Q, R)]$ and $[(g_nP, g_nQ, g_nR)]$ get identified by the $SO(2)$-action and the problem is once again reduced to showing that (A, B, C) and (g_nP, g_nQ, g_nR) get identified, again with strictly fewer isometries relating the triples. We continue in this fashion to verify that (A, B, C) and (P, Q, R) get identified.

□

Below we list the definitions and facts we need, cited from [3].

Definition 9. *Let $f : X \longrightarrow Y$ be continuous. f is said to be proper if for every space Z the map $(x, z) \mapsto (f(x), z)$ from $X \times Z \longrightarrow Y \times Z$ is closed.*

Proposition 4.2.12. *Let $f : X \longrightarrow Y$ be continuous and injective. Then f is proper if and only if f is closed.*

Definition 10. *Let the group G act on the space X. G is said to act properly on X if the map $(g, x) \mapsto (gx, x)$ from $G \times X \longrightarrow X \times X$ is a proper map.*

Definition 11. *A group G acts freely on the space X if $gx = x \Rightarrow g = e$ whenever $g \in G, x \in X$.*

From this we can see that when G acts freely on X the map $(g, x) \mapsto (gx, x)$ is automatically injective. Indeed, $(gx, x) = (hy, y)$ implies $x = y$ whereupon $gx = hx$ implies $h^{-1}g = e$ so that $g = h$. Therefore in the case of a free group action, one need only show that the map $(g, x) \mapsto (gx, x)$ is closed in order to conclude that G acts properly on X.

The group $\mathbb{H}$ acts on the space of vertices $\widetilde{T} = \mathbb{H}^3 - \Sigma$ by the rule

$$G \circ (P_1, P_2, P_3) = (GP_1, GP_2, GP_3) \text{ for } G \in \mathbb{H}, (P_1, P_2, P_3) \in \widetilde{T}.$$

The operation is just to left-translate each vertex by G, and so is a free group action. The group operation induces the map

$$\mu : \mathbb{H} \times \widetilde{T} \longrightarrow \widetilde{T} \times \widetilde{T}$$

defined by

$$(G, P) \mapsto (GP, P).$$

We claim that μ is a closed map. Suppose we are given a sequence

$$\{(G_iP_i, P_i)\}_{i=1}^{\infty} \subset \mu(A)$$

that converges in $\widetilde{T} \times \widetilde{T}$, where $A \subset \mathbb{H} \times \widetilde{T}$ is a closed set. Now

$$(G_iP_i, P_i) \longrightarrow (R, P) \in \widetilde{T} \times \widetilde{T}.$$

We will show that in fact (R, P) is also in the image of A under μ. Thus $\mu(A)$ is closed.

The following notation is cumbersome but necessary. Let

$$P_i = (P_{i1}, P_{i2}, P_{i3}) = (x_{i1}, y_{i1}, t_{i1}; x_{i2}, y_{i2}, t_{i2}; x_{i3}, y_{i3}, t_{i3})$$

and let

$$R = (R_1, R_2, R_3) = (u_1, v_1, s_1; u_2, v_2, s_2; u_3, v_3, s_3)$$

$$P = (P_1, P_2, P_3) = (x_1, y_1, t_1; x_2, y_2, t_2; x_3, y_3, t_3).$$

Now since $(G_iP_i, P_i) \longrightarrow (R, P)$ we have in particular that $P_i \longrightarrow P$ and therefore that $P_{i1} \longrightarrow P_1$, $P_{i2} \longrightarrow P_2$, $P_{i3} \longrightarrow P_3$. From these three we see that

$$x_{i1} \longrightarrow x_1, y_{i1} \longrightarrow y_1, t_{i1} \longrightarrow t_1$$

$$x_{i2} \longrightarrow x_2, y_{i2} \longrightarrow y_2, t_{i2} \longrightarrow t_2$$

$$x_{i3} \longrightarrow x_3, y_{i3} \longrightarrow y_3, t_{i3} \longrightarrow t_3.$$

Next we introduce notation for the G_i's

$$G_i = (x_{G_i}, y_{G_i}, t_{G_i})$$

so that we can write out

$$G_iP_i = (x_{G_i}, y_{G_i}, t_{G_i}) \circ (x_{i1}, y_{i1}, t_{i1}; x_{i2}, y_{i2}, t_{i2}; x_{i3}, y_{i3}, t_{i3})$$

$$= (x_{G_i} + x_{i1}, y_{G_i} + y_{i1}, t_{G_i} + t_{i1} - 2(x_{G_i}y_{i1} - y_{G_i}x_{i1});$$

$$x_{G_i} + x_{i2}, y_{G_i} + y_{i2}, t_{G_i} + t_{i2} - 2(x_{G_i}y_{i2} - y_{G_i}x_{i2});$$

$$x_{G_i} + x_{i3}, y_{G_i} + y_{i3}, t_{G_i} + t_{i3} - 2(x_{G_i}y_{i3} - y_{G_i}x_{i3})).$$

And now from the fact that

$$G_iP_i \longrightarrow R$$

and looking only at the first triple of co-ordinates in this limit, we can deduce that

$$x_{G_i} + x_{i1} \longrightarrow u_1$$

$$y_{G_i} + y_{i1} \longrightarrow v_1$$

$$t_{G_i} + t_{i1} - 2(x_{G_i}y_{i1} - y_{G_i}x_{i1}) \longrightarrow s_1.$$

We then immediately have

$$x_{G_i} \longrightarrow u_1 - x_1$$

$$y_{G_i} \longrightarrow v_1 - y_1$$

and moreover

$$s_1 = \lim_{i\to\infty} t_{G_i} + t_{i1} - 2(x_{G_i}y_{i1} - y_{G_i}x_{i1})$$

$$= \lim_{i\to\infty} t_{G_i} + t_1 - 2((u_1 - x_1)y_1 - (v_1 - y_1)x_1)$$

$$= \lim_{i\to\infty} t_{G_i} + t_1 - 2(u_1y_1 - v_1x_1)$$

or equivalently

$$\lim_{i\to\infty} t_{G_i} = s_1 - t_1 + 2(u_1y_1 - v_1x_1)$$

$$= s_1 - t_1 - 2(u_1(-y_1) - v_1(-x_1)).$$

Thus

$$\lim_{i\to\infty} G_i = (u_1 - x_1, v_1 - y_1, s_1 - t_1 - 2(u_1(-y_1) - v_1(-x_1)))$$

$$= R_1P_1^{-1}.$$

Likewise looking at the second and third triples in the limit $G_iP_i \longrightarrow R$ we obtain

$$\lim_{i\to\infty} G_i = R_2P_2^{-1}$$

$$\lim_{i\to\infty} G_i = R_3P_3^{-1}.$$

Therefore we can define

$$G := \lim_{i\to\infty} G_i = R_1P_1^{-1} = R_2P_2^{-1} = R_3P_3^{-1}.$$

But now $(G_i, P_i) \longrightarrow (G, P)$, so that $(G, P) \in A$ because A is closed. Then

$$(G,P) \mapsto (GP,P) = (G(P_1,P_2,P_3);P) = (R_1P_1^{-1}P_1, R_2P_2^{-1}P_2, R_3P_3^{-1}P_3;P)$$
$$= (R,P)$$

so that $(R, P) \in \mu(A)$.

In summary, if for closed A a sequence $\{(G_iP_i, P_i)\}_{i=1}^{\infty} \subset \mu(A)$ converges to a limit in $\widetilde{T} \times \widetilde{T}$, then the limit is actually in $\mu(A)$. Hence $\mu(A)$ contains all its limit points and is therefore closed. We thus have

Theorem 4.2.13. *The group $\mathbb{H}$ acts properly on the space $\mathbb{H} \times \mathbb{H} \times \mathbb{H} - \Sigma$.*

The second group action consists of the rotations $SO(2)$ acting on the quotient space

$$\mathbb{H}\backslash(\mathbb{H}^3 - \Sigma)$$

obtained from the first group action. We do not need to work directly with this group action, because we can use another proposition from [3].

Proposition 4.2.14. *Let K be a compact group operating continuously on a Hausdorff space X. Then K operates properly on X.*

We thus have

Corollary 4.2.15. *The group $SO(2)$ acts properly on the space $\mathbb{H}\backslash(\mathbb{H}^3 - \Sigma)$.*

The composition of these two group actions on $\mathbb{H}^3 - \Sigma$, first by left-translations $\mathbb{H}$ and then by rotations $SO(2)$, is the same as the action of the group $\mathbb{H} \rtimes SO(2)$ of isometries generated by the left-translations and the rotations. Since the composition of proper group actions is again proper (see [3]), we conclude:

Theorem 4.2.16. *The action of the group $\mathbb{H} \rtimes SO(2)$ on the space of vertices $\widetilde{T} = \mathbb{H}^3 - \Sigma$ is a proper group action.*

At this point we would like to use the characterization of smooth principal bundles, which we cite from [6]:

Theorem 4.2.17. *Let P be a smooth manifold, H a Lie group, and $\mu : P \times H \longrightarrow P$ a smooth, free, proper right action. Then*

(i) P/H with the quotient topology is a topological manifold ($\dim P/H = \dim P - \dim H$),

(ii) P/H has a unique smooth structure for which the canonical projection $P \longrightarrow P/H$ is a submersion,

(iii) $\xi = (P, \pi, M, H)$ is a smooth principal right H bundle.

In the notation of the theorem $M = P/H$ and $\pi : P \to M$ is the canonical projection. There is one technical issue. We have worked with the action of $\mathbb{H} \rtimes SO(2)$ on $\mathbb{H}^3 - \Sigma$ as a left group action. However in general if a group G acts on a set X by a *left* group action $\mu_L : G \times X \to X$, we can define a *right* group action $\mu_R : X \times G \to X$ by the rule

$$\mu_R(x, g) := \mu_L(g^{-1}, x)$$

for $x \in X$ and $g \in G$. The two group actions have identical orbits so that the quotient maps $X \to G\backslash X$ and $X \to X/G$ are identical maps, and it is straightforward to check that when μ_L is smooth, free, and proper so is μ_R. We can therefore use theorems 4.2.16 and 4.2.17 to conclude:

Theorem 4.2.18. *The map $p : \mathbb{H}^3 - \Sigma \longrightarrow T$ is a smooth principal right $\mathbb{H} \rtimes SO(2)$ bundle.*

Note that we have already directly demonstrated some facts we could conclude from theorem 4.2.17, for instance that T is a 5-dimensional smooth manifold.

4.2.5 A cross-section for the bundle $p : \mathbb{H}^3 - \Sigma \longrightarrow T$

We thus have the principal fibre bundle

$$p : \mathbb{H}^3 - \Sigma \longrightarrow T \subset \mathbb{R}^6$$

with structure group $\mathbb{H} \rtimes SO(2)$. A 6-tuple

$$(a, \zeta_A), (b, \zeta_B), (c, \zeta_C)$$

in the parameter space corresponds to the equivalence class $p^{-1}(a, \zeta_A, b, \zeta_B, c, \zeta_C)$ consisting of triangles congruent to one another by an isometry in $\mathbb{H} \rtimes SO(2)$ and having the given parameters. We now describe a way to choose a well-defined representative from each such equivalence class in a continuous fashion. That is, we define a continuous map

$$s : T \longrightarrow \mathbb{H}^3 - \Sigma$$

such that $p \circ s = \mathrm{id}_T$. For $q \in T$ the pre-image $p^{-1}(q)$ is called the fibre over T. Note that the condition $p \circ s = \mathrm{id}_T$ really just says that for each $q \in T$ we have $s(q)$ in the fibre over T. Such a map is called a cross-section for the bundle $p : \mathbb{H}^3 - \Sigma \longrightarrow T$.

The construction of the map s is straightforward. From the equivalence class $p^{-1}(a, \zeta_A, b, \zeta_B, c, \zeta_C)$ of triangles congruent by an isometry in $\mathbb{H} \rtimes SO(2)$ we choose the unique representative with first vertex A at the origin and second vertex B in the yt-plane with $y > 0$. Recall that we call triangles of this form standard triangles and that every triangle is congruent by an isometry in $\mathbb{H} \rtimes SO(2)$ to a unique standard triangle.

To see continuity of the map s we can compute the coordinates of the map directly. In a Euclidean triangle with side lengths

$$\tilde{a} = \frac{a}{\sigma(\zeta_A)} \qquad \tilde{b} = \frac{b}{\sigma(\zeta_B)} \qquad \tilde{c} = \frac{c}{\sigma(\zeta_C)}$$

let $\alpha \in [0, \pi]$ be the angle opposite the side $\tilde{a}$

$$\cos \alpha = \frac{\tilde{b}^2 + \tilde{c}^2 - \tilde{a}^2}{2\tilde{b}\tilde{c}}.$$

If the standard triangle ABC is to project to a Euclidean triangle with side-lengths $\tilde{a}$, $\tilde{b}$, $\tilde{c}$ then we must have projections z_A, z_B, z_C with coordinates

$$z_A = (0,0)$$

$$z_B = (0,\tilde{c})$$

$$z_C = (\tilde{b}\sin\alpha, \tilde{b}\cos\alpha).$$

Now we just choose a height t_B so that the geodesic from A to B has the correct value ζ_C

$$t_B = \tilde{c}^2\mu(\zeta_C)$$

and a height t_C so that the geodesic from C to A has the correct value ζ_B

$$t_C = -\tilde{b}^2\mu(\zeta_B)$$

(note we have the $-$ sign because the geodesic goes from C to A). The map s is therefore given in coordinates by

$$(a,\zeta_A,b,\zeta_B,c,\zeta_C) \mapsto \left(0,0,0;0,\tilde{c},\tilde{c}^2\mu(\zeta_C);\tilde{b}\sin\alpha,\tilde{b}\cos\alpha,-\tilde{b}^2\mu(\zeta_B)\right)$$

and the continuity of the map is clear.

4.3 Topology of the moduli space

We start with the space of vertices

$$\mathbb{H}^3 - \Sigma$$

and successively apply group actions by $\mathbb{H}$

$$\mathbb{H}\backslash\left(\mathbb{H}^3 - \Sigma\right)$$

and by $SO(2)$

$$SO(2)\backslash\left(\mathbb{H}\backslash\left(\mathbb{H}^3 - \Sigma\right)\right)$$

to obtain the moduli space T of Heisenberg triangles up to isometry in $\mathbb{H} \rtimes SO(2)$.

We can additionally let the group $R = \{1, r\}$ act on T where r is the standard involution to get

$$R\backslash T$$

the moduli space of Heisenberg triangles up to isometry in the full group of isometries $\mathbb{H} \rtimes O(2)$.

We proceed with results towards computing the homotopy and homology of the spaces T and $R\backslash T$.

Proposition 4.3.1. *The space $\mathbb{H}^3 - \Sigma$ is homotopy equivalent to the space*

$$\mathbb{C}^3 - \Sigma'$$

where

$$\Sigma' = \left\{(z_1, z_2, z_3) \in \mathbb{C}^3 | z_i = z_j \text{ for some } i \neq j\right\}.$$

Proof. The map

$$(x_A, y_A, t_A; x_B, y_B, t_B; x_C, y_C, t_C) \mapsto (x_A + iy_A, x_B + iy_B, x_C + iy_C; t_A, t_B, t_C)$$

gives a homeomorphism $\mathbb{H}^3 - \Sigma \cong (\mathbb{C}^3 - \Sigma') \times \mathbb{R}^3$, and $(\mathbb{C}^3 - \Sigma') \times \mathbb{R}^3$ is homotopy equivalent to $\mathbb{C}^3 - \Sigma'$ since each $\mathbb{R}$ factor is contractible. □

Note that this means that the fundamental group of the space of vertices $\pi_1(\mathbb{H}^3 - \Sigma) = \pi_1(\mathbb{C}^3 - \Sigma')$ is the *pure braid group* $\mathbf{P}_3$ on 3 strands. See, for example, [2].

Now on the space $\mathbb{H}^3 - \Sigma$ we have the action of the group $\mathbb{H}$ of left-translations and on the space $\mathbb{C}^3 - \Sigma'$ we have the action of the group $\mathbb{C}$ of translations. If we look merely at the x- and y-coordinates of a triple $(A, B, C) \in \mathbb{H}^3 - \Sigma$ we observe that the action of a group element $(z, t) \in \mathbb{H}$ is merely to translate the projections z_A, z_B, z_C

$$z_A \mapsto z + z_A \qquad z_B \mapsto z + z_B \qquad z_C \mapsto z + z_C.$$

Observe that the effect of the homotopy equivalence $\mathbb{H}^3 - \Sigma \simeq \mathbb{C}^3 - \Sigma'$ is precisely to ignore the t-coordinate and thereby identify triples (A, B, C) and (P, Q, R) for which the projections (z_A, z_B, z_C) and (z_P, z_Q, z_R) are equal. We thus obtain a commutative diagram

$$\begin{array}{ccc} \mathbb{H}^3 - \Sigma & \xrightarrow{\simeq} & \mathbb{C}^3 - \Sigma' \\ \downarrow & & \downarrow \\ \mathbb{H}\backslash(\mathbb{H}^3 - \Sigma) & \xrightarrow{\simeq} & \mathbb{C}\backslash(\mathbb{C}^3 - \Sigma') \end{array}$$

where the homotopy equivalence across the bottom of the diagram comes once again from ignoring the t-coordinates.

Proposition 4.3.2. *The space $\mathbb{C}\backslash(\mathbb{C}^3 - \Sigma')$ is homeomorphic to the space*

$$\mathbb{C}^2 - L_1 \cup L_2 \cup L_3$$

where

$$L_1 = \left\{(w, z) \in \mathbb{C}^2 | w = 0\right\}$$

$$L_2 = \left\{(w, z) \in \mathbb{C}^2 | z = 0\right\}$$

$$L_3 = \left\{(w, z) \in \mathbb{C}^2 | w = z\right\}.$$

Proof. Under the group action of $\mathbb{C}$ on $\mathbb{C}^3 - \Sigma'$ acting by the rule

$$z \circ (z_1, z_2, z_3) = (z + z_1, z + z_2, z + z_3)$$

the triples (z_1, z_2, z_3) and (w_1, w_2, w_3) are identified if and only if

$$z_1 - w_1 = z_2 - w_2 = z_3 - w_3.$$

The map

$$\Phi : \mathbb{C}^3 - \Sigma' \longrightarrow \mathbb{C}^2 - L_1 \cup L_2 \cup L_3$$

$$(z_1, z_2, z_3) \mapsto (z_1 - z_3, z_2 - z_3)$$

then induces a homeomorphism from $\mathbb{C}\backslash(\mathbb{C}^3 - \Sigma')$ to $\mathbb{C}^2 - L_1 \cup L_2 \cup L_3$.

Indeed $(z_1 - z_3, z_2 - z_3) = (w_1 - w_3, w_2 - w_3)$ if and only if $z_1 - w_1 = z_2 - w_2 = z_3 - w_3$ so that Φ maps two triples (z_1, z_2, z_3), (w_1, w_2, w_3) to the same image if and only if they are in the same orbit under the $\mathbb{C}$-action. Moreover (z, w) with $z \neq w$, $z \neq 0$, and $w \neq 0$ is the image of the point $(z, w, 0)$ so that Φ is surjective.

□

This shows $\mathbb{H}\backslash(\mathbb{H}^3 - \Sigma) \simeq \mathbb{C}^2 - L_1 \cup L_2 \cup L_3$.

Proposition 4.3.3. *The space $\mathbb{C}^2 - L_1 \cup L_2 \cup L_3$ is homotopy equivalent to*

$$S^3 - L_1 \cup L_2 \cup L_3$$

where S^3 is the 3-sphere viewed as a subspace of $\mathbb{C}^2$.

Proof. Notice that the origin $O(0,0)$ is not in $\mathbb{C}^2 - L_1 \cup L_2 \cup L_3$. Now $\mathbb{C}^2 - O$ can be identified with $\mathbb{R}^4 - O$ which contracts onto the 3-sphere S^3. This contraction restricted to $\mathbb{C}^2 - L_1 \cup L_2 \cup L_3$ then gives a homotopy equivalence

$$\mathbb{C}^2 - L_1 \cup L_2 \cup L_3 \simeq S^3 - L_1 \cup L_2 \cup L_3$$

where on the right-hand side L_i means $L_i \cap S^3$.

□

Now we add the action of the rotations $SO(2)$ on $\mathbb{H}\backslash(\mathbb{H}^3 - \Sigma)$. We identify $SO(2)$ with the unit circle $S^1 \subset \mathbb{C}$ to obtain another commutative diagram

$$\begin{array}{ccc} \mathbb{H}\backslash(\mathbb{H}^3 - \Sigma) & \xrightarrow{\simeq} & \mathbb{C}\backslash(\mathbb{C}^3 - \Sigma') \\ \downarrow & & \downarrow \\ SO(2)\backslash(\mathbb{H}\backslash(\mathbb{H}^3 - \Sigma)) & \xrightarrow{\simeq} & S^1\backslash(\mathbb{C}\backslash(\mathbb{C}^3 - \Sigma')) \end{array}$$

where the action of S^1 on $\mathbb{C}\backslash(\mathbb{C}^3 - \Sigma')$ is induced by the action

$$z \circ (z_1, z_2, z_3) = (zz_1, zz_2, zz_3)$$

of S^1 on $\mathbb{C}^3 - \Sigma'$. The action of $SO(2)$ on $\mathbb{H}\backslash(\mathbb{H}^3 - \Sigma)$ is (as we have seen) likewise induced by the action

$$z \circ (z_1, t_1; z_2, t_2; z_3, t_3) = (zz_1, t_1; zz_2, t_2; zz_3, t_3)$$

of $SO(2)$ on $\mathbb{H}^3 - \Sigma$ so it is apparent that the diagram is commutative.

We continue with the above commutative diagram and extend it to the right with

$$\begin{array}{ccc}
\mathbb{C}\backslash(\mathbb{C}^3 - \Sigma') & \xrightarrow{\cong} & \mathbb{C}^2 - L_1 \cup L_2 \cup L_3 \\
\downarrow & & \downarrow \\
S^1\backslash\big(\mathbb{C}\backslash(\mathbb{C}^3 - \Sigma')\big) & \xrightarrow{\cong} & S^1\backslash\big(\mathbb{C}^2 - L_1 \cup L_2 \cup L_3\big)
\end{array}$$

using the homeomorphism from proposition 4.3.2 and yet further to the right with

$$\begin{array}{ccc}
\mathbb{C}^2 - L_1 \cup L_2 \cup L_3 & \xrightarrow{\simeq} & S^3 - L_1 \cup L_2 \cup L_3 \\
\downarrow & & \downarrow \\
S^1\backslash\big(\mathbb{C}^2 - L_1 \cup L_2 \cup L_3\big) & \xrightarrow{\simeq} & S^1\backslash(S^3 - L_1 \cup L_2 \cup L_3)
\end{array}$$

using the homotopy equivalence from proposition 4.3.3. Following along the bottom of the diagrams the conclusion is

Proposition 4.3.4. *The moduli space of Heisenberg triangles*

$$SO(2)\backslash\big(\mathbb{H}\backslash(\mathbb{H}^3 - \Sigma)\big)$$

is homotopy equivalent to the space

$$S^1 \backslash (S^3 - L_1 \cup L_2 \cup L_3)$$

where

$$L_1 = \{(w,z) \in \mathbb{C}^2 | w = 0\} \cap S^3$$

$$L_2 = \{(w,z) \in \mathbb{C}^2 | z = 0\} \cap S^3$$

$$L_3 = \{(w,z) \in \mathbb{C}^2 | w = z\} \cap S^3.$$

We view S^3 as a subspace of $\mathbb{C}^2$ and the action of S^1 on $S^3 - L_1 \cup L_2 \cup L_3$ is given by

$$\lambda \circ (w,z) = (\lambda w, \lambda z)$$

for $\lambda \in S^1 \subset \mathbb{C}$ and $(w,z) \in S^3 - L_1 \cup L_2 \cup L_3 \subset \mathbb{C}^2$.

4.3.1 The Hopf map

Next we use a famous construction called the Hopf map to produce a homeomorphism $S^1 \backslash S^3 \longrightarrow S^2$.

Definition 12. *For*

$$S^3 = \{(w,z) \in \mathbb{C}^2 : |w|^2 + |z|^2 = 1\}$$

$$S^2 = \{(x,y,z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$$

we define the Hopf map $\rho : S^3 \longrightarrow S^2$ by

$$\rho(w,z) = \big(|w|^2 - |z|^2, 2\,\mathrm{Re}(w\bar{z}), 2\,\mathrm{Im}(w\bar{z})\big).$$

We check that ρ indeed maps S^3 to S^2. If $|w|^2+|z|^2 = 1$ then $|w|^4+2|w|^2|z|^2+|z|^4 =$

1. Then

$$\left[|w|^2 - |z|^2\right]^2 + [2\,\mathrm{Re}(w\overline{z})]^2 + [2\mathrm{Im}(w\overline{z})]^2$$

$$= |w|^4 - 2|w|^2|z|^2 + |z|^4 + 4|w\overline{z}|^2$$

$$= |w|^4 + 2|w|^2|z|^2 + |z|^4 = 1.$$

Observe next that for any $\lambda \in S^1 \subset \mathbb{C}$ we have

$$\rho(\lambda w, \lambda z) = \rho(w, z).$$

Moreover we have the following:

Proposition 4.3.5. *Suppose that for (w, z) and (ξ, η) in S^3 we have $\rho(w, z) = \rho(\xi, \eta)$. Then*

$$(w, z) = (\lambda\xi, \lambda\eta)$$

for some $\lambda \in S^1 \subset \mathbb{C}$.

Proof. From the second and third coordinates of the map ρ we obtain the equalities

$$2\,\mathrm{Re}(w\overline{z}) = 2\,\mathrm{Re}(\xi\overline{\eta})$$

$$2\mathrm{Im}(w\overline{z}) = 2\mathrm{Im}(\xi\overline{\eta}).$$

It follows that $w\overline{z} = \xi\overline{\eta}$.

From here it is straightforward to account for all possibilities where one of w, z, ξ, η is 0. For example $z = 0$ forces $|w| = 1 = |\xi|$ and $\eta = 0$ so that the pairs $(w, 0)$ and $(\xi, 0)$ are indeed related in the desired manner. We thus address the case where w, z, ξ, η are all non-zero.

Then

$$\frac{w}{\xi} = \frac{\overline{\eta}}{\overline{z}} = \lambda \neq 0$$

or

$$w = \lambda\xi \qquad\qquad \overline{\eta} = \lambda\overline{z}.$$

Our goal is to show that $|\lambda| = 1$. Now from these we have the four equations

$$|w|^2 + |z|^2 = 1 \qquad |\xi|^2 + |\eta|^2 = 1$$

$$\frac{1}{|\lambda|^2}|w|^2 + |\lambda|^2|z|^2 = 1 \qquad |\lambda|^2|\xi|^2 + \frac{1}{|\lambda|^2}|\eta|^2 = 1.$$

Multiplying the second equation on the left by $|\lambda|^2$ and subtracting the first equation we obtain

$$\big(|\lambda|^4 - 1\big)|z|^2 = |\lambda|^2 - 1$$

and similarly on the right we obtain

$$\big(|\lambda|^4 - 1\big)|\xi|^2 = |\lambda|^2 - 1.$$

If $|\lambda| = 1$ we are done. Otherwise we can cancel the quantity $|\lambda|^2 - 1$ and solve to obtain

$$|z|^2 = \frac{1}{|\lambda|^2 + 1} \qquad |\xi|^2 = \frac{1}{|\lambda|^2 + 1}$$

whereupon

$$|w|^2 = \frac{|\lambda|^2}{|\lambda|^2 + 1} \qquad |\eta|^2 = \frac{|\lambda|^2}{|\lambda|^2 + 1}.$$

But now from the first coordinate of the map ρ we use $|w|^2 - |z|^2 = |\xi|^2 - |\eta|^2$ or $|w|^2 + |\eta|^2 = |\xi|^2 + |z|^2$ to get

$$\frac{2|\lambda|^2}{|\lambda|^2 + 1} = \frac{2}{|\lambda|^2 + 1}$$

contradicting the assumption that $|\lambda|^2 \neq 1$. We conclude that $|\lambda| = 1$. Finally we already have $w = \lambda\xi$ and then

$$\overline{\eta} = \lambda\overline{z} \Rightarrow \eta = \overline{\lambda}z$$

$$\Rightarrow z = \frac{1}{\overline{\lambda}}\eta = \lambda\eta$$

because λ has modulus 1.

□

Proposition 4.3.6. *The Hopf map is surjective.*

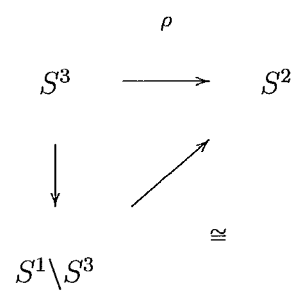
Proof. This can be checked directly. Suppose $(x, y, t) \in S^2$ with $x \neq 1$. We want to find $(w, z) \in S^3$ such that $\rho(w, z) = (x, y, t)$ and by taking advantage of the symmetry $\rho(w, z) = \rho(\lambda w, \lambda z)$ for $\lambda \in S^1$ we may assume z is real, in fact strictly positive. With this restriction it is routine to compute the solution

$$(w, z) = \left(\frac{y}{\sqrt{2}\sqrt{1-x}} + \frac{t}{\sqrt{2}\sqrt{1-x}} i, \sqrt{\frac{1-x}{2}} \right)$$

and check that $\rho(w, z) = (x, y, t)$. For the case $x = 1$ we can take $(w, z) = \left(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} i, 0 \right)$.

$\square$

In fact the map $(x, y, t) \mapsto (w, z)$ constructed in proposition 4.3.6 is continuous and provides a continuous inverse for the Hopf map factored through the S^1 -action on S^3



The Hopf map thus induces a homeomorphism $S^1 \backslash S^3 \cong S^2$.

4.3.2 A homotopy equivalence for the moduli space

This homeomorphism $S^1 \backslash S^3 \cong S^2$ in turn induces a homeomorphism

$$S^1 \backslash (S^3 - L_1 \cup L_2 \cup L_3) \cong S^2 - \rho(L_1) \cup \rho(L_2) \cup \rho(L_3).$$

Now one readily checks that

$$\rho(L_1) = \{(-1,0,0)\}$$

$$\rho(L_2) = \{(1,0,0)\}$$

$$\rho(L_3) = \{(0,1,0)\}\,.$$

We have

Proposition 4.3.7. *The spaces*

$$T = SO(2)\backslash\big(\mathbb{H}\backslash(\mathbb{H}^3 - \Sigma)\big)$$

and

$$X = S^2 - \{(-1,0,0),(1,0,0),(0,1,0)\}$$

are homotopy equivalent.

Thus T has the homotopy type of a thrice-punctured sphere $S^2 - \{p_1, p_2, p_3\}$. Citing [5] we can conclude that T is *aspherical*, that is, T has trivial higher homotopy. For the fundamental group of T, observe that a thrice-punctured sphere is homotopy equivalent to the wedge sum of two circles $S^1 \vee S^1$ and therefore has fundamental group free on two generators:

- $\pi_1(T) = \langle a, b\rangle$
- $\pi_i(T) = 0,\ i \geq 2.$

We account last for the action of the group $R = \{1, r\}$ generated by the involution. On a triple $(A, B, C) \in \mathbb{H}^3 - \Sigma$ the involution r has the effect

$$\big(z_A, t_A; z_B, t_B; z_C, t_C\big) \mapsto \big(\overline{z}_A, -t_A; \overline{z}_B, -t_B; \overline{z}_C, -t_C\big).$$

When we contract away the t-coordinates the action of R on $\mathbb{C}^3 - \Sigma'$ becomes merely that of component-wise conjugation $(z_1, z_2, z_3) \mapsto (\overline{z}_1, \overline{z}_2, \overline{z}_3)$. Tracing through the commutative diagrams we check that the action of R behaves as component-wise conjugation and gives the homotopy equivalence

$$\begin{array}{ccc} T & \longrightarrow & S^1\backslash(S^3 - L_1 \cup L_2 \cup L_3) \\ \downarrow & & \downarrow \\ R\backslash T & \longrightarrow & R\backslash\big(S^1\backslash(S^3 - L_1 \cup L_2 \cup L_3)\big) \end{array}$$

However when we apply the Hopf map $\rho : S^3 \longrightarrow S^2$ we find that the R-action changes its behaviour. The R-action identifies conjugates on S^3 but

$$\rho(\overline{w}, \overline{z}) = \big(|w|^2 - |z|^2, 2\,\mathrm{Re}(\overline{\overline{w}}z), 2\mathrm{Im}(\overline{\overline{w}}z)\big)$$

$$= \big(|w|^2 - |z|^2, 2\,\mathrm{Re}(w\overline{z}), -2\mathrm{Im}(w\overline{z})\big).$$

On S^2 the R-action therefore becomes

$$r \circ (x, y, z) = (x, y, -z).$$

The R-action thus identifies the upper and lower hemispheres of S^2, turning S^2 into a disk with boundary where the equator was. On $X = S^2 - \{(-1,0,0), (1,0,0), (0,1,0)\}$ the three removed points are all on the equator, so the R-action creates a disk with three boundary points removed. Such a disk remains contractible, so we can conclude that $R\backslash T$ is contractible.

4.4 Unordered Heisenberg triangles

4.4.1 The S_3-action

At last we let the symmetric group S_3 act on $\mathbb{H}^3 - \Sigma$ in order to remove the ordering on vertices (A, B, C). We produce triangles $\triangle ABC$.

Proposition 4.4.1. *The S_3 action on vertices induces a well-defined action on $\mathbb{H} \rtimes O(2)\backslash(\mathbb{H}^3 - \Sigma)$.*

Proof. Indeed given $\sigma \in S_3$ suppose two triples of vertices are related by an isometry $g \in \mathbb{H} \rtimes O(2)$

$$(P_1, P_2, P_3) \sim (gP_1, gP_2, gP_3).$$

Then the permuted vertices are related by the same isometry

$$(P_{\sigma(1)}, P_{\sigma(2)}, P_{\sigma(3)}) \sim (gP_{\sigma(1)}, gP_{\sigma(2)}, gP_{\sigma(3)}).$$

□

In the action we have just constructed S_3 acts on $\mathbb{H} \rtimes O(2)\backslash(\mathbb{H}^3 - \Sigma)$ by the map

$$\sigma \circ [(P_1, P_2, P_3)] = \left[(P_{\sigma(1)}, P_{\sigma(2)}, P_{\sigma(2)})\right].$$

Alternatively we could let S_3 act directly on $\mathbb{H}^3 - \Sigma$ to produce (unordered) sets of three distinct vertices. Now we let $\mathbb{H} \rtimes O(2)$ act on $S_3\backslash(\mathbb{H}^3 - \Sigma)$ by the map

$$g \circ \{P_1, P_2, P_3\} = \{gP_1, gP_2, gP_3\}.$$

But if the triple (A, B, C) can be made to coincide with another triple (X, Y, Z) first by an isometry and then by a permutation, so that $\sigma \circ g \circ (A, B, C) = (X, Y, Z)$, then $\sigma \circ (A, B, C)$ and (X, Y, Z) have the same triangle parameters up to sign on the ζ-values

$$a = x, b = y, c = z, \zeta_A = \pm\zeta_X, \zeta_B = \pm\zeta_Y, \zeta_C = \pm\zeta_Z$$

and therefore $\sigma \circ (A, B, C)$ and (X, Y, Z) coincide by an isometry. Conversely if the triple (A, B, C) can be made to coincide with another triple (X, Y, Z) first by a permutation and then by an isometry, so that $g \circ \sigma \circ (A, B, C) = (X, Y, Z)$, then $g \circ (A, B, C)$ and (X, Y, Z) consist of the same vertices, but not necessarily in the same order, so that $g \circ (A, B, C)$ and (X, Y, Z) coincide by a permutation. Thus we have shown

Proposition 4.4.2. *The successive group actions*

$$S_3\backslash \left(\mathbb{H} \rtimes O(2)\backslash(\mathbb{H}^3 - \Sigma)\right)$$

$$\mathbb{H} \rtimes O(2)\backslash \left(S_3\backslash(\mathbb{H}^3 - \Sigma)\right)$$

identify precisely the same triples $(A, B, C) \in \mathbb{H}^3 - \Sigma$.

Therefore we need not worry about the order in which the group actions are applied.

4.4.2 The moduli space for unordered Heisenberg triangles

We now give a definition for families of unordered Heisenberg triangles. Our definitions for families of ordered and unordered Heisenberg triangles bear the same relation as the definitions for families of ordered and unordered Euclidean triangles, with one significant exception. Since we allow geodesically colinear vertices for Heisenberg triangles, we require an unordered triple of sections to specify vertices.

Definition 13. *For a topological space S, a family of unordered Heisenberg triangles over S is a subspace $X \subset S \times \mathbb{H}$ such that*

(1) projection onto the first coordinate is a (continuous, proper) fibre bundle projection

(2) each fibre $X_s = p^{-1}(s)$ is (after projection onto the second coordinate) a Heisenberg triangle $\gamma_A \cup \gamma_B \cup \gamma_C \subset \mathbb{H}$

(3) there is an unordered triple of sections $V, V', V'' : S \to X$ which specify the vertices of each fibre X_s.

Morphisms between families $X \to S$ and $X' \to S'$ are pairs of continuous maps making the diagram commute

$$\begin{array}{ccc} X & \xrightarrow{f} & X' \\ \downarrow & & \downarrow \\ S & \xrightarrow{g} & S' \end{array}$$

and such that for each $s \in S$ the map f restricts to an isometry in Isom($\mathbb{H}$) from X_s to $X'_{g(s)}$ respecting vertices in the sense that the sets $\{f(V(s)), f(V'(s)), f(V''(s))\}$ and $\{V(g(s)), V'(g(s)), V''(g(s))\}$ are identical.

Any family $X \to S$ determines a canonical map $S \to S_3 \backslash (\mathrm{Isom}(\mathbb{H}) \backslash \mathbb{H}^3 - \Sigma)$ taking $s \in S$ to the isometry class of the fibre X_s. But just as in the Euclidean case, the existence of non-trivial self isometries for equilateral triangles[1] means the space $S_3 \backslash (\mathrm{Isom}(\mathbb{H}) \backslash \mathbb{H}^3 - \Sigma)$ cannot have a universal family over it. Therefore the quotient $S_3 \backslash (\mathrm{Isom}(\mathbb{H}) \backslash \mathbb{H}^3 - \Sigma)$ is a coarse moduli space for unordered Heisenberg triangles.

4.4.3 Topology of the moduli space

We proceed to trace the effect of the S_3-action down through our commutative diagrams of homotopy equivalences and homeomorphisms.

First the S_3-action has the effect on $\mathbb{H}^3 - \Sigma$ of identifying sets of six distinct triples

$$(A, B, C), (B, C, A), (C, A, B), (A, C, B), (C, B, A), (B, A, C).$$

After retracting the t-coordinates to get the homotopy equivalance $\mathbb{H}^3 - \Sigma \cong \mathbb{C}^3 - \Sigma'$ the S_3-action identifies the distinct triples

$$(z_A, z_B, z_C), (z_B, z_C, z_A), (z_C, z_A, z_B),$$

$$(z_A, z_C, z_B), (z_C, z_B, z_A), (z_B, z_A, z_C).$$

We pause to point out that this means $\pi_1(S_3 \backslash \mathbb{H}^3 - \Sigma) = \pi_1(S_3 \backslash \mathbb{C}^3 - \Sigma')$ is the *braid group* B_3 on 3 strands. Again see [2].

[1]That is, a triangle having parameters $(a, \zeta, a, \zeta, a, \zeta)$. We show in section 4.7.1 that such triangles exist.

After the $\mathbb{C}$-action on $\mathbb{C}^3 - \Sigma'$ the six equivalence classes

$$[(z_A, z_B, z_C)], [(z_B, z_C, z_A)], [(z_C, z_A, z_B)],$$
$$[(z_A, z_C, z_B)], [(z_C, z_B, z_A)], [(z_B, z_A, z_C)]$$

are identified. One checks that in fact these are precisely the induced identifications because if (A, B, C) and (X, Y, Z) are related by a translation then so are their permutations.

Next the homeomorphism $\mathbb{C}\backslash\mathbb{C}^3 \cong \mathbb{C}^2$ (note that for convenience we will often suppress the fact that we have removed subsets) given by the map $[(z_1, z_2, z_3)] \mapsto (z_1 - z_3, z_2 - z_3)$ induces the identification of the pairs

$$(z_A - z_C, z_B - z_C), (z_B - z_A, z_C - z_A), (z_C - z_B, z_A - z_B),$$
$$(z_A - z_B, z_C - z_B), (z_C - z_A, z_B - z_A), (z_B - z_C, z_A - z_C).$$

If $(w, z) \in \mathbb{C}^2$ is given we find (z_A, z_B, z_C) with $(z_A - z_C, z_B - z_C) = (w, z)$. The six pairs above written in terms of w, z are

$$(w, z), (z - w, -w), (-z, w - z), (w - z, -z), (-w, z - w), (z, w). \quad (4.4.1)$$

One checks that the algorithm which produces these six pairs out of the pair (w, z) also produces the same six pairs if one starts, for example, with $(w - z, -z) = (\xi, \eta)$ and computes the points $(\xi, \eta), (\eta - \xi, -\xi), \ldots$. Thus at this stage the effect of the S_3-action on $\mathbb{C}^2$ is to identify all sets of six pairs as in 4.4.1. (On our space the six pairs are always distinct; we have removed precisely the points where any of the pairs could coincide.)

Next we apply the retraction of $\mathbb{C}^2 - (0,0)$ onto S^3 to our subset of $\mathbb{C}^2$. Notice that (w, z) and (ξ, η) retract to the same point on S^3 if and only if $(\xi, \eta) = (cw, cz)$ for some $c \in \mathbb{R}^+$. In this case the six pairs identified by the S_3-action are likewise related by this positive real scalar c, for example, $(\xi - \eta, -\eta) = \big(c(w - z), c(-z)\big)$.

Thus if we start with $(w, z) \in S^3$, so that $|w|^2 + |z|^2 = 1$, then in $\mathbb{C}^2$ our point gets identified as usual with

$$(w, z), (z, w),$$

$$(w - z, -z), (-z, w - z),$$

$$(-w, z - w), (z - w, -w)$$

but these points are not all on S^3. After the retraction the identification is therefore between

$$(w, z), (z, w),$$

$$\frac{1}{\sqrt{|w - z|^2 + |z|^2}}(w - z, -z), \frac{1}{\sqrt{|w - z|^2 + |z|^2}}(-z, w - z),$$

$$\frac{1}{\sqrt{|w|^2 + |z - w|^2}}(-w, z - w), \frac{1}{\sqrt{|w|^2 + |z - w|^2}}(z - w, -w).$$

The images under the Hopf map $\rho : S^3 \longrightarrow S^2$ of these six points must be identified. Also these are precisely the identifications on S^2 caused by the S_3-action, because if for $\lambda \in S^1$ we have $(w, z), (\lambda w, \lambda z) \in S^3$ in the same S^1-fibre, then $(w - z, -z)$ and $\big(\lambda(w - z), \lambda(-z)\big)$ are also in the same S^1-fibre and so have the same image under ρ, and likewise for the remaining pairs produced by (w, z) and $(\lambda w, \lambda z)$ respectively.

We can describe the identifications on S^2 by starting with a point $(x, y, t) \in S^2$. We have seen that the point

$$(w, z) = \left(\frac{y}{\sqrt{2}\sqrt{1 - x}} + \frac{t}{\sqrt{2}\sqrt{1 - x}} i, \sqrt{\frac{1 - x}{2}} \right)$$

maps to (x, y, t) under ρ. For this (w, z) the six images under ρ of the points in 4.4.1

are

$$(x, y, t)$$

$$(-x, y, -t)$$

$$\left(\frac{2(1-y)-(1-x)}{2(1-y)+(1-x)}, 2\frac{1-x-y}{2(1-y)+(1-x)}, \frac{-2t}{2(1-y)+(1-x)}\right)$$

$$\left(\frac{(1-x)-2(1-y)}{2(1-y)+(1-x)}, 2\frac{1-x-y}{2(1-y)+(1-x)}, \frac{2t}{2(1-y)+(1-x)}\right)$$

$$\left(\frac{(1+x)-2(1-y)}{2(1-y)+(1+x)}, 2\frac{1+x-y}{2(1-y)+(1+x)}, \frac{-2t}{2(1-y)+(1+x)}\right)$$

$$\left(\frac{2(1-y)-(1+x)}{2(1-y)+(1+x)}, 2\frac{1+x-y}{2(1-y)+(1+x)}, \frac{2t}{2(1-y)+(1+x)}\right).$$

One checks that the algorithm that produces these six points out of the triple (x, y, t) produces the same six points if one begins with any other triple in the list; that is, suppose we take the third triple and treat it as the original, writing

$$(u, v, s)$$

$$= \left(\frac{2(1-y)-(1-x)}{2(1-y)+(1-x)}, 2\frac{1-x-y}{2(1-y)+(1-x)}, \frac{-2t}{2(1-y)+(1-x)}\right)$$

and then list the six points $(u, v, s), (-u, v, -s), \ldots$. It is tedious to check, but we get the same set of six points produced by (x, y, t).

Recall that the action of the involution group R had the effect on S^2 of identifying points (x, y, t) and $(x, y, -t)$. One way to look at this is that $R\backslash S^2$ can be obtained as the image of the map $(x, y, t) \mapsto (x, y)$ since (x, y, t) and $(x, y, -t)$ are precisely the points mapping to (x, y). In this picture $R\backslash S^2$ is viewed as the unit disc D^2 in the plane and the six identifications above become

$$(x, y)$$

$$(-x, y)$$

$$\left(\frac{2(1-y)-(1-x)}{2(1-y)+(1-x)}, 2\frac{1-x-y}{2(1-y)+(1-x)}\right)$$

$$\left(\frac{(1-x)-2(1-y)}{2(1-y)+(1-x)}, 2\frac{1-x-y}{2(1-y)+(1-x)}\right)$$

$$\left(\frac{(1+x)-2(1-y)}{2(1-y)+(1+x)}, 2\frac{1+x-y}{2(1-y)+(1+x)}\right)$$

$$\left(\frac{2(1-y)-(1+x)}{2(1-y)+(1+x)}, 2\frac{1+x-y}{2(1-y)+(1+x)}\right)$$

when we drop the t-coordinate. One checks as before that the algorithm producing these six points out of (x, y) produces the same six points if we begin at any other point in the list.

Part of this identification, evident in the first two pairs, is that the left half of the disk gets identified with the right half when (x, y) and $(-x, y)$ get identified. So we work now on the half disk

$$D^2_+ = \left\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1, x \geq 0\right\}$$

where now only three identifications are necessary

$$(x, y)$$

$$\left(\frac{(1-x)-2(1-y)}{2(1-y)+(1-x)}, 2\frac{1-x-y}{2(1-y)+(1-x)}\right)$$

$$\left(\frac{(1+x)-2(1-y)}{2(1-y)+(1+x)}, 2\frac{1+x-y}{2(1-y)+(1+x)}\right).$$

The points (x, y) in the half disk D^2_+ where

$$\frac{(1-x)-2(1-y)}{2(1-y)+(1-x)} \geq 0$$

are those points for which $x \leq 2y-1$, and the points (x, y) in the half disk D^2_+ where

$$\frac{(1+x)-2(1-y)}{2(1-y)+(1+x)} \geq 0$$

are those points for which $x \geq -2y+1$. We divide the half disk D^2_+ into the regions

$$R_1 = D^2_+ \cap \{x \leq 2y-1\}$$

$$R_2 = D^2_+ \cap \{x \geq 2y-1\} \cap \{x \geq -2y+1\}$$

$$R_3 = D^2_+ \cap \{x \leq -2y+1\}.$$

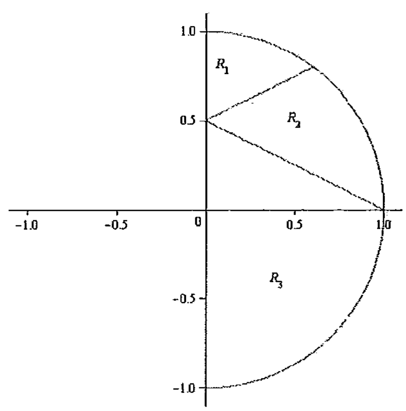


Figure 4.1: D^2_+ and the regions R_1, R_2, R_3

The identification

$$(x,y) \sim \left(\frac{(1+x)-2(1-y)}{2(1-y)+(1+x)}, 2\frac{1+x-y}{2(1-y)+(1+x)}\right)$$

identifies the regions R_1 and R_2, folding (and stretching) R_1 onto R_2 along the common boundary segment $x = 2y - 1$ and matching up the line segments $x = 0$ and $x = -2y + 1$ as shown.

The identification

$$(x,y) \sim \left(\frac{(1-x)-2(1-y)}{2(1-y)+(1-x)}, 2\frac{1-x-y}{2(1-y)+(1-x)}\right)$$

identifies the regions R_1 and R_3, which can be visualized as R_1 swinging clockwise (and stretching) to coincide with R_3, or alternatively as R_2 folding over onto R_3 after R_1 and R_2 are identified. The end result is just the space R_3, shaped like a distorted pie slice, which remains contractible.

Notice that the points $(-1,0,0)$, $(1,0,0)$, $(0,1,0)$ removed from S^2 end up at $(-1,0)$, $(1,0)$, $(0,1)$ on the disk D^2 and finally, after all identifications, at the corner $(1,0)$ on R_3, leaving R_3 contractible.

4.5 Local charts for the parameter space T

Observe that the section $s : T \longrightarrow \mathbb{H}^3 - \Sigma$ gives a homeomorphism onto its image: s is surjective onto its image; s is injective because it maps distinct points into distinct fibres; we have shown that s is continuous; and s^{-1} is continuous because s^{-1} is the restriction of the bundle projection p to the image of s.

Thus the image $s(T)$ provides yet another characterization of the moduli space of Heisenberg triangles. Described explicitly this image is the space of standard triangles

$$s(T) = \{(O,B,C) \in \mathbb{H}^3 - \Sigma : x_B = 0, y_B > 0\}.$$

The $s(T)$ moduli space is useful because we can directly obtain local charts. Given a point $(O,B,C) \in s(T)$. We choose any connected 2-dimensional open subset V_B of the yt-plane containing B, such that $y > 0$ for all points in V_B. Then we choose any connected 3-dimensional open subset V_C containing C, such that no point of V_C projects to the origin O. Additionally we need to make sure that V_B and V_C do not intersect, and moreover that no two points in V_B and V_C project to the same point in the xy-plane.

With such V_B and V_C we obtain the smooth local chart

$$\{O\} \times V_B \times V_C \subset s(T) \longrightarrow V_B \times V_C \subset \mathbb{R}^5$$

by simply dropping the origin $(O, B', C') \mapsto (B', C')$. The advantage of such a chart is that it offers a direct picture of the topology of the moduli space of Heisenberg triangles. We can visualize a path in $s(T)$ (and thereby equivalently a path in T) by visualizing the vertices B and C moving within V_B and V_C respectively.

4.6 T is embedded in $\mathbb{R}^6$

We have seen that T is a 5-dimensional injectively immersed submanifold of $\mathbb{R}^6$. Now we show that in addition to being injectively immersed in $\mathbb{R}^6$, the topology on T is the same as the subspace topology on T induced by the inclusion $T \subset \mathbb{R}^6$. Thus T is embedded in $\mathbb{R}^6$.

The space T sits inside the open subset

$$U := \mathbb{R} \times (-\pi, \pi) \times \mathbb{R} \times (-\pi, \pi) \times \mathbb{R} \times (-\pi, \pi) \subset \mathbb{R}^6.$$

For this discussion only we rearrange the order of the coordinates on $\mathbb{R}^6$ so that a 6-tuple in T looks like

$$(a, b, c; \zeta_A, \zeta_B, \zeta_C).$$

In fact it will be easier to work with the side-lengths of the projected Euclidean triangles than with the side-lengths of the Heisenberg triangles, so on U we apply the transformation

$$f : U \to \mathbb{R}^6$$

$$(a, b, c; \zeta_A, \zeta_B, \zeta_C) \mapsto \left(\frac{a}{\sigma(\zeta_A)}, \frac{b}{\sigma(\zeta_B)}, \frac{c}{\sigma(\zeta_C)}; \zeta_A, \zeta_B, \zeta_C \right).$$

It is straightforward to check that the map f is a continuous bijection onto its image. The differential Df of the map is

$$\begin{bmatrix} \frac{1}{\sigma(\zeta_A)} & 0 & 0 & & & \\ 0 & \frac{1}{\sigma(\zeta_B)} & 0 & & X & \\ 0 & 0 & \frac{1}{\sigma(\zeta_C)} & & & \\ & & & 1 & 0 & 0 \\ & 0 & & 0 & 1 & 0 \\ & & & 0 & 0 & 1 \end{bmatrix}$$

which has rank 6 so that f is an open map. The point is that f maps the open set $U \subset \mathbb{R}^6$ diffeomorphically to the open set $f(U) \subset \mathbb{R}^6$. We will show that $f(T)$ is embedded in the open set $f(U) \subset \mathbb{R}^6$ and therefore is embedded in $\mathbb{R}^6$ itself. It follows that T is likewise embedded in the open set $U \subset \mathbb{R}^6$ and therefore is embedded in $\mathbb{R}^6$ itself.

Given a fixed point

$$P_0 = \left(\tilde{a}_0, \tilde{b}_0, \tilde{c}_0; \zeta_{A_0}, \zeta_{B_0}, \zeta_{C_0}\right) \in f(T).$$

This 6-tuple of triangle parameters represents an equivalence class of Heisenberg triangles up to isometry in $\mathbb{H} \rtimes SO(2)$ with the unique standard triangle representative

$$\begin{aligned}&\left(0,0,0;0,\tilde{c}_0,\tilde{c}_0^2\mu(\zeta_{C_0});\tilde{b}_0\sin\alpha_0,\tilde{b}_0\cos\alpha_0,-\tilde{b}_0^2\mu(\zeta_{B_0})\right)\\ &= (O, B_0, C_0)\end{aligned}$$

where α_0 is the angle opposite the side of length $\tilde{a}_0$ in a Euclidean triangle with side-lengths $\tilde{a}_0$, $\tilde{b}_0$, $\tilde{c}_0$. (We assume the 6-tuple of parameters satisfies the defining equation with positive sign; if the sign were negative, the coordinate x_C would be $-\tilde{b}_0 \sin\alpha_0$.)

We now construct a chart for T as in section 4.5 of a particularly convenient form. We want the parameters $\tilde{b}$, $\tilde{c}$, ζ_B, ζ_C to have complete freedom to range through open intervals

$$\tilde{b}_0 - \varepsilon_1 < \tilde{b} < \tilde{b}_0 + \varepsilon_1$$

$$\tilde{c}_0 - \varepsilon_2 < \tilde{c} < \tilde{c}_0 + \varepsilon_2$$

$$\zeta_{B_0} - \delta_1 < \zeta_B < \zeta_{B_0} + \delta_1$$

$$\zeta_{C_0} - \delta_2 < \zeta_C < \zeta_{C_0} + \delta_2.$$

These requirements determine the neighborhood V_{B_0} around B_0. Indeed it is clear that B must project to a point $(0, \tilde{c}, 0)$ on the positive y-axis with $\tilde{c}_0 - \varepsilon_2 < \tilde{c} < \tilde{c}_0 + \varepsilon_2$ so that B lies within a rectangular strip in the yt-plane.

Now since

$$t_B = \tilde{c}^2 \mu(\zeta_C)$$

we see that B lies in the region of the yt-plane defined by the boundary conditions

$$\tilde{c}_0 - \varepsilon_2 < y_B < \tilde{c}_0 + \varepsilon_2$$

$$y_B^2 \mu(\zeta_{C_0} - \delta_2) < t_B < y_B^2 \mu(\zeta_{C_0} + \delta_2).$$

This is the neighborhood V_{B_0}.

To describe the region V_{C_0} recall that for the projected frame triangle of OB_0C_0 the side-lenghts $\tilde{a}_0$, $\tilde{b}_0$, $\tilde{c}_0$ determine an angle α_0 at the vertex O. Since the vertex C is free to move in 3 dimensions we can vary α freely through an open range around α_0

$$\alpha_0 - \tau < \alpha < \alpha_0 + \tau$$

while still letting $\tilde{b}$ and ζ_B vary freely through their respective open ranges. The resulting neighborhood V_{C_0} is the 3-dimensional region produced by rotating a region shaped like V_{B_0} about the origin through an angle of 2τ.

We take these choices of V_{B_0} and V_{C_0} as the regions giving a local chart around OB_0C_0. We now describe the image of this region in the parameter moduli space $f(T)$. We already know that the parameters $\tilde{b}$, $\tilde{c}$, ζ_B, ζ_C range freely in open intervals determining a four-dimensional open rectangle. The third side-length $\tilde{a}$ is determined by $\tilde{b}$, $\tilde{c}$ and α. We have

$$\tilde{a} = \sqrt{\tilde{b}^2 + \tilde{c}^2 - 2\tilde{b}\tilde{c}\cos\alpha}$$

and it is straightforward to check that with this formula the map

$$(\alpha, \tilde{b}, \tilde{c}, \zeta_B, \zeta_C) \mapsto (\tilde{a}, \tilde{b}, \tilde{c}, \zeta_B, \zeta_C)$$

is a rank 5 map except where both $\tilde{b} = \tilde{c}$ and $\alpha = 0$, and therefore is rank 5 on $\{O\} \times V_{B_0} \times V_{C_0}$. This 5-dimensional portion of the image parameters $(\tilde{a}, \tilde{b}, \tilde{c}, \zeta_B, \zeta_C)$ is therefore an open set in $\mathbb{R}^5$ which we call W'.

We now use this open set $W' \subset \mathbb{R}^5$ to construct an open subset of $\mathbb{R}^6$. To do this we choose $\delta_3 > 0$ such that the interval $(\zeta_{A_0} - \delta_3, \zeta_{A_0} + \delta_3)$ is contained in $(-\pi, \pi)$. Then we produce the following open subset of $\mathbb{R}^6$

$$W := W' \times (\zeta_{A_0} - \delta_3, \zeta_{A_0} + \delta_3)$$

(above we really mean that the ζ_A interval is in the 4th coordinate slot but to avoid cumbersome notation we do not explicitly indicate this). Thus in fact W is an open subset of $f(U)$ and in particular W contains the point $(\tilde{a}_0, \tilde{b}_0, \tilde{c}_0; \zeta_{A_0}, \zeta_{B_0}, \zeta_{C_0})$.

Suppose now that the point

$$P = (\tilde{a}, \tilde{b}, \tilde{c}; \zeta_A, \zeta_B, \zeta_C)$$

is contained in $W \cap f(T)$. Then the five parameters

$$\tilde{a}, \tilde{b}, \tilde{c}, \zeta_B, \zeta_C$$

are contained in W'.

By construction there is a standard Euclidean frame triangle with side-lengths $\tilde{a}$, $\tilde{b}$, $\tilde{c}$ such that the three vertices lie respectively at the origin, in the projection of V_{B_0} onto the positive y-axis, and in the projection of V_{C_0} onto the xy-plane. If we choose heights

$$t_B = \tilde{c}^2\mu(\zeta_C)$$

$$t_C = -\tilde{b}^2\mu(\zeta_B)$$

over the corresponding vertices of the frame triangle, then we have constructed a standard Heisenberg triangle with vertices B and C inside V_{B_0} and V_{C_0} respectively. This triangle OBC has the five parameters

$$\tilde{a}, \tilde{b}, \tilde{c}, \zeta_B, \zeta_C.$$

Now the sixth parameter ζ for the triangle OBC is completely determined by the points B and C. But this parameter will satisfy precisely the same defining equation as that satisfied by the parameters corresponding to the point $P \in W \cap f(T)$. It follows that ζ and ζ_C coincide.

Therefore the unique standard Heisenberg triangle corresponding to $P \in W \cap f(T)$ is actually inside the neighborhood $V_{B_0} \times V_{C_0}$. That is

$$W \cap f(T) \cong V_{B_0} \times V_{C_0}.$$

We have therefore found an open set $W \subset f(U) \subset \mathbb{R}^6$ such that $W \cap f(T)$ is a local chart around the point P_0. The topology on $f(T)$ induced by the ambient space $\mathbb{R}^6$ is thus the same as the moduli space topology on $f(T)$ induced by the original map $p : \mathbb{H}^3 - \Sigma \to T$.

Therefore $f(T)$ is an embedded submanifold of $\mathbb{R}^6$ and likewise therefore T is an embedded submanifold of $\mathbb{R}^6$.

4.7 The boundary of T

As a 5-dimensional embedded submanifold $T \subset \mathbb{R}^6$, every point of T is a boundary point of T in the topological sense, when T is viewed as a subset of $\mathbb{R}^6$. Nonetheless some points of $\mathbb{R}^6$ will be in the closure $\overline{T}$ but not in T, and such points can be viewed as a kind of boundary for T. Henceforth when we refer to the boundary of T, we will mean $\overline{T} - T$.

To describe this boundary, it will be useful to have an explicit criterion to decide whether the parameters $(\zeta_A, \zeta_B, \zeta_C) \in (-\pi, \pi)^3$ can be obtained as the ζ-values for a triangle ABC.

4.7.1 The space of ζ-values

What triples $(\zeta_A, \zeta_B, \zeta_C)$ are possible? Equivalently, what triples $(\mu(\zeta_A), \mu(\zeta_B), \mu(\zeta_C)) = (\lambda_1, \lambda_2, \lambda_3)$ are possible?

Whatever triples $(\lambda_1, \lambda_2, \lambda_3)$ we can get, we can get from triangles constructed from points of the form

$$O(0,0,0)$$

$$A(x,0,t)$$

$$B(u,v,s)$$

with $x \neq 0$, $(u,v) \neq (x,0)$, $u^2+v^2=1$, by taking advantage of the isometries of $\mathbb{H}$. We think of $(\lambda_1, \lambda_2, \lambda_3)$ as fixed, and decide whether we can find values for x, t, u, v, s giving these λ's. Thus we set

$$\lambda_1 = \mu(\zeta_{OA}) = \frac{t}{x^2}$$

$$\lambda_2 = \mu(\zeta_{AB}) = \frac{s-t+2xv}{(u-x)^2+v^2}$$

$$\lambda_3 = \mu(\zeta_{BO}) = \frac{-s}{u^2+v^2} = -s.$$

We obtain the system of equations

$$t = \lambda_1 x^2$$

$$s - t + 2xv = \lambda_2\left[(u-x)^2+v^2\right]$$

$$s = -\lambda_3.$$

Substituting the first and third equations into the second, using the relation $u^2+v^2=1$, and simplifying, we obtain

$$(\lambda_1+\lambda_2)x^2 - 2(v+u\lambda_2)x + (\lambda_2+\lambda_3) = 0$$

viewed as a quadratic in x, still with $x \neq 0$, $(x, 0) \neq (u, v)$, $u^2 + v^2 = 1$. The case $\lambda_1 + \lambda_2 = 0$ will be dealt with separately; for now we assume this is not so and we have a genuine quadratic in x. Therefore if, for fixed $(\lambda_1, \lambda_2, \lambda_3)$, we can find such x, u, v to satisfy the quadratic, then $(\lambda_1, \lambda_2, \lambda_3)$ can be obtained as a triple of $\mu(\zeta)$'s; and likewise if $(\lambda_1, \lambda_2, \lambda_3)$ can be obtained as a triple of $\mu(\zeta)$'s then we can find such x, u, v. So the question becomes, what conditions on $(\lambda_1, \lambda_2, \lambda_3)$ are necessary and sufficient to find such x, u, v satisfying the quadratic?

The quadratic can be solved for real x if and only if the discriminant $b^2 - 4ac$ is non-negative; this condition becomes

$$(v + u\lambda_2)^2 \geq (\lambda_1 + \lambda_2)(\lambda_2 + \lambda_3).$$

So, thinking of λ_1, λ_2, λ_3 as fixed, do there exist u, v such that $u^2 + v^2 = 1$ and the inequality is satisfied? To answer this, we determine the maximum value that $(v + u\lambda_2)^2$ can attain, subject to $u^2 + v^2 = 1$.

For the case $\lambda_2 = 0$ it is immediate that $(v + u\lambda_2)^2$ attains the maximum value 1 when $v = 1$. The condition on the discriminant becomes $1 \geq \lambda_1\lambda_3$ which in this case is equivalent to

$$1 \geq \lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1$$

because $\lambda_2 = 0$.

For the case $\lambda_2 \neq 0$, let

$$(u, v) = (\cos\theta, \sin\theta), \quad -\pi \leq \theta \leq \pi,$$

and let

$$f(\theta) = (\sin\theta + \lambda_2 \cos\theta)^2$$

$$f'(\theta) = 2(\sin\theta + \lambda_2\cos\theta)(\cos\theta - \lambda_2\sin\theta).$$

Setting $f'(\theta) = 0$ we obtain critical points where

$$\sin\theta + \lambda_2\cos\theta = 0, \text{ i.e., } \tan\theta = -\lambda_2$$

and where

$$\cos\theta - \lambda_2\sin\theta = 0, \text{ i.e., } \tan\theta = \frac{1}{\lambda_2}.$$

For $\tan\theta = -\lambda_2$ we compute $f(\theta) = 0$, and for $\tan\theta = \frac{1}{\lambda_2}$ we compute $f(\theta) = 1 + \lambda_2^2$. At the endpoints $\theta = -\pi, \pi$ we compute $f(\theta) = \lambda_2^2$. Thus we see that the largest possible value attained by $(v + u\lambda_2)^2$ subject to $u^2 + v^2 = 1$ is $1 + \lambda_2^2$. Thus the condition on the discriminant becomes

$$1 + \lambda_2^2 \geq (\lambda_1 + \lambda_2)(\lambda_2 + \lambda_3)$$

which simplifies to

$$1 \geq \lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1.$$

Last we address the case $\lambda_1 + \lambda_2 = 0$. Now the equation for which we must find a solution is no longer a quadratic in x but the linear equation

$$2x(v + u\lambda_2) = \lambda_2 + \lambda_3$$

with $x \neq 0$ and $u^2 + v^2 = 1$. If $\lambda_2 + \lambda_3 \neq 0$ then we can always find u and v with $u^2 + v^2 = 1$ such that $v + u\lambda_2 \neq 0$, whereupon we can solve for $x \neq 0$. If $\lambda_2 + \lambda_3 = 0$ then solutions with $x \neq 0$ exist so long as we can find u and v with $u^2 + v^2 = 1$ and $v + u\lambda_2 = 0$. This we can always do by setting $(u, v) = (\cos\theta, \sin\theta)$ and solving $\tan\theta = -\lambda_2$. Thus solutions always exist for the case $\lambda_1 + \lambda_2 = 0$. But in this case

$$\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1 = -\lambda_1^2 \leq 0$$

so that in particular $\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1 \leq 1$ as well.

Theorem 4.7.1. *The triple $(\zeta_A, \zeta_B, \zeta_C)$ occurs as the ζ-values for a triangle ABC if and only if*

$$\mu(\zeta_A)\mu(\zeta_B) + \mu(\zeta_B)\mu(\zeta_C) + \mu(\zeta_C)\mu(\zeta_A) \leq 1.$$

We pause to point out that the work done in this section serves to find equilateral triangles.

Proposition 4.7.2. *There exist triangles having parameters $(a, \zeta, a, \zeta, a, \zeta)$, where $a > 0$ is arbitrary and $\zeta = \mu^{-1}\left(\pm\frac{1}{\sqrt{3}}\right)$.*

Proof. Substituting the parameters in the defining equation 4.2.1 gives $\mu(\zeta) = \pm\frac{1}{\sqrt{3}}$. In the notation above we have $\lambda_1 = \lambda_2 = \lambda_3 = \pm\frac{1}{\sqrt{3}}$, and the work already done makes it straightforward to solve for x, t, u, v, s to obtain the vertices

$$O(0,0,0), A\left(1,0,\frac{1}{\sqrt{3}}\right), B\left(\frac{1}{2},\frac{\sqrt{3}}{2},-\frac{1}{\sqrt{3}}\right)$$

of an equilateral triangle. □

4.7.2 Finding points in the boundary

To obtain points in the boundary $\overline{T} - T$ we use the map

$$p : \mathbb{H}^3 - \Sigma \longrightarrow T \subset \mathbb{R}^6.$$

We take a sequence of triples $(A_n, B_n, C_n) \in \mathbb{H}^3 - \Sigma$ that approach a removed triple $(A, B, C) \in \Sigma$. The sequence of images $p(A_n, B_n, C_n) \in T$ approaches a 6-tuple of parameters that is in $\overline{T}$ but not in T, as we will see. (Alternatively, we take a path in $\mathbb{H}^3 - \Sigma$ with limit $(A, B, C) \in \Sigma$.)

We divide the set of removed triples Σ into types. Recall that Σ was the set of triples

$$A(z_A, t_A), B(z_B, t_B), C(z_C, t_C)$$

for which at least one pair out of the three projections z_A, z_B, z_C coincided. Thus Σ decomposes into

- triples for which exactly two projections coincide
- triples for which all three projections coincide.

The first group further decomposes into

- TYPE I: triples with three distinct vertices
- TYPE II: triples where precisely two vertices coincide (necessarily the two points with coinciding projections)

and the second group further decomposes into

- TYPE III: triples with three distinct vertices
- TYPE IV: triples where precisely two vertices coincide
- TYPE V: triples where all three vertices coincide.

TYPE I:

We left-translate such a triple so that the vertex with distinct projection is at the origin. A typical triple of TYPE I then has the form

$$A(0,0,0)$$

$$B(x,y,t_1)$$

$$C(x,y,t_2)$$

where $(x,y) \neq (0,0)$ and $t_1 < t_2$. All other TYPE I triples (after left-translation to the origin) are obtained as permutations of triples with the form specified above. The parameters

$$\zeta_B = \mu^{-1}\left(\frac{-t_2}{x^2+y^2}\right)$$

$$b = \sigma(\zeta_B)\sqrt{x^2+y^2}$$

and

$$\zeta_C = \mu^{-1}\left(\frac{t_1}{x^2+y^2}\right)$$

$$c = \sigma(\zeta_C)\sqrt{x^2+y^2}$$

are well-defined, and as we approach (A,B,C) via points (A_n,B_n,C_n) in $\mathbb{H}^3 - \Sigma$, the sequences (b_n,ζ_{B_n}), (c_n,ζ_{C_n}) approach (b,ζ_B), (c,ζ_C) respectively, by continuity.

Moreover the lengths $a_n = d(B_n, C_n)$ will approach the length of the geodesics from B to C, which exist but are not unique. Thus

$$a = \sqrt{\pi}\sqrt{t_2 - t_1}.$$

Finally as B_n and C_n approach B and C, $\zeta_{A_n} = \zeta_{B_n C_n} \to \pi$. Therefore we obtain the parameters

$$\left(\sqrt{\pi}\sqrt{t_2 - t_1}, \pi\right), \left(\sigma(\zeta_B)\sqrt{x^2+y^2}, \zeta_B\right), \left(\sigma(\zeta_C)\sqrt{x^2+y^2}, \zeta_C\right) \in \overline{T} - T.$$

Notice that the parameters do not depend directly on $(x, y) \neq (0, 0)$ but only on the value $0 < r^2 = x^2 + y^2$. We can set up a map sending the parameters (r, t_1, t_2) to the 6-tuple

$$\left(\sqrt{\pi}\sqrt{t_2 - t_1}, \pi\right), \left(\sigma(\zeta_B)r, \mu^{-1}(-t_2/r^2)\right), \left(\sigma(\zeta_C)r, \mu^{-1}(t_1/r^2)\right).$$

This map, defined on the open subset

$$\left\{(r, t_1, t_2) \in \mathbb{R}^3 | r > 0, t_1 < t_2\right\}$$

is of rank 3. We need look only at the 1st, 4th, and 6th rows of the resulting 6×3 differential matrix

$$\begin{bmatrix} 0 & -\frac{\sqrt{\pi}}{2}\frac{1}{\sqrt{t_2-t_1}} & \frac{\sqrt{\pi}}{2}\frac{1}{\sqrt{t_2-t_1}} \end{bmatrix}$$

$$\begin{bmatrix} \frac{2t_2}{r^3}(\mu^{-1})'\left(-\frac{t_2}{r^2}\right) & 0 & -\frac{1}{r^2}(\mu^{-1})'\left(-\frac{t_2}{r^2}\right) \end{bmatrix}$$

$$\begin{bmatrix} -\frac{2t_1}{r^3}(\mu^{-1})'\left(\frac{t_1}{r^2}\right) & \frac{1}{r^2}(\mu^{-1})'\left(\frac{t_1}{r^2}\right) & 0 \end{bmatrix}$$

to find three linearly independent rows. The linear independence follows from the facts that $r > 0$, that $(\mu^{-1})' > 0$, and that $t_1 < t_2$, so in particular $t_2 - t_1 > 0$ and t_1 and t_2 are not both 0. Therefore this portion of the boundary $\overline{T} - T$ sits in $\mathbb{R}^6$ as a 3-dimensional immersed submanifold.

Note also that we have described only one component of the portion of the boundary obtained from points of this type. There are six such components, corresponding to the six permutations of a typical triple ABC. These components can also be counted by considering the three possible positions for the two values $\zeta = \pm\pi$.

TYPE II:

We again left-translate and write down a typical triple of this type (of which all other left-translated triples of the same type are permutations)

$$A(0,0,0)$$

$$B(x,y,t)$$

$$C(x,y,t)$$

with $(x, y) \neq (0,0)$. As before the parameters b, ζ_B, c, ζ_C are determined by continuity, no matter how $(A_n, B_n, C_n) \in \mathbb{H}^3 - \Sigma$ approach (A, B, C):

$$\zeta_B = \mu^{-1}\left(\frac{-t}{x^2+y^2}\right)$$

$$b = \sigma(\zeta_B)\sqrt{x^2+y^2}$$

and

$$\zeta_C = \mu^{-1}\left(\frac{t}{x^2+y^2}\right)$$

$$c = \sigma(\zeta_C)\sqrt{x^2+y^2}.$$

Likewise the lengths $a_n = d(B_n, C_n)$ must approach 0 as B_n and C_n approach the mutual limit $B = C$. However the sequence ζ_{A_n} can be made to approach any limit in $[-\pi, \pi]$ by judicious choice of the paths by which B_n and C_n approach the point $B = C$. In fact we can fix $A = (0,0,0)$ and $B = (x, y, t)$ and let C approach B continuously:

$$A = (0,0,0)$$

$$B = (x,y,t)$$

$$C = (x,y,t) \circ h(a,b,1)$$

where $a^2 + b^2 \neq 0$, and we take the limit as h approaches 0. Then

$$\begin{aligned} C &= (x, y, t) \circ h(a, b, 1) \\ &= (x, y, t) \circ (ha, hb, h^2) \\ &= \left(x + ha, y + hb, t + h^2 - 2h(xb - ya)\right). \end{aligned}$$

Now

$$\begin{aligned} \zeta_A &= \zeta \text{ for } B \text{ to } C \\ &= \zeta \text{ for } O \text{ to } B^{-1}C \end{aligned}$$

and

$$\begin{aligned} B^{-1}C &= (-x, -y, -t) \circ \left(x + ha, y + hb, t + h^2 - 2h(xb - ya)\right) \\ &= (ha, hb, h^2) \end{aligned}$$

so

$$\zeta_A = \mu^{-1}\left(\frac{h^2}{(ha)^2 + (hb)^2}\right) = \mu^{-1}\left(\frac{1}{a^2 + b^2}\right)$$

is constant along the path. By ranging $a^2 + b^2$ on $(0, \infty)$ we can obtain any value in $(0, \pi)$ for ζ_A.

Similarly, by using $h(a, b, -1)$ for the continuous path towards C, we can obtain any value in $(-\pi, 0)$ for ζ_A. By using $h(a, b, 0)$ for the continuous path towards C, we get $\zeta_A = 0$.

To get $\zeta_A = \pm\pi$, we simply choose a path where C approaches B more steeply, from above for $\zeta = \pi$ and from below for $\zeta = -\pi$. For example the path

$$C = (x + h^2, y + h^2, t + h)$$

taking $h \to 0^+$ gives

$$B^{-1}C = \left(h^2, h^2, h + 2h^2(x - y)\right)$$

so that

$$\mu(\zeta_A) = \frac{h + 2h^2(x - y)}{2h^4} = \frac{\frac{1}{h^3} + \frac{2(x-y)}{h^2}}{2} \longrightarrow \infty$$

as $h \to 0^+$. Hence $\zeta_A = \pi$.

Now that we have categorized the boundary points obtained by approaching TYPE II points in Σ we can parametrize such points by the map sending (ζ_A, r, t) to the image

$$(0, \zeta_A), (\sigma(\zeta_B)r, \mu^{-1}(-t/r^2)), (\sigma(\zeta_C)r, \mu^{-1}(t/r^2)).$$

We obtain maps from the domains

$$\{(\zeta_A, r, t) \in \mathbb{R}^3 | -\pi < \zeta_A < \pi, r > 0\}$$

$$\{(\pm\pi, r, t) | r > 0\} \cong \{(r, t) | r > 0\}$$

of ranks 3 and 2 respectively, giving one 3-dimensional immersed submanifold component and two 2-dimensional immersed submanifold components of the boundary $\overline{T} - T$.

As before, permutations of the vertices ABC give more components of the boundary. In this case the six permutations give rise to only three components. These can be counted by considering the three possible positions for the side-length 0.

TYPE III:

After left-translation a typical triple is

$$A(0,0,0)$$

$$B(0,0,t_1)$$

$$C(0,0,t_2)$$

with $t_1 < t_2$. All other such TYPE III triples are permutations of such a triple. The same arguments presented above show such a triple corresponds to a 6-tuple

$$(\sqrt{\pi}\sqrt{t_2 - t_1}, \pi), (\sqrt{\pi}\sqrt{t_2}, -\pi), (\sqrt{\pi}\sqrt{t_1}, \pi) \in \overline{T} - T.$$

We obtain a rank 2 map defined on the open subset

$$\{(t_1, t_2) \in \mathbb{R}^2 | t_1 < t_2\}$$

mapping to a 2-dimensional immersed submanifold component of the boundary $\overline{T}-T$ in $\mathbb{R}^6$.

Once again, we have described only one of six such components resulting from the permutations of the triple we have written above.

TYPE IV:

After left-translation a typical triple is

$$A(0,0,0)$$

$$B(0,0,0)$$

$$C(0,0,t)$$

with $t > 0$, which gives a 6-tuple

$$\left(\sqrt{\pi}\sqrt{t},\pi\right),\left(\sqrt{\pi}\sqrt{t},-\pi\right),(0,\zeta_C) \in \overline{T}-T$$

where ζ_C can be in the range $-\pi \leq \zeta_C \leq \pi$.

We obtain maps from the domains

$$\{(\zeta_C,t)| -\pi < \zeta_C < \pi, t > 0\}$$

$$\{(\pm\pi,t)|t > 0\} \cong \{t|t > 0\}$$

of ranks 2 and 1 respectively, giving one 2-dimensional immersed submanifold component and two 1-dimensional immersed submanifold components of the boundary $\overline{T}-T$.

Once again, we have described only one of six such components coming from permutations of the triples.

TYPE V:

All vertices A, B, C converge to the same point, say $(0,0,0)$. All side lengths a, b, c converge to 0. Using the non-isotropic dilation by λ and taking $\lambda \to 0$ we can obtain any parameters

$$(0, \zeta_A), (0, \zeta_B), (0, \zeta_C)$$

where the ζ's are realizable as the ζ-values for a triangle.

This portion of $\overline{T} - T$ can be obtained by first looking at the space

$$\{(\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3 | \lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1 \leq 1\}$$

which is a 3-dimensional manifold, with boundary where $\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1 = 1$. The image of this space under the map

$$(\lambda_1, \lambda_2, \lambda_3) \mapsto (\mu^{-1}(\lambda_1), \mu^{-1}(\lambda_2), \mu^{-1}(\lambda_3))$$

is then the union of a 3-dimensional manifold (the image of where $\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1 < 1$) and a 2-dimensional manifold (the image of where $\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1 = 1$). With this construction we do not need to account for permutations for this component of the boundary.

In fact we can obtain ζ-values of $\pm\pi$, so long as they work together to keep

$$\mu(\zeta_A)\mu(\zeta_B) + \mu(\zeta_B)\mu(\zeta_C) + \mu(\zeta_C)\mu(\zeta_A) \leq 1$$

in the limit. Thus, we can obtain triples of the form

$$(0, \zeta_A), (0, \pi), (0, -\pi)$$

$$(0, \pm\pi), (0, \pi), (0, -\pi)$$

where $-\pi < \zeta_A < \pi$. This portion of the boundary $\overline{T} - T$ is a 1-dimensional manifold with boundary where $\zeta_A = \pm\pi$. And here we obtain 6 components of the boundary.

4.7.3 Do we have them all?

We repeat the full description of the space T as a subset of $\mathbb{R}^6$. The space T consists of all 6-tuples

$$(a, \zeta_A), (b, \zeta_B), (c, \zeta_C)$$

satisfying the inequalities

$$a, b, c > 0$$

$$-\pi < \zeta_A, \zeta_B, \zeta_C < \pi$$

$$a + b \geq c, b + c \geq a, c + a \geq b$$

and the defining equation (4.2.1)

$$\tilde{a}^2\mu(\zeta_A) + \tilde{b}^2\mu(\zeta_B) + \tilde{c}^2\mu(\zeta_C) = \pm\sqrt{(\tilde{a}^2 + \tilde{b}^2 + \tilde{c}^2)^2 - 2(\tilde{a}^4 + \tilde{b}^4 + \tilde{c}^4)}$$

where

$$\tilde{a} = \frac{a}{\sigma(\zeta_A)}, \tilde{b} = \frac{b}{\sigma(\zeta_B)}, \tilde{c} = \frac{c}{\sigma(\zeta_B)}.$$

Additionally we know that for such points the following inequalities are satisfied

$$\tilde{a} + \tilde{b} \geq \tilde{c}, \tilde{b} + \tilde{c} \geq \tilde{a}, \tilde{c} + \tilde{a} \geq \tilde{b}$$

$$\mu(\zeta_A)\mu(\zeta_B) + \mu(\zeta_B)\mu(\zeta_C) + \mu(\zeta_C)\mu(\zeta_A) \leq 1.$$

An equivalent characterization of equation (4.2.1) will be useful

$$\tilde{a}^2\mu(\zeta_A) + \tilde{b}^2\mu(\zeta_B) + \tilde{c}^2\mu(\zeta_C) = \pm 4 \text{ area } \triangle$$

where the triangle has side lengths $\tilde{a}$, $\tilde{b}$, $\tilde{c}$.

For points in the closure $\overline{T} \subset \mathbb{R}^6$, all the same inequalities must still be satisfied, except that strict inequalites are replaced with non-strict inequalites. Roughly speaking, this means that now the side-lengths a, b, c can be 0 and that the ζ-values can be $\pm\pi$.

When (a, ζ_A) makes part of a 6-tuple in $\overline{T}$, we can still define $\tilde{a} = \frac{a}{\sigma(\zeta_A)}$ and this value $\tilde{a}$ agrees with the limit of the "tilde"-lengths for whatever sequence of points (a_n, ζ_{A_n}) limited to (a, ζ_A) in the closure. Notice then that for a point in the closure, $a = 0 \Rightarrow \tilde{a} = 0$.

Points in the closure will still satisfy equation (4.2.1) *so long as it is still defined.* For closure points with a, b, or c of length 0 the equation makes sense, but for closure points with ζ-values of $\pm\pi$ the equation no longer makes sense and we must be more careful.

Having made these remarks, we will attempt to classify all points in $\overline{T} - T$ and make sure that we have hit them all by our previous analysis based on the map p. We will anchor the analysis by focusing on the ζ-values. There are four cases:

- (i) none of the ζ's are $\pm\pi$
- (ii) exactly one of the ζ's is $\pm\pi$
- (iii) exactly two of the ζ's are $\pm\pi$
- (iv) all three of the ζ's are $\pm\pi$.

Case (i): $(a, \zeta_A), (b, \zeta_B), (c, \zeta_C) \in \overline{T} - T$ with $-\pi < \zeta_A, \zeta_B, \zeta_C < \pi$. Now if $a, b, c > 0$ then our point is in fact in T and is therefore not a point in the boundary $\overline{T} - T$.

Suppose WLOG that $a = 0$ and $b, c > 0$. Then $\tilde{a} = 0$ and from the triangle inequalities on $\tilde{a}$, $\tilde{b}$, $\tilde{c}$ we see that $\tilde{b} = \tilde{c}$. Moreover the triangle with side lengths $\tilde{a}$, $\tilde{b}$, $\tilde{c}$ has area 0. Hence equation (4.2.1) becomes

$$\tilde{b}^2\mu(\zeta_B) + \tilde{c}^2\mu(\zeta_C) = 0$$

and with $\tilde{b} = \tilde{c} \neq 0$ this gives $\zeta_B = -\zeta_C$, whereupon $b = c$ also. We have thus obtained points of the form

$$(0, \zeta_A), (b, \zeta_B), (b, -\zeta_B)$$

all of which we can hit by approaching triples in Σ of type II.

The other possibility in case (i) is WLOG $a = 0$ and one of b, c is also 0. The triangle inequalities then force the remaining side-length to be 0. We obtain points of the form

$$(0, \zeta_A), (0, \zeta_B), (0, \zeta_C)$$

all of which can be hit by approaching triples in Σ of type V.

Case (ii): $(a, \zeta_A), (b, \zeta_B), (c, \zeta_C) \in \overline{T} - T$ with exactly one of the ζ's at $\pm\pi$, WLOG $\zeta_A = \pi$. Then $\tilde{a} = \frac{a}{\sigma(\pi)} = 0$. The triangle inequalities on $\tilde{a}$, $\tilde{b}$, $\tilde{c}$ then force $\tilde{b} = \tilde{c}$.

We obtain points of the form

$$(a,\pi),(\tilde{b}\sigma(\zeta_B),\zeta_B),(\tilde{b}\sigma(\zeta_C),\zeta_C).$$

When $\tilde{b}\neq 0$ these can be hit by approaching triples in Σ of type I. Indeed, we want to find (r,t_1,t_2) such that

$$(\sqrt{\pi}\sqrt{t_2-t_1},\pi),(\sigma(\zeta_B)r,\zeta_B),(\sigma(\zeta_C)r,\zeta_C)$$

$$=(a,\pi),(\tilde{b}\sigma(\zeta_B),\zeta_B),(\tilde{b}\sigma(\zeta_C),\zeta_C).$$

Here we are viewing the 6-tuple on the right as an arbitrary case (ii) point, and we are trying to show that it can be obtained by the process of approaching removed triples in Σ. The expression on the left is the general form of such images. Recall that on the right ζ_B is not arbitrary, but is determined by $\mu(\zeta_B)=t_1/r^2$, and likewise $\mu(\zeta_C)=-t_2/r^2$. We show that the resulting system of equations has a solution.

First we want

$$a=\sqrt{\pi}\sqrt{t_2-t_1}$$

so we obtain

$$t_2=t_1+a^2/\pi.$$

From $\mu(\zeta_B)=t_1/r^2$ we get $t_1=r^2\mu(\zeta_B)$ and from $\mu(\zeta_C)=-t_2/r^2$ we get $-t_2=r^2\mu(\zeta_C)$. These relations in $t_2=t_1+a^2/\pi$ end up giving

$$r^2\left(\mu(\zeta_B)+\mu(\zeta_C)\right)=-a^2/\pi.$$

Thus r is determined, whereupon t_1 and t_2 are determined, and a quick check shows that these values work.

When $\tilde{b}=0=\tilde{c}$ the fact that $-\pi<\zeta_B,\zeta_C<\pi$ forces $b=0=c$. Then the triangle inequalities on a, b, c force $a=0$ as well and we obtain points of the form

$$(0,\pi),(0,\zeta_B),(0,\zeta_C)$$

all of which can be hit by approaching points of type V.

Case (iii): $(a,\zeta_A),(b,\zeta_B),(c,\zeta_C)\in\overline{T}-T$ where (WLOG) $\zeta_A=\pm\pi$, $\zeta_B=\pm\pi$ and $-\pi<\zeta_C<\pi$. Now whatever points in T limited to our point in $\overline{T}-T$ we have

$$\zeta_A\longrightarrow\pm\pi,\zeta_B\longrightarrow\pm\pi$$

so

$$\mu(\zeta_A) \longrightarrow \pm\infty, \mu(\zeta_B) \longrightarrow \pm\infty.$$

If the $\pm$ signs agree then in the limit of the inequality

$$1 + \mu(\zeta_C)\left(\frac{1}{\mu(\zeta_A)} + \frac{1}{\mu(\zeta_B)}\right) \leq \frac{1}{\mu(\zeta_A)\mu(\zeta_B)}$$

we obtain $1 \leq 0$, a contradiction. Thus $\zeta_A = \pm\pi$ and $\zeta_B = \mp\pi$ must have opposite signs.

Now $\tilde{a} = \frac{a}{\sigma(\pm\pi)} = 0$ and $\tilde{b} = \frac{b}{\sigma(\mp\pi)} = 0$, whereupon the triangle inequalities on $\tilde{a}$, $\tilde{b}$, $\tilde{c}$ force $\tilde{c} = 0$. Since $\sigma(\zeta_C)$ is finite, this in turn forces $c = \tilde{c}\sigma(\zeta_C) = 0$. Then the triangle inequalities on a, b, c force $a = b$. We have obtained points of the form

$$(a, \pm\pi), (a, \mp\pi), (0, \zeta_C)$$

all of which can be hit by approaching triples in Σ of type IV.

Case (iv): $(a, \zeta_A), (b, \zeta_B), (c, \zeta_C) \in \bar{T} - T$ where ζ_A, ζ_B, ζ_C are all $\pm\pi$. The inequality

$$\mu(\zeta_A)\mu(\zeta_B) + \mu(\zeta_B)\mu(\zeta_C) + \mu(\zeta_C)\mu(\zeta_A) \leq 1$$

shows that ζ_A, ζ_B, ζ_C cannot all have the same sign. WLOG we assume $\zeta_A = \zeta_C = \pi$ and $\zeta_B = -\pi$.

Now $\tilde{a} = \frac{a}{\sigma(\zeta_A)} = 0$. For whatever points (a_n, ζ_{A_n}) limit to (a, ζ_A), we have $\tilde{a}_n = \frac{a_n}{\sigma(\zeta_{A_n})} \longrightarrow 0$.

We also have

$$\frac{\zeta_{A_n}}{\sin\zeta_{A_n}}\tilde{a}_n = a_n \longrightarrow a.$$

Since $\zeta_{A_n} \longrightarrow \pi$ we in particular have

$$\frac{\tilde{a}_n}{\sin\zeta_{A_n}} \longrightarrow \frac{a}{\pi}.$$

Now we have

$$\tilde{a}_n^2\mu(\zeta_{A_n}) + \tilde{b}_n^2\mu(\zeta_{B_n}) + \tilde{c}_n^2\mu(\zeta_{C_n}) = \pm 4 \text{ area } \triangle \longrightarrow 0$$

since the side-lengths $\tilde{a}$, $\tilde{b}$, $\tilde{c}$ of the triangle all go to 0. The left-hand-side of this equation is

$$\tilde{a}_n^2\mu(\zeta_{A_n}) + \tilde{b}_n^2\mu(\zeta_{B_n}) + \tilde{c}_n^2\mu(\zeta_{C_n})$$

$$=\tilde{a}_n^2\left(\frac{\zeta_{A_n} - \sin\zeta_{A_n}\cos\zeta_{A_n}}{\sin^2\zeta_{A_n}}\right)$$

$$+\tilde{b}_n^2\left(\frac{\zeta_{B_n} - \sin\zeta_{B_n}\cos\zeta_{B_n}}{\sin^2\zeta_{B_n}}\right)$$

$$+\tilde{c}_n^2\left(\frac{\zeta_{C_n} - \sin\zeta_{C_n}\cos\zeta_{C_n}}{\sin^2\zeta_{C_n}}\right)$$

$$=\left(\frac{\zeta_{A_n}^2}{\sin^2\zeta_{A_n}}\tilde{a}_n^2\right)\frac{1}{\zeta_{A_n}} - \tilde{a}_n\left(\frac{\tilde{a}_n}{\sin\zeta_{A_n}}\right)\cos\zeta_{A_n}$$

$$+\left(\frac{\zeta_{B_n}^2}{\sin^2\zeta_{B_n}}\tilde{b}_n^2\right)\frac{1}{\zeta_{B_n}} - \tilde{b}_n\left(\frac{\tilde{b}_n}{\sin\zeta_{B_n}}\right)\cos\zeta_{B_n}$$

$$+\left(\frac{\zeta_{C_n}^2}{\sin^2\zeta_{C_n}}\tilde{c}_n^2\right)\frac{1}{\zeta_{C_n}} - \tilde{c}_n\left(\frac{\tilde{c}_n}{\sin\zeta_{C_n}}\right)\cos\zeta_{C_n}.$$

These are all limits we have seen. As $n \longrightarrow \infty$ these go to

$$a^2\frac{1}{\pi} - 0\cdot\frac{a}{\pi}\cdot(-1)$$

$$+b^2\frac{1}{-\pi} - 0\cdot\frac{b}{-\pi}\cdot(-1)$$

$$+c^2\frac{1}{\pi} - 0\cdot\frac{c}{\pi}\cdot(-1).$$

Since the overall limit is 0 we obtain

$$a^2\frac{1}{\pi} - b^2\frac{1}{\pi} + c^2\frac{1}{\pi} = 0$$

or simply $b^2 = a^2 + c^2$. We obtain points of the form

$$(a, \pi), (\sqrt{a^2 + c^2}, -\pi), (c, \pi)$$

all of which can be hit by approaching points in Σ of type III.

4.8 Compactification of T up to dilation

Recall for $\lambda > 0$ the non-isotropic dilation

$$\lambda : \mathbb{H} \longrightarrow \mathbb{H}$$

$$(x, y, t) \mapsto (\lambda x, \lambda y, \lambda^2 t).$$

We can let λ map triples of vertices $(A, B, C) \mapsto (\lambda A, \lambda B, \lambda C)$ to obtain an $\mathbb{R}_+$ action on the space of vertices $\mathbb{H}^3 - \Sigma$.

Between arbitrary points $P(x, y, t)$ and $Q(u, v, s)$ we have

$$P^{-1}Q = \big(u - x, v - y, s - t - 2(uy - xv)\big)$$

while

$$(\lambda P)^{-1}(\lambda Q) = \big(\lambda(u - x), \lambda(v - y), \lambda^2(s - t - 2(uy - xv))\big).$$

In particular we compute that

$$\zeta_{(\lambda P)(\lambda Q)} = \zeta_{PQ}$$

and

$$d(\lambda P, \lambda Q) = \lambda d(P, Q).$$

Thus if the vertices (A, B, C) give a triangle with the parameters

$$(a, \zeta_A), (b, \zeta_B), (c, \zeta_C)$$

then the vertices $(\lambda A, \lambda B, \lambda C)$ give a triangle with the parameters

$$(\lambda a, \zeta_A), (\lambda b, \zeta_B), (\lambda c, \zeta_C).$$

In this way the $\mathbb{R}_+$ action on $\mathbb{H}^3 - \Sigma$ induces an $\mathbb{R}_+$ action on the parameter moduli space T

$$\lambda \circ (a, \zeta_A, b, \zeta_B, c, \zeta_C) = (\lambda a, \zeta_A, \lambda b, \zeta_B, \lambda c, \zeta_C).$$

We introduce the map

$$\Lambda : T \longrightarrow T$$

$$(a, \zeta_A), (b, \zeta_B), (c, \zeta_C) \mapsto \left(\frac{a}{a+b+c}, \zeta_A\right), \left(\frac{b}{a+b+c}, \zeta_B\right), \left(\frac{c}{a+b+c}, \zeta_C\right).$$

Two 6-tuples of triangle parameters in the same orbit under the $\mathbb{R}_+$ action have the same image under Λ, i.e., for $q \in T$ we have $\Lambda(\lambda q) = \Lambda(q)$. Moreover suppose that $\Lambda(q) = \Lambda(q')$, i.e., that

$$\left(\frac{a}{a+b+c}, \zeta_A\right), \left(\frac{b}{a+b+c}, \zeta_B\right), \left(\frac{c}{a+b+c}, \zeta_C\right)$$

$$= \left(\frac{x}{x+y+z}, \zeta_X\right), \left(\frac{y}{x+y+z}, \zeta_Y\right), \left(\frac{z}{x+y+z}, \zeta_Z\right).$$

Then $q' = \lambda q$ for $\lambda = \frac{x+y+z}{a+b+c}$. Therefore

Theorem 4.8.1. *The quotient space $\mathbb{R}_+ \backslash T$ is homeomorphic to the subspace*

$$\Lambda(T) = \{(a, \zeta_A, b, \zeta_B, c, \zeta_C) \in T : a+b+c = 1\}.$$

The space $\Lambda(T)$ is bounded but not closed. The closure $\overline{\Lambda(T)}$ is closed and bounded, therefore compact. The space $\overline{\Lambda(T)}$ can be viewed as the compactification of the moduli space of Heisenberg triangles up to dilation. Points in the boundary $\overline{\Lambda(T)} - \Lambda(T)$ are precisely the points in $\overline{T} - T$ with perimeter equal to 1. For if points in T converge $q_n = (a_n, \zeta_{A_n}, b_n, \zeta_{B_n}, c_n, \zeta_{C_n}) \to q = (a, \zeta_A, b, \zeta_B, c, \zeta_C) \in \overline{T} - T$ with $a + b + c = 1$, then $\Lambda(q_n) \to q \in \overline{\Lambda(T)} - \Lambda(T)$ also; while if points in $\Lambda(T)$ converge to a limit in $\overline{\Lambda(T)}$ then that limit has perimeter 1 and is in $\overline{T}$ since $\Lambda(T) \subset T$.

We can obtain the boundary points in another way, namely by applying the map Λ to those points in $\overline{T} - T$ for which the perimeter is non-zero, for if a sequence

$q_n \in T$ converges to $q \in \overline{T}$ with the perimeter of q non-zero, then for all large n the perimeter of q_n is non-zero also, whereupon $\Lambda(q_n)$ is a sequence in $\Lambda(T)$ converging to $\Lambda(q) \in \overline{\Lambda(T)}$.

Chapter 5

Appendix: Some properties of plane triangles

5.1 Orientation

Consider three distinct points in the plane

$$A(x_A, y_A), B(x_B, y_B), C(x_C, y_C).$$

We describe a criterion to check whether the points A,B,C are colinear. Indeed, A,B,C are colinear if and only if the vectors

$$\overrightarrow{AB} = (x_B - x_A, y_B - y_A)$$

$$\overrightarrow{AC} = (x_C - x_A, y_C - y_A)$$

are dependent. Thus A,B,C are colinear if and only if

$$0 = \begin{vmatrix} x_B - x_A & y_B - y_A \\ x_C - x_A & y_C - y_A \end{vmatrix}$$

$$= (x_B - x_A)(y_C - y_A) - (y_B - y_A)(x_C - x_A).$$

Next consider a triangle determined by an ordered triple of distinct non-colinear points in the plane, again labelled

$$A(x_A, y_A), B(x_B, y_B), C(x_C, y_C).$$

We describe a criterion to check whether $\triangle ABC$ is oriented clockwise or counter-clockwise. Since A,B,C are not colinear, the vectors $\overrightarrow{AB}$, $\overrightarrow{AC}$ are independent and therefore provide a basis for $\mathbb{R}^2$. Now $\triangle ABC$ is oriented clockwise if $\{\overrightarrow{AB}, \overrightarrow{AC}\}$ is a basis for $\mathbb{R}^2$ with negative orientation, and is oriented counter-clockwise if the basis has positive orientation. We thus obtain the following criterion for the orientation of a triangle:

$$(x_B - x_A)(y_C - y_A) - (y_B - y_A)(x_C - x_A) < 0 \Leftrightarrow \triangle ABC \text{ oriented clockwise}$$

$$(x_B - x_A)(y_C - y_A) - (y_B - y_A)(x_C - x_A) > 0 \Leftrightarrow \triangle ABC \text{ oriented counter-clockwise.}$$

5.2 Area

By a well-known formula the vectors $\overrightarrow{AB}$, $\overrightarrow{AC}$ are two sides of a parallelogram with area equal to the absolute value of the determinant of the matrix $\begin{bmatrix} \overrightarrow{AB} \\ \overrightarrow{AC} \end{bmatrix}$. Thus we also have the formula

$$\text{area } \triangle ABC = \pm\frac{1}{2}\begin{vmatrix} x_B - x_A & y_B - y_A \\ x_C - x_A & y_C - y_A \end{vmatrix}.$$

Additionally we record two expressions for the area R of a triangle having side-lengths a, b, c:

$$R = \sqrt{s(s-a)(s-b)(s-c)}$$

$$= \frac{1}{4}\sqrt{(a^2 + b^2 + c^2)^2 - 2(a^4 + b^4 + c^4)}$$

where $s = \frac{a+b+c}{2}$ is the *semi-perimeter*.

Bibliography

[1] Kai Behrend, Brian Conrad, Dan Edidin, Barbara Fantechi, William Fulton, Lothar Göttsche, and Andrew Kresch, *Algebraic stacks*, `http://www.math.uzh.ch/index.php?pr_vo_det&key1=1287&key2=580&no_cache=1`.

[2] Joan S. Birman and Tara E. Brendle, *Braids: a survey*, Handbook of knot theory, Elsevier B. V., Amsterdam, 2005, pp. 19–103. MR MR2179260 (2007a:57004)

[3] Nicolas Bourbaki, *Elements of mathematics. General topology. Part 1*, Hermann, Paris, 1966. MR MR0205210 (34 #5044a)

[4] Ovidiu Calin, Der-Chen Chang, and Peter Greiner, *Geometric mechanics on the Heisenberg group*, Bull. Inst. Math. Acad. Sinica **33** (2005), no. 3, 185–252. MR MR2165285 (2006f:53039)

[5] Edward Fadell and Lee Neuwirth, *Configuration spaces*, Math. Scand. **10** (1962), 111–118. MR MR0141126 (25 #4537)

[6] R. W. Sharpe, *Differential geometry*, Graduate Texts in Mathematics, vol. 166, Springer-Verlag, New York, 1997, Cartan's generalization of Klein's Erlangen program, With a foreword by S. S. Chern. MR MR1453120 (98m:53033)

[7] Michael Spivak, *A comprehensive introduction to differential geometry. Vol. I*, second ed., Publish or Perish Inc., Wilmington, Del., 1979. MR MR532830 (82g:53003a)

[8] Kang-Hai Tan and Xiao-Ping Yang, *Characterisation of the sub-Riemannian isometry groups of H-type groups*, Bull. Austral. Math. Soc. **70** (2004), no. 1, 87–100. MR MR2079363 (2005c:53033)