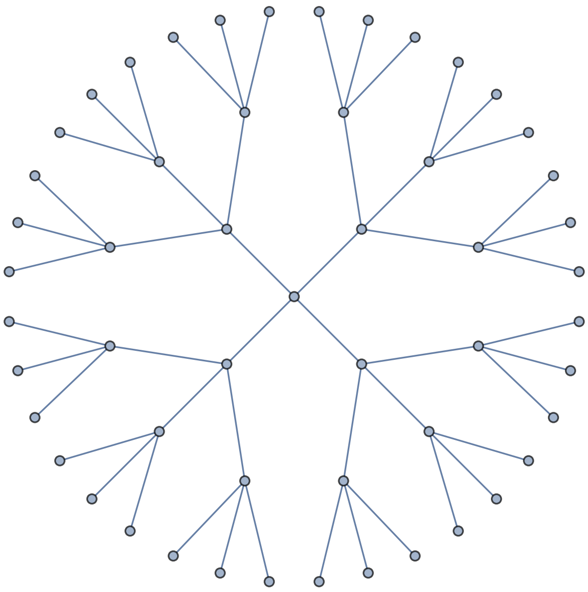
Geometry



Jouni Parkkonen

Lectures at the University of Jyväskylä

fall semester 2024

Contents

Introduction

This text is an introduction to negatively curved spaces. Part I begins with general background on geodesic metric spaces. After this, we study Euclidean and spherical geometry to set the stage for a quick tour of the basics of hyperbolic geometry.

Notations and conventions

For any mapping $f\colon X \to X$, the *fixed point set* of f is

$$\operatorname{fix} f = \{x \in X : f(x) = x\}\,.$$

If a group G acts on a space X and A is a nonempty subset of X, the *stabilizer* of A in G is

$$\operatorname{Stab}_G A = \{g \in G : gA = A\}\,.$$

Clearly, stabilisers are subgroups of G.

- $\mathbb{N} = \{0, 1, 2, \ldots\}$.
- $\#(A) \in \mathbb{N} \cup \{\infty\}$ cardinality of A.
- $A - B = \{a \in A : a \notin B\}$.
- $f|_A$ is the restriction of. mapping $f\colon X \to Y$ to a subset $A \subset X$, $f|_A(a) = f(a)$ for all $a \in A$.
- $A \subsetneq B$ means A is a *proper subset* of B: $A \subset B$ and $A \neq B$.
- $\coprod_{j\in J} X_j = \{(x, j) : x \in X,\ j \in J\}$ is the *disjoint union* of the family of sets $(X_j)_{j\in J}$.
- $\operatorname{diag}(a_1, a_2, \ldots, a_n)$ is the $n \times n$-diagonal matrix with $a_1, a_2, \ldots, a_n$ on the diagonal.
- $\operatorname{diag}(A_1, A_2, \ldots, A_n)$ is the block diagonal matrix with square matrices $A_1, A_2, \ldots, A_n$ on the diagonal.
- $I_n = \operatorname{diag}(1, 1, \ldots, 1)$.
- ${}^tA$ is the transpose of a matrix A.
- $\operatorname{Homeo}(X)$ the group of homeomorphisms of a topological space X.
- $\operatorname{Isom}(X)$ the group of isometries of a metric space X.
- $\mathrm{C}(X, Y)$ space of continuous functions from a topological space X to a metric space Y with the topology of compact convergence.

Definitions are boxed like this and not numbered.

A box like this has some remark or convention that is good to notice!

Part I

Elements

Chapter 1

Geodesic metric spaces

In this chapter, we collect background material on metric spaces, in particular on geodesic spaces. We also introduce some convenient terminology to be used throughout the course.

1.1 Metric spaces

In this section, for the convenience of the reader, we collect some standard definitions, notations and examples on metric spaces. For more details and background, see for example [Bou1, Bou2, Mun].

Let $X \neq \emptyset$. A function $d\colon X \times X \to [0, \infty[$ is a *metric* in X if

(1) $d(x, x) = 0$ for all $x \in X$ and $d(x, y) > 0$ if $x \neq y$,

(2) $d(x, y) = d(y, x)$ for all $x, y \in X$, and

(3) $d(x, y) \leqslant d(x, z) + d(z, y)$ for all $x, y, z \in X$ (the *triangle inequality*).

The pair (X, d) is a *metric space*.

Example 1.1. (a) Any normed space is a metric space. In particular, the space $\mathbb{R}^n$ with the Euclidean distance is a metric space.
(b) The circle $\mathbb{S}^1$ with the distance between two points defined as their angle as vectors in $\mathbb{E}^2$ is a metric space, see Section 3.1 for details and generalisations.
(c) Let $X \neq \emptyset$. The discrete metric $d\colon X \times X \to [0, \infty[$ is defined by setting $d(x, x) = 0$ for all $x \in X$ and $d(x, y) = 1$ for all $x, y \in X$ if $x \neq y$.

Open and closed balls in a metric space, continuity of maps between metric spaces and other "metric properties" are defined in the usual manner. In particular, if X is a metric space, $x \in X$ and $r > 0$,

$$B(x_0, r) = B_d(x_0, r) = \{x \in X : d(x, x_0) < r\}$$

is the open ball of radius r and

$$\overline{B}(x_0, r) = \overline{B}_d(x_0, r) = \{x \in X : d(x, x_0) \leqslant r\}$$

is the closed ball of radius r.

A metric space is *proper* if its closed balls are compact.

Euclidean spaces are proper metric spaces by the theorem of Heine and Borel, see for example [Str, Theorem (3.40)].

1.2 Isometric embeddings and isometries

If (X_1, d_1) and (X_2, d_2) are metric spaces, then a map $i\colon X \to Y$ is an *isometric embedding*, if

$$d_2(i(x), i(y)) = d_1(x, y)$$

for all $x, y \in X_1$.

A map $i\colon X \to Y$ is a *locally isometric embedding* if each point $x \in X$ has a neighbourhood U such that the restriction of i to U is an isometric embedding.

Lemma 1.2. *(a) Isometric embeddings are continuous injective mappings.*

(b) If $f\colon X \to Y$ and $g\colon Y \to Z$ are isometric embeddings, then $g \circ f$ is an isometric embedding.

(b) If $f\colon X \to Y$ and $g\colon Y \to Z$ are locally isometric embeddings, then $g \circ f$ is a locally isometric embedding.

Proof. Exercise. □

If an isometric embedding $i\colon X \to Y$ is a bijection, then it is called an *isometry* between X and Y.

An isometry $i\colon X \to X$ is called an isometry of X.

If (X, d) is a metric space, Y is a set and $f\colon Y \to X$ is a bijection, then we get a metric in Y by setting $d_f(y_1, y_2) = d(f(y_1), f(y_2))$ for all y_1, y_2. Now $f\colon (Y, d_f) \to (X, d)$ is an isometry and it is natural to think of (Y, d_f) as a model of (X, d). We will see concrete examples in Chapter 5 when we consider models of hyperbolic space.

We consider two isometric metric spaces to be *models* of the same abstract metric space.

Proposition 1.3. *The isometries of a metric space X form a group* $\operatorname{Isom}(X)$ *with the composition of mappings as the group law.*

Proof. Exercise 1.1. □

1.3 Group actions

Let Perm(A) be the *group of permutations* of a set A. A group G *acts on* A if there is a homomorphism $\phi\colon G \to \mathrm{Perm}(A)$. The homomorphism ϕ is an *action* of G on A.

Let X be a topological space. A group G *acts on* (X, d) *by homeomorphisms* if there is a homomorphism $\phi\colon G \to \mathrm{Homeo}(X, d)$.

Let (X, d) be a metric space. A group G *acts on* (X, d) *by isometries* if there is a homomorphism $\phi\colon G \to \mathrm{Isom}(X, d)$.

If a group G acts on a set A, we use the notation

$$g \cdot a = \phi(g)(a) = \big(\phi(g)\big)(a)$$

for all $g \in G$ and all $a \in A$. If the group is a subgroup of the permutation group of A, the notation $g(a)$ is natural to use, and if we have an action of a group of matrices on a vector space with a fixed basis,[1] the usual notation of matrix multiplication is used.

In this course, we are mainly interested in actions by isometries but linear action is also used for example in chapters 2 to 4

Let X be a set and let G be a group that acts on X. The *stabilizer (in G)* of a point $x \in X$ is

$$\mathrm{Stab}_G\, x = \{g \in G : g \cdot x = x\}\,.$$

Proposition 1.4. *Let X be a metric space and let $x \in X$. Then* $\mathrm{Stab}\, X$ *is a subgroup of* $\mathrm{Isom}\, X$.

Proof. Exercise 1.1. □

Example 1.5. We will see in section 2.3 that the *Euclidean group*

$$\mathrm{E}(n) = \{x \mapsto Ax + b : A \in \mathrm{O}(n), b \in \mathbb{R}^n\}$$

is the group of isometries of the n-dimensional Euclidean space $\mathbb{E}^n$.[2] The stabilizer of $0 \in \mathbb{E}^n$ in $\mathrm{E}(n)$ is $\mathrm{O}(n)$.

If a group G acts on a space X, and x is a point in X, the set

$$G(x) = \{g \cdot x : g \in G\}$$

is the *G-orbit* of x. The action of a group is said to be *transitive* if $G(x) = X$ for some (and therefore for any) $x \in X$.

A more elementary way to express this is that a group G acts transitively on X if for all $x, y \in X$ there is some $g \in G$ such that $g \cdot x = y$.

[1] We call such an action a *linear action*.

[2] See Theorem, 2.8.

1.4 Geodesics

In this section, we give names to a particularly important class of isometric and locally isometric embeddings and use these objects to define the class of metric spaces that plays a central role in this course.

Let $I \subset \mathbb{R}$ be an interval. A (locally) isometric embedding $i\colon I \to X$ is a *(local) geodesic*. More precisely, it is

(1) a *(locally) geodesic segment*, if $I \subset \mathbb{R}$ is a (closed) bounded interval,

(2) a *(locally) geodesic ray*, if $I = [0, +\infty[$, and

(3) a *(locally) geodesic line*, if $I = \mathbb{R}$.

Note that in Riemannian geometry, the definition of a geodesic is different from the above: If (M, g) is a Riemannian manifold and I is an open interval, a *Riemannian geodesic* $\gamma\colon I \to M$ is a differentiable path whose acceleration is 0. If $\gamma\colon I \to M$ is a Riemannian geodesic, then there is some $c > 0$ and such that the mapping $t \mapsto g(\frac{t}{c})$ is a local geodesic according to our definition. See for example [Lee, Chapter 6] or [Pet, Chapter 5] for more information.

If $\gamma\colon [a, b] \to X$ is a path, then γ *connects* the points $\gamma(a)$ to $\gamma(b)$.

If γ is a geodesic segment that connects $x \in X$ to $y \in X$, the points x and y are the *endpoints* of γ.

Sometimes it is convenient to use more precise terminology and, for instance, refer to the endpoint $j(0)$ as the *origin* of j and to the other endpoint as the *terminal point* or the *terminus* of j.

A metric space (X, d) is a *geodesic metric space*, if for any $x, y \in X$ there is a geodesic segment that connects x to y.

Example 1.6. Any normed space is a geodesic metric space: Let $(V, \|\cdot\|)$ be a normed space. For any two distinct points $x, y \in V$, the map

$$t \overset{j}{\mapsto} x + t\,\frac{y - x}{\|y - x\|},$$

is a geodesic line that passes through the points x and y. Indeed, for any $s, t \in \mathbb{R}$, we have

$$\|j(t) - j(s)\| = \left\| x_0 + t\frac{y - x}{\|y - x\|} - \left(x_0 + s\frac{y - x}{\|y - x\|}\right)\right\| = \left\|(t - s)\frac{y - x}{\|y - x\|}\right\| = |t - s|\,.$$

The restriction $j|_{[0, \|x-y\|]}$ is a geodesic segment that connects x to y.

Example 1.7. It can be shown that $h_\alpha(s, t) = |s - t|^\alpha$ is a metric in $\mathbb{R}$ if $0 < \alpha \leqslant 1$. The metric space $(\mathbb{R}, h_\alpha)$ is homeomorphic to $\mathbb{R}$ with the usual metric given by the expression h_1 but it is not a geodesic metric space if $0 < \alpha < 1$.

A metric space (X,d) is *uniquely geodesic*, if for any $x, y \in X$ there is exactly one geodesic segment that connects x to y.

If X is a uniquely geodesic metric space and $x, y \in X$, $x \neq y$, we denote the (image of the) unique geodesic segment connecting x to y by $[x, y]$.[a]

[a] This notation is often used even in spaces that are not uniquely geodesic.

Note that the inverse path of a geodesic that connects x to y is a geodesic that connects y to x so even in a uniquely geodesic space there are two geodesic segments with endpoints x and y if we do not specify the order of the endpoints.

Proposition 1.8. *Any inner product space is a uniquely geodesic metric space.*

Proof. Let V be an inner product space and let $x, y \in V$. We show that the geodesic segment $j|_{[0,\|x-y\|]}$ constructed in Example 1.6 is the only geodesic segment that connects x to y. [3]

Let $\tilde{\jmath}\colon [0, \|x-y\|] \to V$ be a geodesic segment with $\tilde{\jmath}(0) = x$ and $\tilde{\jmath}(\|x-y\|) = y$. If $0 < t < \|x-y\|$, then $\|x - \tilde{\jmath}(t)\| = \|\tilde{\jmath}(t) - \tilde{\jmath}(t)\| = t$ and $|\tilde{\jmath}(t) - y\| = 1 - t$. Thus, we have the equality

$$\|x - \tilde{\jmath}(t)\| + \|\tilde{\jmath}(t) - y\| = \|x - y\|$$

in the triangle inequality. We may assume for simplicity that $x = 0$. Squaring, the equation $\|y - \tilde{\jmath}(t)\| = \|y\| - \|\tilde{\jmath}(t)\|$, we get after simplification $(y \,|\, \tilde{\jmath}(t)) = \|y\|\|\tilde{\jmath}(t)\|$. The equality case of Cauchy's inequality implies that $y - x$ is in the linear segment from x to y. Thus, $\tilde{\jmath}(t) = j(t)$. □

Let X be a uniquely geodesic metric space. A nonempty subset $K \subset X$ is *convex* if $[x, y] \subset K$ for all $x, y \in K$.

A convex set $K \subset X$ is *strictly convex* if $[x, y] \cap \partial K \subset \{x, y\}$ for any $x, y \in K$.

Example 1.9. A normed space is uniquely geodesic if and only if its unit ball is strictly convex. See [BH, Prop. I.1.6]. Thus, for example the normed spaces $(\mathbb{R}^2, \|\cdot\|_p)$ with

$$\|x\|_p = \sqrt[p]{x_1^p + x_2^p}$$

are uniquely geodesic metric spaces if $1 < p < \infty$.

There are plenty of examples of metric spaces arising from normed spaces that are not uniquely geodesic. For example, the unit balls of the norms

$$\|x\|_\infty = \max\{|x_1|, |x_2|\}$$

and

$$\|x\|_\infty = |x_1| + |x_2|$$

in $\mathbb{R}^2$ are not strictly convex.

[3] up to replacing the interval of definition $[0, \|x-y\|]$ of the geodesic by $[a, a + \|x-y\|]$ for some $a \in \mathbb{R}$.

It is easy to check that, among many others, the mappings $j_1, j_2 \colon [0,1] \to (\mathbb{R}^2, d_\infty)$ defined by $j_1(t) = t(1,0)$ and

$$j_2(t) = \begin{cases} t(1,1), & \text{if } 0 \leqslant t \leqslant \frac{1}{2}, \\ (t, 1-t), & \text{if } \frac{1}{2} \leqslant t \leqslant 1 \end{cases}$$

are both geodesic segments in $(\mathbb{R}^2, d_\infty)$ connecting 0 to $(1,0)$.

Example 1.10. The Euclidean round circle $\mathbb{S}^1 \subset \mathbb{E}^2$ with the induced metric is not a geodesic metric space because the ambient space $\mathbb{E}^2$ is uniquely geodesic and the geodesic segment in $\mathbb{E}^2$ that connects any two distinct points of $\mathbb{S}^1$ intersects $\mathbb{S}^1$ only at these two points.

In certain contexts,[4] it is convenient to use mappings that multiply distances with a fixed constant.

Let X be a metric space, let $I \subset \mathbb{R}$ be a compact interval and let $K > 0$. A mapping $j \colon I \to X$ is an *affinely reparametrized geodesic segment* if $d(j(s), j(t)) = K\,|s-t|$ for all $s, t \in I$.

1.5 Metric graphs

Metric graphs and, in particular, metric trees are important examples in this course. The definition, based on see [Ser, Sect. 2.1], is somewhat involved.

Let $E\mathbb{X}$ and $V\mathbb{X}$ be two nonempty sets and let $o, t \colon E\mathbb{X} \to V\mathbb{X}$ and $\bar{\cdot} \colon E\mathbb{X} \to E\mathbb{X}$ be mappings that satisfy $\overline{e} \neq e$, $\overline{\overline{e}} = e$ and $o(\overline{e}) = t(e)$ for all $e \in E\mathbb{X}$. The quintuple $\mathbb{X} = (V\mathbb{X}, E\mathbb{X}, t, o, \bar{\cdot})$ is a *graph*.

The sets $E\mathbb{X}$ and $V\mathbb{X}$, called the *set of vertices* and the *set of edges* of $\mathbb{X}$.

The elements $o(e)$, $t(e)$ and $\overline{e}$ are called the *initial vertex*, the *terminal vertex* and the *opposite edge* of an edge $e \in E\mathbb{X}$. The quotient of $E\mathbb{X}$ by the equivalence relation induced by the involution $e \mapsto \overline{e}$ is called the set of *nonoriented edges* of $\mathbb{X}$.

The cardinality of the preimage $o^{-1}(v)$ is the *degree* $\deg v$ of the vertex $v \in V\mathbb{X}$. If $\deg \colon \mathbb{X} \to \mathbb{N}$ is a constant mapping, then $\mathbb{X}$ is a *regular graph*.

Note that we assume that the sets of vertices and edges are not empty but we make no further assumptions on the cardinalities of these sets. Often, graphs are defined in a different way, taking the set of nonoriented edges to be a set consisting of pairs of distinct vertices. The above definition allows for *loops* where $o(e) = t(e)$ for some edge e, and for multiple edges with equal initial and terminal vertices.

A graph is not a geometrical or topological object but one can associate natural spaces to it as follows. Recall that an equivalence relation $\sim$ is finer than $\simeq$ if $x \sim y$ implies $x \simeq y$.

[4] See the proof of Theorem **??** and the definition of metric convexity in section **??**.

The *topological realisation* $|\mathbb{X}|$ of a graph $\mathbb{X}$ is the topological space obtained from the disjoint union of the family $(I_e)_{e\in E\mathbb{X}}$ of closed unit intervals I_e and $V\mathbb{X}$ by the finest equivalence relation that identifies intervals corresponding to an edge I_e and its opposite edge $I_{\bar{e}}$ by the map $t \mapsto 1-t$ and identifies $0 \in I_e$ with $o(e) \in V\mathbb{X}$ for all $e \in E\mathbb{X}$.

More precisely, let $\coprod_{e\in E\mathbb{X}} I_e$ be the disjoint union of a family $(I_e)_{e\in E\mathbb{X}}$ of closed unit intervals I_e with the topology of the disjoint union.[5] Let $\sim$ be the equivalence relation in $\coprod_{e\in E\mathbb{X}}$ generated[6] by the identifications $(t, e) \sim (1-t, \bar{e})$ for all $t \in [0, 1]$ and all $e \in E\mathbb{X}$ and $(0, e) \sim (0, e')$ if and only if $o(e) = o(e') \in V\mathbb{X}$.

A graph is *connected* if its topological realisation is path connected as a topological space.

A connected graph is a *tree* if its topological realisation is uniquely arcwise connected.[a]

[a] Recall that the image of an injective path defined on a compact interval is an *arc*. A topological space X is *uniquely arcwise connected* if for any two distinct points $x, y \in X$ there is a unique arc $|\gamma|$ whose endpoints are x and y.

Example 1.11. (1) If $V\mathbb{X} = \mathbb{Z}$, $E\mathbb{X} = \mathbb{Z}\times\{0, 1\}$, $o(k, j) = k + j$, $t(k, j) = k + 1 - j$ and $\overline{(k, j)} = (k, 1-j)$, then it is easy to check using Figure 1.1 that the topological realization of $\mathbb{X}$ is homeomorphic to $\mathbb{E}^1$. If we replace $\mathbb{Z}$ by $\mathbb{N}$ in the construction, we obtain a graph $\mathbb{X}'$ whose topological realization is homeomorphic to $[0, \infty[$.

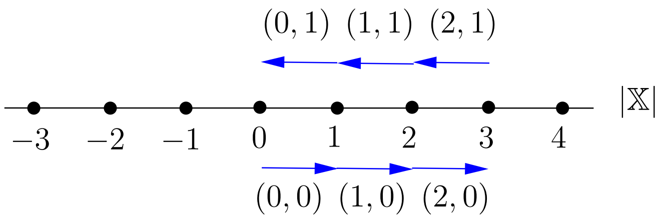


Figure 1.1 — $\mathbb{E}^1$ as a metric graph

(2) Let $A \neq \varnothing$ be any nonempty set and let $V\mathbb{X} = \{0\} \cup A$ and $E\mathbb{X} = A \times \{0, 1\}$. Let $o(a, 0) = 0 = t(a, 1)$ and $o(a, 1) = a = t(a, 0)$ for all $a \in A$ and define $\overline{(a, k)} = (a, 1-k)$ for all $a \in A$. If A is an infinite set, for example $A = \mathbb{S}^1$, the geometric realization of $\mathbb{X}$ is a *hedgehog space* that is not locally compact at the vertex 0.

(3) Often, we describe a graph more informally, for example by drawing a picture of the geometric realization or a sufficiently large part of it if the structure repeats itself in a reasonable manner.

A *metric graph* $(\mathbb{X}, \lambda)$ is a pair consisting of a connected graph $\mathbb{X}$ and *edge length map* $\lambda : E\mathbb{X} \to]0, +\infty]$ such that $\lambda(\bar{e}) = \lambda(e)$.

A *simplicial graph* $\mathbb{X}$ is a metric graph whose edge length map is constant equal to 1.

[5] This is the finest topology for which all the natural injections $I_e \hookrightarrow \coprod_{e\in E\mathbb{X}} I_e$ are continuous.

[6] The equivalence relation generated by a relation R on a set X is the smallest equivalence relation on X that contains R.

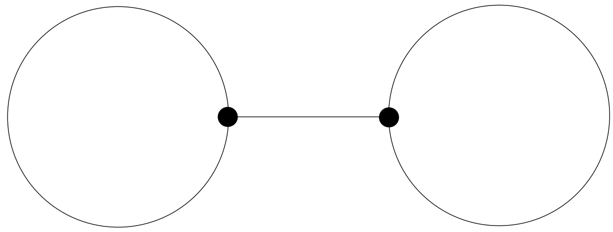


Figure 1.2 — The topological realization of a graph with two vertices and three undirected edges that has two loops.

Let $(\mathbb{X}, \lambda)$ be a metric graph and let $\pi_{\mathbb{X}} \colon \coprod_{e \in E\mathbb{X}} I_e \to |\mathbb{X}|$ be the canonical projection. A continuous mapping $c \colon [0,1] \to |\mathbb{X}|$ is a *piecewise linear path* if there is a subdivision $0 = t_0 < t_1 < \cdots < t_n = 1$ of $[0,1]$, a collection of edges $e_1, \ldots, e_n \in E\mathbb{X}$ and affine mappings $c_i \colon [t_{i-1}, t_i] \to I_{e_i}$ such that $c|_{[t_{i-1}, t_i]} = \pi_{\mathbb{X}} \circ c_i$. The *length* of c is

$$\ell_\lambda(c) = \sum_{i=1}^{n} |c_i(t_i) - c_i(t_{i-1})| \lambda(e_i) \,.$$

If $x, y \in |\mathbb{X}|$, let

$$\mathrm{PL}(x,y) = \big\{ c \colon [0,1] \to |\mathbb{X}| : c \text{ piecewise linear}, \; c(0) = x, \; c(1) = y \big\} \,.$$

We will now study a useful method to construct geodesic metric spaces from metric graphs. In some cases, this construction would not produce a metric space. Such problems do not arise if, for example, the edge length map has a positive lower bound as in the following result. In Exercise 1.2, we see there are examples of metric graphs that define metric spaces even if the edge lengths have no positive lower bound.

Proposition 1.12. *Let $(\mathbb{X}, \lambda)$ be a metric graph such that any two points in $|\mathbb{X}|$ can be connected by a piecewise linear path and λ has a positive lower bound. The expression*

$$d_\lambda(x,y) = \inf_{c \in \mathrm{PL}(x,y)} \ell_\lambda(c) \tag{1.1}$$

defines a metric on the topological realization of $\mathbb{X}$.

Proof. Exercise 1.3. □

Let $(\mathbb{X}, \lambda)$ be a metric graph such that d_λ is a metric.[a] The *geometric realisation* of $(\mathbb{X}, \lambda)$ is the metric space $(|\mathbb{X}|, d_\lambda)$.

[a] See equation 1.1.

From now on, we usually assume that the edge length map of a metric graph has a positive lower bound.

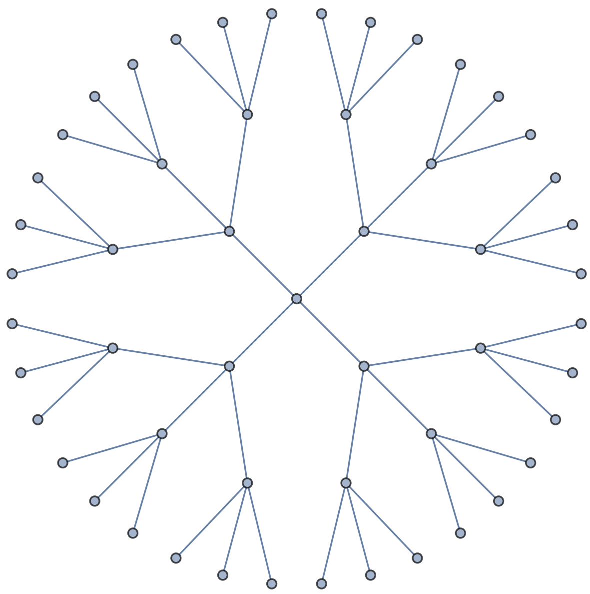


Figure 1.3 — Part of the geometric realization of a regular infinite simplicial tree such that the degree of each vertex is 4. Imagine all the branches extending indefinitely with the same branching at every vertex.

The metric space X determines $(\mathbb{X}, \lambda)$ up to subdivisions of edges, hence we will often not make a strict distinction between X and $(\mathbb{X}, \lambda)$. In particular, we identify $V\mathbb{X}$ with its image in X and we will refer to convex subsets of $(\mathbb{X}, \lambda)$ as convex subsets of X, etc.

A uniquely arcwise connected geodesic metric space is an $\mathbb{R}$-*tree*.

Example 1.13. (1) For any $x, y \in \mathbb{R}$, let

$$d_{\mathrm{SNCF}}(x,y) = \begin{cases} \|x-y\| & \text{if } x \text{ and } y \text{ are linearly dependent,} \\ \|x\| + \|y\| & \text{otherwise.} \end{cases}$$

The *French railroad space* $(\mathbb{R}^2, d_{\mathrm{SNCF}})$ is an $\mathbb{R}$-tree.[7] The closed unit ball $\overline{B}(0,1)$ of this space is a geometric realisation of the simplicial hedgehog space of Example 1.11(2).

(2) Figure 1.3 shows a simplicial tree.

1.6 Triangles

The definitions of negatively curved spaces in Chapters 6 and **??** are based on the properties of triangles and we will also treat classical properties of triangles in the Euclidean, spherical and hyperbolic spaces. A precise definition of this fundamental object is therefore in order:

[7]SNCF=Société nationale des chemins de fer français is the French national railroad company.

Let X be a metric space. A *triangle* in X is a triple $\Delta = \{j_1, j_2, j_3\}$ of geodesic segments such that the terminus of j_i is the origin of j_{i+1} with the index i considered cyclically mod 3.

The geodesic segments j_1, j_2 and j_3 are the *sides* of Δ.

A triangle Δ is *degenerate* if it is contained in the image of one of its sides.

The endpoints of the geodesic arcs j_1, j_2 and j_3 are the *vertices* of Δ.

A triangle Δ in a uniquely geodesic metric space is determined by its vertices but in general,[8] one has to specify the sides.

If X is a uniquely geodesic metric space and $x, y, z \in X$, then

$$\Delta(x, y, z) = \big\{[x, y], [y, x], [z, x]\big\}$$

is the *triangle with vertices* x, y and z.

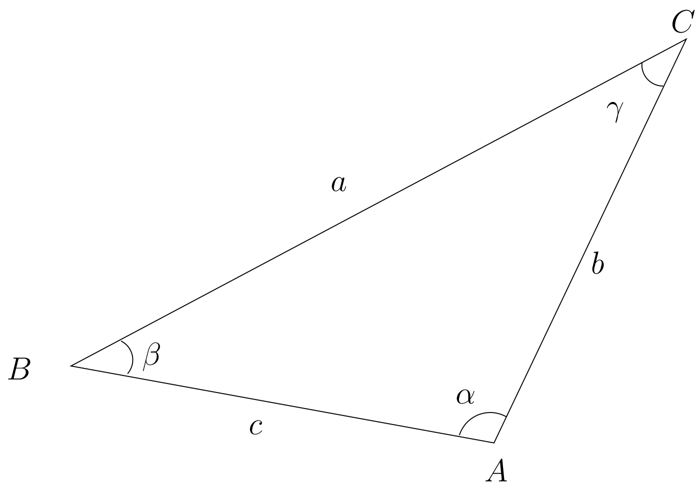


Figure 1.4 — A triangle in the Euclidean plane with a standard notation for the vertices, the lengths of the edges and the angles.

If X is a geodesic metric space and three points $A, B, C \in X$ are the vertices of a triangle, we denote the lengths of the sides with endpoints A and B, B and C and C and A, in the corresponding order, by c, a and b. If the angles at the vertices are defined,[a] the angles between the sides at the vertices A, B and C be α, β and γ. See Figure 1.4.

[a]for example in Chapters 2, 3 and 4

[8]See Example 1.9

Exercises

1.1. Prove Propositions 1.3 and 1.4.

1.2. Give examples of metric graphs $(\mathbb{X}_1, \lambda_1)$ and $(\mathbb{X}_2, \lambda_2)$ such that
(1) the edge length maps λ_1 and λ_2 do not have a positive lower bound.
(2) the geometric realization of $(\mathbb{X}_1, \lambda_1)$ is isometric with $\mathbb{E}^1$, and
(3) the geometric realization of $(\mathbb{X}_2, \lambda_2)$ is isometric with a bounded half-open interval.

1.3. Prove Proposition 1.12. Why do we assume that the length function has a positive lower bound?

1.4. Prove that $(\mathbb{R}^2, d_{\text{SNCF}})$ is not a proper metric space.[9] Describe the isometry group of $(\mathbb{R}^2, d_{\text{SNCF}})$.

1.5. For any $x, y \in \mathbb{R}^2$, let

$$d(x,y) = \begin{cases} |x_2| + |x_1 - y_1| + |y_2| & , \text{ if } x_1 \neq y_1, \\ |x_2 - y_2| & , \text{ if } x_1 = y_1, \end{cases}$$

(a) Prove that $(\mathbb{R}^2, d)$ is an $\mathbb{R}$-tree.
(b) Draw the sphere $\partial B(0,1)$ of $(\mathbb{R}^2, d)$. Is it compact or connected?

Let $[a,b] \subset \mathbb{R}$ be a compact interval. An ordered finite sequence

$$\sigma = (a = \sigma_0 < \sigma_1 < \cdots < \sigma_n = b)$$

is a *partition* of $[a,b]$. Let $\mathscr{P}_{a,b}$ be the *set of partitions* of $[a,b]$.

Let X be a metric space and let $\gamma\colon [a,b] \to X$ be a path. The *variation* of γ with respect to a partition $\sigma = (a = \sigma_0 < \sigma_1 < \cdots < \sigma_n = b)$ is

$$V_a^b(\gamma, \sigma) = \sum_{i=1}^{n} d(\gamma(\sigma_i), \gamma(\sigma_{i-1}))\,.$$

The *length* of γ is its *total variation*

$$\ell(\gamma) = V_a^b(\gamma) = \sup_{\sigma \in \mathscr{P}_{a,b}} V_a^b(\gamma, \sigma)\,.$$

1.6. Let X be a metric space and let $\gamma\colon [0,b] \to X$ be a geodesic segment.
(a) Compute the length of γ.
(b) Prove that γ is a shortest path from $\gamma(0)$ to $\gamma(b)$.

1.7. Fill in the details in Example 1.7.

[9] See Example 1.13 for the definition.

Chapter 2

Euclidean geometry

This chapter collects background information on Euclidean spaces. Most of this should be known in some form from linear algebra and elementary geometry.

2.1 Euclidean space

As we use various different structures on the space $\mathbb{R}^n$, it is convenient to have a fixed notation for the situation where we use the standard Euclidean structure. The notation $\mathbb{R}^n$ therefore does not carry the Euclidean structure, it is just the n-fold Cartesian product of $\mathbb{R}^n$, and we usually consider it with the standard structure of a vector space over $\mathbb{R}$.

Let us denote the *Euclidean inner product* of $\mathbb{R}^n$ by

$$(x \,|\, y) = \sum_{i=1}^{n} x_i y_i \,.$$

The *Euclidean norm* $\|x\| = \sqrt{(x|x)}$ defines the *Euclidean distance* $d(x, y) = \|x - y\|$. The triple $\mathbb{E}^n = (\mathbb{R}^n, (\cdot \,|\, \cdot), \| \cdot \|)$ is n-dimensional *Euclidean space.*

Proposition 2.1. *Euclidean space is a uniquely geodesic metric space.*

Proof. See Proposition 1.8 □

2.2 Euclidean triangles

The first two results are classical formulae that connect the side lengths and angles of triangles in Euclidean space.

Proposition 2.2 (The Euclidean law of cosines)**.** *The relation*

$$c^2 = a^2 + b^2 - 2ab\cos\gamma$$

holds for all triangles in $\mathbb{E}^n$*.*

Proof. The proof is linear algebra:

$$\begin{aligned} c^2 &= \|B - A\|^2 = \|B - C + C - A\|^2 = b^2 + 2(B - C \mid C - A) + a^2 \\ &= b^2 + 2(B - C \mid C - A) + a^2 = b^2 - 2ab\cos\gamma + a^2 \,. \end{aligned}$$

□

Proposition 2.3 (The Euclidean law of sines)**.** *The relation*

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

holds for all triangles in $\mathbb{E}^n$*.*

Proof. Exercise. □

The following result will be useful in Chapter **??** when we discuss comparison geometry and CAT(-1) spaces. The content is this: Given any three positive numbers that satisfy the conditions arising from the triangle inequality to be the sides of a triangle in a geodesic metric space, there is a triangle in $\mathbb{E}^2$ with precisely these side lengths.

Proposition 2.4. *Let* $a, b, c > 0$ *and assume that* $a + b \geqslant c$*,* $a + c \geqslant b$ *and* $b + c \geqslant a$*. There is a triangle in* $\mathbb{E}^2$ *with side lengths* a*,* b *and* c*.*

Proof. The inequality $a+b \geqslant c$ implies $\frac{a^2+b^2-c^2}{2ab} \geqslant -1$ and the inequality $a+c \geqslant b$ implies $\frac{a^2+b^2-c^2}{2ab} \leqslant 1$. Thus, we can solve the equation $c^2 = a^2 + b^2 - 2ab\cos\gamma$ to find $\gamma \in [0, \pi]$. Placing two segments of lengths a and b starting at 0 with the angle γ at the vertex 0 determines a triangle in $\mathbb{E}^2$. The Euclidean law of cosines implies that the length of the third edge is c. □

Proposition 2.5. *The sum of the angles of a triangle in* $\mathbb{E}^2$ *is* π*.*

Proof. There are many different proofs, here is one that uses complex numbers: Note that

$$\frac{C-A}{B-A} = \left\|\frac{C-A}{B-A}\right\| e^{i\alpha} \,, \quad \frac{A-B}{C-B} = \left\|\frac{A-B}{C-B}\right\| e^{i\beta} \,, \quad \frac{B-C}{A-C} = \left\|\frac{B-C}{A-C}\right\| e^{i\gamma} \,.$$

The product of the left sides of these equations is -1, and therefore, $e^{i(\alpha+\beta+\gamma)} = -1$. Thus, $\alpha + \beta + \gamma = \pi + k\,2\pi$ for some $k \in \mathbb{Z}$. As $0 \leqslant \alpha, \beta, \gamma \leqslant \pi$ and at most one of them can be π, we get the claim. □

2.3 Isometries of $\mathbb{E}^n$

We will now study the isometries of Euclidean space more closely.

The *(Euclidean) orthogonal group* of dimension n is

$$\begin{aligned} \mathrm{O}(n) &= \{A \in \mathrm{GL}_n(\mathbb{R}) : (Ax \mid Ay) = (x \mid y) \ \text{ for all } x, y \in \mathbb{E}^n\} \\ &= \{A \in \mathrm{GL}_n(\mathbb{R}) : A^T A = I_n\} \,. \end{aligned}$$

Recall the following basic result from linear algebra:

Lemma 2.6. *An $n \times n$-matrix $A = (a_1, \dots, a_n)$ is in $\mathrm{O}(n)$ if and only if the vectors $a_1, \dots, a_n$ form an orthonormal basis of $\mathbb{E}^n$.* □

It is easy to check that elements of $\mathrm{O}(n)$ give isometries on $\mathbb{E}^n$ for any $n \in \mathbb{N}$: Let $A \in \mathrm{O}(n)$ and let $x, y \in \mathbb{E}^n$. Now

$$\begin{aligned} d(Ax, Ay)^2 &= (Ax - Ay \,|\, Ax - Ay) = (A(x-y) \,|\, A(x-y)) \\ &= (A^T A(x-y) \,|\, x-y) = (x-y \,|\, x-y) \\ &= d(x-y)^2 \,. \end{aligned}$$

For any $b \in \mathbb{R}^n$, let $t_b(x) = x + b$ be the translation by b. Again, it is easy to see that translations are isometries of $\mathbb{E}^n$. The *translation group* is

$$\mathrm{T}(n) = \{t_b : b \in \mathbb{R}^n\} \,.$$

Orthogonal maps and translations generate the *Euclidean group*

$$\mathrm{E}(n) = \{x \mapsto Ax + b : A \in \mathrm{O}(n), b \in \mathbb{R}^n\} = T(n)O(n)$$

which consists of isometries of $\mathbb{E}^n$.

Proposition 2.7. $\mathrm{E}(n)$ *acts transitively by isometries on $\mathbb{E}^n$. In particular,* $\mathrm{Isom}(\mathbb{E}^n)$ *acts transitively on $\mathbb{E}^n$.*

Proof. The Euclidean group of $\mathbb{E}^n$ contains the group of translations $\mathrm{T}(n)$ as a subgroup. This subgroup acts transitively because for any $x, y \in \mathbb{R}^n$, we have $T_{y-x}(x) = y$. □

Next, we want to prove that all isometries of Euclidean space $\mathbb{E}^n$ are elements of the Euclidean group.

Theorem 2.8. $\mathrm{Isom}(\mathbb{E}^n) = \mathrm{E}(n)$.

The proof of this theorem and the introduction of the tools needed in the proof takes up the rest of this section.

An *affine hyperplane* of $\mathbb{E}^n$ is a subset of the form

$$H = H(P, u) = P + u^\perp \,,$$

where $P, u \in \mathbb{E}^n$ and $\|u\| = 1$. The *reflection* in H is the map

$$r_H(x) = x - 2(x - P \,|\, u)u \,.$$

Lemma 2.9. *The definition of r_H is independent of the choice of $P \in H$.*

Proof. If $P, Q \in H$, then $P - Q \in u^\perp$. Thus,

$$x - 2(x - P \,|\, u)u = x - 2(x - P \,|\, u)u - 2(P - Q \,|\, u)u = x - 2(x - Q \,|\, u)u \,. \tag{2.1}$$

□

Reflections are very useful isometries, the following results give some of their basic properties.

Proposition 2.10. *Let H be an hyperplane in $\mathbb{E}^n$. Then*
(1) $r_H \circ r_H$ is the identity.
(2) $r_H \in \mathrm{E}(n)$. In particular, r_H is an isometry, and if $0 \in H$, then $r_H \in \mathrm{O}(n)$.
(3) $d(r_H(x), y) = d(x, y)$ for all $x \in \mathbb{E}^n$ and all $y \in H$.
(4) The fixed point set of r_H is H.

Proof. We will prove (3) and leave the rest as exercises. Let $x \in \mathbb{E}^n$ and $y \in H$. We have $r_H(x) = x - 2(x - y \,|\, u)u$, which implies

$$\begin{aligned} d(r_H(x), y)^2 &= (r_H(x) - y \,|\, r_H(x) - y) = (x - y - 2(x - y \,|\, u)u \,|\, x - y - 2(x - y \,|\, u)u) \\ &= (x - y \,|\, x - y) - 4(x - y \,|\, (x - y \,|\, u)u) + 4\big((x - y \,|\, u)u \,\big|\, (x - y \,|\, u)u\big) \\ &= (x - y \,|\, x - y) = d(x, y)^2 \,. \end{aligned}$$ □

The *bisector* of two distinct points p and q in $\mathbb{E}^n$ is the affine hyperplane

$$\mathrm{bis}(p, q) = \{x \in \mathbb{E}^n : d(x, p) = d(x, q)\} \,.$$

Lemma 2.11. *If $p, q \in \mathbb{E}^n$, $p \neq q$, then*

$$\mathrm{bis}(p, q) = \frac{p + q}{2} + (p - q)^{\perp}.$$

Proof. Exercise. □

Proposition 2.12. *(1) If $r_H(x) = y$ and $x \notin H$, then $H = \mathrm{bis}(x, y)$.*
(2) If $p, q \in \mathbb{E}^n$, $p \neq q$, then $r_{\mathrm{bis}(p,q)}(p) = q$.
(3) Let $\phi \in \mathrm{Isom}(\mathbb{E}^n)$, $\phi \neq \mathrm{id}$. If $a \in \mathbb{E}^n$, $\phi(a) \neq a$, then the fixed points of ϕ are contained in $\mathrm{bis}(a, \phi(a))$.
(4) Let $\phi \in \mathrm{Isom}(\mathbb{E}^n)$, $\phi \neq \mathrm{id}$. If H is a hyperplane such that $\phi|_H$ is the identity, then $\phi = r_H$.

Proof. (1) follows from Proposition 2.10(3).
(2) From the definitions we get

$$r_{\mathrm{bis}(p,q)}(p) = p - 2\big(p - \frac{p + q}{2} \,\big|\, p - q\big) \frac{p - q}{\|p - q\|^2} = q \,.$$

(3) If $\phi(b) = b$, then $d(a, b) = d(\phi(a), \phi(b)) = d(\phi(a), b)$, so that $b \in \mathrm{bis}(a, \phi(a))$.
(4) Let $a \notin H$ be a point that is not fixed by ϕ. Claim (3) implies that H is contained in $\mathrm{bis}(a, \phi(a))$. As H and $\mathrm{bis}(a, \phi(a))$ are both hyperplanes, we have $H = \mathrm{bis}(a, \phi(a))$. Thus, by Claim (2), $r_H(a) = \phi(a)$. But this holds for all $a \notin H$. As $r_H|_H = \phi_H = \mathrm{id}_H$, we have $\phi = r_H$. □

The idea of the proof of Theorem 2.8 is to show that each isometry of $\mathbb{E}^n$ is the composition of reflections in affine hyperplanes. In order to do this, we show that the isometry group has a stronger transitivity property than what was noted above.

Proposition 2.13. *Let $p_1, p_2, \ldots, p_k, q_1, q_2, \ldots, q_k \in \mathbb{E}^n$ be points that satisfy*

$$d(p_i, p_j) = d(q_i, q_j)$$

for all $i, j \in \{1, 2, \ldots, k\}$. Then, there is an isometry $\phi \in \mathrm{E}(n) \leqslant \mathrm{Isom}(\mathbb{E}^n)$ such that $\phi(p_i) = q_i$ for all $i \in \{1, 2, \ldots, k\}$. Furthermore, the isometry ϕ is the composition of at most k reflections in affine hyperplanes.

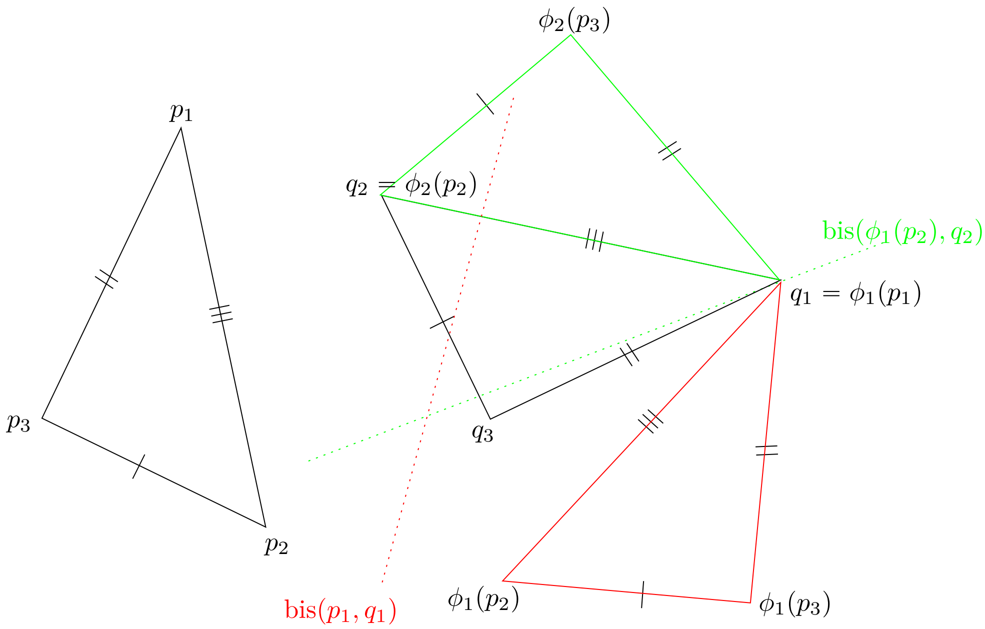


Figure 2.1 —

Proof. We construct the isometry by induction. If $p_1 = q_1$, let ϕ_1 be the identity, otherwise, let ϕ_1 be the reflection in the bisector of p_1 and q_1. Let $m > 1$ and assume that there is an isometry ϕ_m such that $\phi_m(p_i) = q_i$ for all $i \in \{1, 2, \ldots, m\}$, which is the composition of at most m reflections. The mapping ϕ is in $\mathrm{E}(n)$ by Proposition 2.10.

Assume that $\phi_m(p_{m+1}) \neq q_{m+1}$. Now, $q_1, \ldots q_m \in \mathrm{bis}(\phi_m(p_{m+1}), q_{m+1})$ because for each $1 \leqslant i \leqslant m$, we have

$$d(q_i, \phi_m(p_{m+1})) = d(\phi_m(p_i), \phi_m(p_{m+1})) = d(p_i, p_{m+1}) = d(q_i, q_{m+1})\,.$$

Thus, the map

$$\phi_{m+1} = r_{\mathrm{bis}(\phi_m(p_{m+1}), q_{m+1})} \circ \phi_m$$

satisfies $\phi_{m+1}(p_i) = q_i$ for all $1 \leqslant i \leqslant m + 1$. □

Corollary 2.14. *If T and T' are two triangles in $\mathbb{E}^n$ with equal side lengths, then there is an isometry ϕ of $\mathbb{E}^n$ such that $\phi(T) = T'$.* □

Corollary 2.15. *Any isometry of $\mathbb{E}^n$ can be represented as the composition of at most $n + 1$ reflections.*

Proof. Let $\phi \in \text{Isom}(\mathbb{E}^n)$. Proposition 2.13 implies that there is an isometry $\phi_0 \in \text{E}(n)$ such that $\phi_0(\phi(e_i)) = e_i$ for all $1 \leqslant i \leqslant n$ and $\phi_0(\phi(0)) = 0$. The set of fixed points of $\phi_0 \circ \phi$ contains the points $0, e_1, \ldots, e_n$. In particular, the fixed point set of $\phi_0 \circ \phi$ is not contained in any affine hyperplane. Proposition 2.12(3) implies that $\phi_0 \circ \phi = \text{id}$. Thus, $\phi = \phi_0^{-1}$. □

Proof of Theorem 2.8. The elements of $\text{E}(n)$ are isometries by Proposition 2.7. The opposite inclusion follows from Corollary 2.15 and Proposition 2.10(2). □

Proposition 2.16. *The stabiliser in* $\text{Isom}(\mathbb{E}^n)$ *of any point* $x \in \mathbb{E}^n$ *is isomorphic to* $\text{O}(n)$*. An isometry* F *of* $\mathbb{E}^n$ *fixes* $b \in \mathbb{E}^n$ *if and only if there is an orthogonal linear map* F_0 *such that* $F = T_b \circ F_0 \circ T_b^{-1}$*.*

Proof. An element of $\text{E}(n)$ fixes the origin if and only if it is an orthogonal linear transformation. Thus the claim holds for 0. If $b \in \mathbb{E}^n - \{0\}$ and $F \in \text{Stab}\, b$, then $T_b^{-1} \circ F \circ T_b \in \text{O}(n)$ and for any $A \in \text{O}(n)$, $\quad T_b \circ A \circ T_b^{-1} \in \text{fix}\, b$ □

Proposition 2.17. *For each affine k-plane P, there is an isometry* $\phi \in \text{Isom}(\mathbb{E}^n)$ *such that*

$$\phi(P) = \{x \in \mathbb{E}^n : x^{k+1} = x^{k+2} = \cdots = x^n = 0\}\,.$$

Each affine k-plane of $\mathbb{E}^n$ *is isometric with* $\mathbb{E}^k$*.*

Proof. This is a direct generalisation of Proposition 2.7. The details are left as an exercise. □

Exercises

2.1. Prove Proposition 2.3.

2.2. Let $x_0 \in \mathbb{E}^n$ and let $u, v \in \mathbb{S}^n$. Let $F \colon \mathbb{E}^n \to \mathbb{E}^n$ be an isometry.

(1) Show that $F \circ j_{x_0,u}$ and $F \circ j_{x_0,v}$ are geodesic lines.

(2) Show that $F \circ j_{x_0,u}$ and $F \circ j_{x_0,v}$ intersect and that the angle of intersection is the same as for $j_{x_0,u}$ and $j_{x_0,v}$.

2.3. Find an isometry F of $\mathbb{E}^2$ such that $F(0) = (1,0)$, $F(1,0) = (1,1)$ and $F(0,1) = (2,0)$.

2.4. Let $H(0,u)$ be a line in $\mathbb{E}^2$ that forms an angle $\frac{\phi}{2}$with the positiv x_1-axis. Let r_u be the reflection in $H(0,u)$.

(1) Compute the matrix of r_u in the standard basis.

(2) Let $u_1, u_2 \in \mathbb{S}^1$. Compute the matrix of $r_{u_2} \circ r_{u_1}$ in the standard basis.

(3) Write the rotation by $\frac{\pi}{2}$ as the composition of two reflections.

2.5. Prove the remaining parts of Proposition 2.10.

2.6. Prove Lemma 2.11.

2.7. Prove Proposition 2.17.

Chapter 3

Spherical geometry

3.1 The sphere

The unit sphere in $(n-1)$-dimensional Euclidean space is

$$\mathbb{S}^n = \{x \in \mathbb{E}^{n+1} : \|x\| = 1\} .$$

Let us show that the *angle distance*

$$d_{\mathbb{S}^n}(x, y) = \arccos(x \mid y) \in [0, \pi] \tag{3.1}$$

is a metric. In order to do this, we will use the analog of the Euclidean law of cosines, but first we have to define the objects that are studied in spherical geometry.

Each 2-dimensional linear subspace $T \subset \mathbb{R}^{n+1}$ intersects $\mathbb{S}^n$ in a *great circle*. If $A \in \mathbb{S}^n$ and $u \in \mathbb{S}^n$ is orthogonal to A ($u \in A^\perp$), then the path $j_{A,u} \colon \mathbb{R} \to \mathbb{S}^n$,

$$j_{A,u}(t) = A \cos t + u \sin t ,$$

parametrises the great circle $\langle A, u\rangle \cap \mathbb{S}^n$, where $\langle A, u\rangle$ is the linear span of A and u. The vectors A and u are linearly independent, so $\langle A, u\rangle$ is a 2- plane.

Lemma 3.1. *If $d_{\mathbb{S}^n}$ is a metric, then $j_{A,u}$ is a locally geodesic line.*

Proof. Observe that as A and u are unit vectors such that $(A|u) = 0$, we have

$$\begin{aligned}(j_{A,u}(s) \mid j_{A,u}(t)) &= (A \cos s + u \sin s \mid A \cos t + u \sin t) \\ &= \|A\|^2 \cos s \cos t + (\cos s \sin t + \sin s \cos t)(A \mid u) + \sin s \sin t \|u\|^2 \\ &= \cos s \cos t + \sin s \sin t = \cos(s - t) .\end{aligned} \tag{3.2}$$

Thus, if $||s - t| \leqslant \pi$, we have

$$d(j_{A,u}(s), j_{A,u}(t)) = \arccos(j_{A,u}(s) \mid j_{A,u}(t)) = \arccos \cos(s - t) = |s - t| ,$$

which implies that the restriction of $j_{A,u}$ to any segment of length less than π is an isometric embedding. □

Note that the computation (3.2) applied with $s = t$ implies that the image of the mapping $j_{A,u}$ is contained in $\mathbb{S}^1$.

If $A, B \in \mathbb{S}^n$ such that $B \neq \pm A$, then there is a unique plane that contains both points. Thus, there is unique great circle that contains A and B, in the remaining cases, there are infinitely many such planes. The great circle is parametrised by the map $j_{A,u}$, with

$$u = \frac{B - (B \,|\, A)A}{\|B - (B \,|\, A)A\|} = \frac{B - (A \,|\, B)A}{\sqrt{1 - (A \,|\, B)^2}}\,. \tag{3.3}$$

Now $j(0) = A$ and $j(d(A, B)) = B$.

If $B = -A$, then there are infinitely many great circles through A and B: the map $j_{A,u}$ parametrises a great circle through A and B for any $u \in A^\perp$.

We call the restriction of any $j_{A,u}$ as above to any compact interval $[0, s]$ a *spherical segment*, and u is called the *direction* of $j_{A,u}$. Once we have proved that d is a metric, it is immediate that a spherical segment is a geodesic segment.

Our proof showing that the expression (3.1) defines a metric is based on the spherical law of cosines.

A triangle in $\mathbb{S}^n$ is defined as in the Euclidean case but now the sides of the triangle are the spherical segments connecting the vertices.

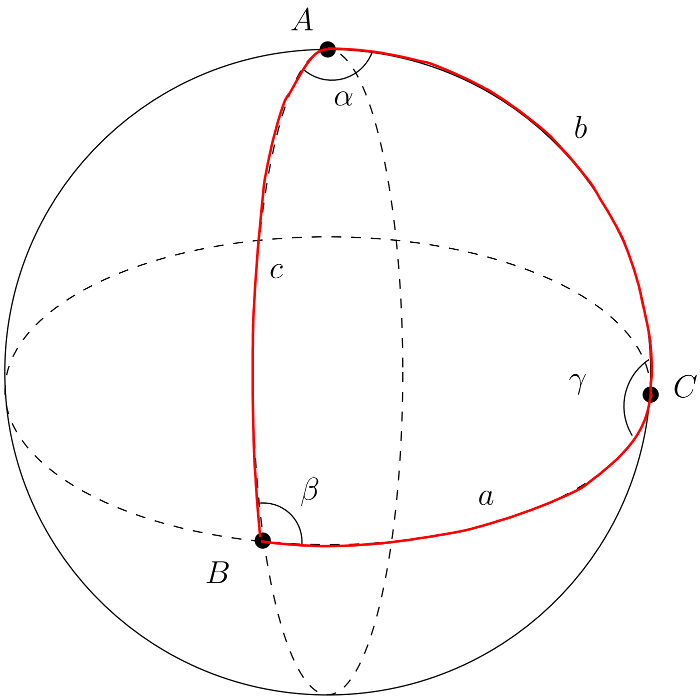


Figure 3.1 — A triangle in $\mathbb{S}^2$.

Let $j_{C,u}([0, d(C, A)])$ be the side between C and A, and let $j_{C,v}([0, d(C, B)])v$ be the side between C and B. The angle between the sides $j_{C,u}([0, d(C, A)])$ and $j_{C,v}([0, d(C, B)])$ is $\arccos(u \,|\, v)$, which is the angle at A between the segments $j_{C,u}([0, d(C, A)])$ and $j_{C,v}([0, d(C, B)])$ in the ambient space $\mathbb{E}^{n+1}$.

Now we can state and prove

Proposition 3.2 (The spherical law of cosines)**.** *In spherical geometry, the relation*

$$\cos c = \cos a \cos b + \sin a \sin b \cos \gamma$$

holds for any triangle.

Proof. Let u and v be the initial tangent vectors of the spherical segments $j_{C,u}$ from C to A and $j_{C,v}$ from C to B. As u and v are orthogonal to C, we have

$$\begin{aligned}\cos c &= (A \mid B) = (\cos(b)C + \sin(b)u \mid \cos(a)C + \sin(a)v)\\ &= \cos(a)\cos(b) + \sin(b)\sin(a)(u \mid v)\,.\end{aligned}$$ □

Proposition 3.3. *The angle distance is a metric on $\mathbb{S}^n$.*

Proof. Clearly, the triangle inequality is the only property that needs to be checked to show that the angle metric is a metric. Let $A, B, C \in \mathbb{S}^n$ be three distinct points and use the notation introduced above for triangles. The function

$$\gamma \mapsto f(\gamma) = \cos(a)\cos(b) + \sin(a)\sin(b)\cos(\gamma)$$

is strictly decreasing on the interval $[0, \pi]$, and

$$f(\pi) = \cos(a)\cos(b) - \sin(b)\sin(a) = \cos(a + b)\,.$$

Thus, the law of cosines implies that for all $\gamma \in [0, \pi]$, we have

$$\cos(c) = \cos(a)\cos(b) + \sin(a)\sin(b)\cos(\gamma) \geqslant \cos(a + b), \tag{3.4}$$

which implies $c \leqslant a + b$. Thus, the angle distance is a metric. □

Note that the inequality (3.4) is strict unless $\gamma = \pi$. This also implies that for triangles that are not completely contained in a great circle,

$$c < a + b < 2\pi - c\,. \tag{3.5}$$

We return to this observation in Section 3.4.

Theorem 3.4. *$(\mathbb{S}^n, d_{\mathbb{S}^n})$ is a geodesic metric space. If $0 < d_{\mathbb{S}^n}(A, B) < \pi$,*[1] *then there is a unique geodesic segment from A to B.*

Proof. If $x, y \in \mathbb{S}$ with $y \neq \pm x$, then, by Lemma 3.1, the spherical segment with direction given by the equation (3.3) is a geodesic segment that connects x to y. If the points x and y are antipodal, then it is immediate from the expression of the spherical segment that $j_{x,u}(\pi) = -x$ for all $u \in x^{\perp}$ with $\|u\| = 1$. Thus, in this case there are infinitely many geodesic segments connecting x to y.

If j is a geodesic segment connecting A to B, then any C in $j([0, d(A, B)])$ satisfies

$$d_{\mathbb{S}^n}(A, C) + d_{\mathbb{S}^n}(C, B) = d_{\mathbb{S}^n}(A, B)$$

by definition of a geodesic segment. In the proof of Proposition 3.3, we saw that equality holds in the triangle inequality if and only if $\gamma = \pi$. In this case, all the points A, B and C lie on the same great circle and C is contained in the side connecting A to B. Thus, the spherical segments are the only geodesic segments connecting A and B. If $A \neq \pm B$, then there is exactly one 2-plane containing both points. This proves the second claim. □

Note that the sphere has no geodesic lines or rays because the diameter of the sphere is π.

[1]This condition is equivalent with $B \neq \pm A$.

3.2 More on cosine and sine laws

The law of cosines implies that a triangle in $\mathbb{E}^n$ or $\mathbb{S}^n$ is uniquely determined up to an isometry of the space, if the lengths of the three sides are known. In Euclidean space the angles are given by

$$\cos\gamma = \frac{a^2 + b^2 - c^2}{2ab}$$

and the corresponding equations for α and β obtained by permuting the sides and angles, and in the sphere we have

$$\cos\gamma = \frac{\cos c - \cos a \cos b}{\sin a \sin b}\,.$$

In Euclidean space, the three angles of a triangle do not determine the triangle uniquely because dilations of $\mathbb{E}^n$ preserve angles. In $\mathbb{S}^n$ the angles determine a triangle uniquely up to isomorphism. This is the content of

Proposition 3.5 (The second spherical law of cosines)**.** *In spherical geometry, the relation*

$$\cos c = \frac{\cos\alpha\cos\beta + \cos\gamma}{\sin\alpha\sin\beta}$$

holds for any triangle.

Proof. This formula follows from the first law of cosines by manipulation. The first law of cosines implies

$$\sin^2\gamma = 1 - \cos^2\gamma = \frac{1 + 2\cos a\cos b\cos c - (\cos^2 + \cos^2 b + \cos^2 c)}{\sin^2 a \sin^2 b} = \frac{D}{\sin^2 a\sin^2 b}\,,$$

and D is symmetric in a, b and c. Thus, using the law of cosines, we get

$$\frac{\cos\alpha\cos\beta + \cos\gamma}{\sin\alpha\sin\beta} = \frac{\frac{\cos a - \cos b\cos c}{\sin b\sin c}\frac{\cos b - \cos a\cos c}{\sin a\sin c} + \frac{\cos c - \cos a\cos b}{\sin a\sin b}}{\frac{D}{\sin a\sin b\sin^2 c}} = \cos c\,.$$

□

Spherical geometry even has its own sine law

Proposition 3.6 (The spherical law of sines)**.** *In spherical geometry, the relation*

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

holds for any triangle.

Proof. In the proof of the second law of cosines we saw that he first law of cosines implies that

$$\left(\frac{\sin c}{\sin\gamma}\right)^2 = \frac{\sin^2 a\sin^2 b\sin^2 c}{D}\,.$$

The claim follows because this expression is symmetric in a, b and c. □

3.3 Isometries of $\mathbb{S}^n$

Proposition 3.7. *The orthogonal group* $\mathrm{O}(n+1)$ *acts transitively by isometries on* $\mathbb{S}^n$*. In particular,* $\mathrm{Isom}(\mathbb{S}^n)$ *acts transitively on* $\mathbb{S}^n$*.*

Proof. Let $A \in \mathrm{O}(n+1)$ and let $x \in \mathbb{E}^{n+1}$. By definition of orthogonal matrices, we have $\|Ax\|^2 = (Ax \,|\, Ax) = \|x\|^2$. Thus, A defines a bijection of the sphere $\mathbb{S}^n$ to itself. Furthermore, for any $x, y \in \mathbb{S}^{n+1}$, again by the definition of orthogonal matrices,

$$\cos d_{\mathbb{S}^n}(Ax, Ay) = (Ax \,|\, Ay) = (x \,|\, y) = \cos d_{\mathbb{S}^n}(x, y)\,,$$

which implies that the above mapping is an isometry.

Transitivity follows from the fact that any element of $\mathbb{S}^n$ can be taken as the first element of an orthonormal basis of $\mathbb{E}^n$ or, equivalently, as the first column of an orthogonal matrix. □

Theorem 3.8. $\mathrm{Isom}(\mathbb{S}^n) = \mathrm{O}(n+1)$

Proof. The claim follows from Proposition 3.7 and Corollary 3.13 and Proposition 3.9 below in the same way as its Euclidean analog, Theorem 2.8, was proven. □

Let H_0 be a linear hyperplane in $\mathbb{E}^n$. The intersection $H = H_0 \cap \mathbb{S}^n$ is a *hyperplane* of $\mathbb{S}^n$.

The *reflection* r_H in H is the restriction of the reflection in H_0 to the sphere: $r_H = r_{H_0}|_{\mathbb{S}^n}$.

Note that each hyperplane of $\mathbb{S}^n$ is isometric with $\mathbb{S}^{n-1}$ and that, by Propositions 2.10(2) and 3.7, the image of $r_{H_0}|_{\mathbb{S}^n}$ is contained in $\mathbb{S}^n$.

Proposition 3.9. *Let* H *be an hyperplane in* $\mathbb{S}^n$*. Then*

(1) $r_H \circ r_H$ *is the identity.*

(2) $r_H \in \mathrm{O}(n+1)$*. In particular,* r_H *is an isometry of* $\mathbb{S}^n$*.*

(3) $d_{\mathbb{S}^n}(r_H(x), y) = d_{\mathbb{S}^n}(x, y)$ *for all* $x \in \mathbb{S}^n$ *and all* $y \in H$*.*

(4) The fixed point set of r_H *is* H*.*

Proof. Claims (1), (2) and (4) are direct consequences of Proposition 2.10. Claim (3) is Exercise 3.1. □

The *bisector* of two distinct points $p, q \in \mathbb{S}^n$ is

$$\mathrm{bis}(p, q) = \{x \in \mathbb{S}^n : d_{\mathbb{S}^n}(x, p) = d_{\mathbb{S}^n}(x, q)\}\,.$$

Lemma 3.10. *Let* $p, q \in \mathbb{S}^n$*,* $p \neq q$*. Then* $\mathrm{bis}(p, q) = (p - q)^{\perp} \cap \mathbb{S}^n$*. In particular, the bisector is a hyperplane, it is the intersection of the Euclidean bisector of* p *and* p *with the* $\mathbb{S}^n$*.*

Proof. The points $p, q, x \in \mathbb{S}^n$ satisfy $d_{\mathbb{S}^n}(x, p) = d_{\mathbb{S}^n}(x, q)$ if and only if $(p \,|\, x) = (q \,|\, x)$, which is equivalent with $(p - q \,|\, x) = 0$. □

Proposition 3.11. *Let $x, y \in \mathbb{S}^n$ and let H be a hyperplane of $\mathbb{S}^n$.*
(1) If $r_H(x) = y$ and $x \notin H$, then $H = \mathrm{bis}(x, y)$.
(2) If $p, q \in \mathbb{S}^n$, $p \neq q$, then $r_{\mathrm{bis}(p,q)}(p) = q$.
(3) Let $\phi \in \mathrm{Isom}(\mathbb{S}^n)$, $\phi \neq \mathrm{id}$. If $a \in \mathbb{S}^n$, $\phi(a) \neq a$, then the fixed points of ϕ are contained in $\mathrm{bis}(a, \phi(a))$.
(4) Let $\phi \in \mathrm{Isom}(\mathbb{S}^n)$, $\phi \neq \mathrm{id}$. If H is a hyperplane such that $\phi|_H$ is the identity, then $\phi = r_H$.

Proof. (1) follows from Proposition 3.9(3).
(2) Using the definitions and the fact that $\frac{p+q}{2}$ is in the Euclidean bisector of p and q, we get

$$r_{\mathrm{bis}(p,q)}(p) = p - 2\big(p - \frac{p+q}{2} \,\big|\, p - q\big) \frac{p-q}{\|p-q\|^2} = q\,.$$

The proofs of (3) and (4) are formally the same as in the Euclidean case. □

We leave it as an exercise to check that the following result is proved in the same way as their Euclidean counterparts.

Proposition 3.12. *Let $p_1, p_2, \dots, p_k, q_1, q_2, \dots, q_k \in \mathbb{S}^n$ be points that satisfy*

$$d(p_i, p_j) = d(q_i, q_j)$$

for all $i, j \in \{1, 2, \dots, k\}$. Then, there is an isometry $\phi \in \mathrm{Isom}(\mathbb{S}^n)$ such that $\phi(p_i) = q_i$ for all $i \in \{1, 2, \dots, k\}$. □

Corollary 3.13. *Any isometry of $\mathbb{S}^n$ can be represented as the composition of at most $n + 1$ reflections.* □

Proposition 3.14. *The stabilizer in $\mathrm{Isom}(\mathbb{S}^n)$ of any point $x \in \mathbb{S}^n$ is isomorphic to $\mathrm{O}(n)$.*

Proof. The north pole e_{n+1} is stabilized by the subgroup of $\mathrm{O}(n)$ that consists of block diagonal matrices $\mathrm{diag}(A, 1)$, where $A \in \mathrm{O}(n)$. Proposition 3.7 implies the claim as in the Euclidean case, see Proposition 2.16. □

Proposition 3.15. *Each k-plane of $\mathbb{S}^n$ is isometric with $\mathbb{S}^k$. For each k-plane P, there is an isometry $\phi \in \mathrm{Isom}(\mathbb{S}^n)$ such that*

$$\phi(P) = \{x \in \mathbb{S}^n : x^{k+2} = x^{k+3} = \dots = x^{n+1} = 0\}\,.$$

Proof. The proof is similar to that of the Euclidean analog, Proposition 2.17, Exercise 3.6 □

3.4 Triangles in the sphere

In this section, we prove among other results that the sum of the angles of a nondegenerate triangle in $\mathbb{S}^2$ is greater than π. In order to do this, we introduce the polar triangle of a spherical triangle.

Let $A, B, C \in \mathbb{S}^2$ be points that do not all lie on the same great circle, and let Δ be the triangle with vertices A, B and C. The *polar points* $A^*, B^*, C* \in \mathbb{S}^2$ of A, B and C are the unique points that satisfy the conditions

$$\begin{aligned} (A^* \mid B) =&0 = (A^* \mid C), \quad (A^* \mid A) > 0 \\ (B^* \mid C) =&0 = (B^* \mid A), \quad (B^* \mid B) > 0 \\ (C^* \mid A) =&0 = (C^* \mid B), \quad (C^* \mid C) > 0\,. \end{aligned} \tag{3.6}$$

The triangle Δ^* with vertices A^*, B^* and C^* is the *polar triangle* of Δ. Let a^*, b^* and c^* be the side lengths and let α^*, β^* and γ^* be the angles of $(ABC)^*$.

Geometrically, for each vertex of the triangle, the dual vertex is the intersection point of the line orhogonal to the plane that contains the other two vertices, on the same side of the plane as the original vertex.

Lemma 3.16. *The polar points of the vertices of a nondegenerate triangle Δ in $\mathbb{S}^2$ are linearly independent and $(\Delta^*)^* = \Delta$.*

Proof. Exercise 3.7. □

Proposition 3.17. *Let ABC be a triangle in $\mathbb{S}^2$ such that the vertices do not all lie on the same great circle. Then*

$$a + \alpha^* = b + \beta^* = c + \gamma^* = a^* + \alpha = b^* + \beta = c^* + \gamma = \pi\,.$$

Proof. The situation is completely symmetric so it suffices to prove $a + \alpha^* = \pi$. Let $u, v \in A^\perp = \langle B^*, C^* \rangle$ be the directions of the edges AB and AC, respectively. Recall that $(u \mid v) = \cos\alpha$ and $(B^* \mid C^*) = \cos a^*$.

Now, $u \in \langle A, B \rangle$ implies that $(u \mid C^*) = 0$ and similarly we have $(v \mid B^*) = 0$. Furthermore,

$$(u \mid B^*) = \left(\frac{B - (B \mid A)A}{\|B - (B \mid A)A\|} \,\Big|\, B^* \right) = \frac{(B \mid B^*)}{\|B - (B \mid A)A\|} > 0$$

and similarly $(v \mid C^*) > 0$. Thus, we have either the points u, B^*, C^* and v on the circle $\langle B^*, C^* \rangle$ in this order or in the order B^*, u, v and C^* with the right angles between u and C^* and v and B^* overlapping in both cases. The claim follows easily. □

Lemma 3.18. *The perimeter of a spherical triangle is at most 2π. If the perimeter is 2π, then the vertices are all contained in the same great circle.*

Proof. This follows from the inequality (3.5) and the fact that this inequality is an equality if and only if $\gamma = \pi$. □

Proposition 3.19. *The sum of the angles of a nondegenerate triangle in $\mathbb{S}^2$ is greater than π.*

Proof. Proposition 3.17 implies that $\alpha + \beta + \gamma + a^* + b^* + c^* = 3\pi$. As $a^* + b^* + c^* < 2\pi$ by Lemma 3.18, we get the claim of Proposition 3.19. □

The following is the spherical analog of Proposition 2.4.

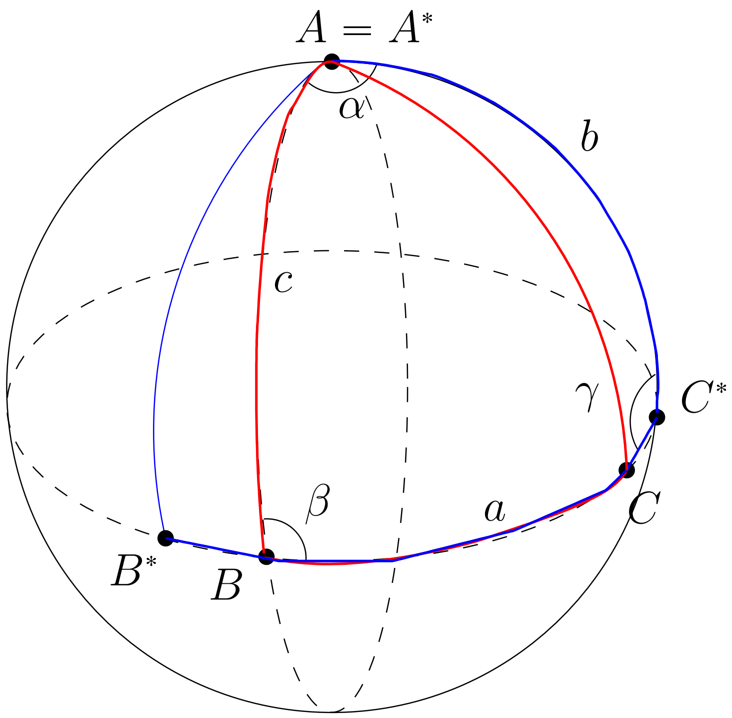


Figure 3.2 — If A is the north pole and B and C are on the equator, then $A = A^*$.

Proposition 3.20. *Let $0 < a, b, c < \pi$. If $a+b > c$, $b+c > a$, $c+a > b$ and $a+b+c < 2\pi$, then there is a triangle in $\mathbb{S}^2$ with side lengths a, b and c. All such triangles are isometric.*

Proof. We use the law of cosines in the construction: Note that if such a triangle exists, then the angle at C satisfies the cosine law. Therefore, we can compute it if we know that

$$\left|\frac{\cos c - \cos a \cos b}{\sin a \sin b}\right| < 1\,, \tag{3.7}$$

because then $\frac{\cos c - \cos a \cos b}{\sin a \sin b}$ is in the range of cos, and we can proceed with the construction. The pair of inequalities $c < a + b < 2\pi - c$ implies

$$\cos c > \cos(a + b) = \cos a \cos b - \sin a \sin b\,.$$

The inequalities $b + c > a$ and $c + a > b$ give $|a - b| < c$, which implies

$$\cos c < \cos(a - b) = \cos a \cos b + \sin a \sin b\,.$$

These two inequalities give

$$-\sin a \sin b < \cos c - \cos a \cos b < \sin a \sin b\,,$$

which implies the inequality (3.7). Now we can place the sides of length a and b starting at C in the correct angle γ. The cosine law implies that the lengths of the side opposite to C is indeed c.

The triangles are isometric by Proposition 3.12 □

3.5 Some elementary Riemannian geometry on $\mathbb{S}^2$.

Let $x \in \mathbb{S}^2$. The *latitude* of x is

$$\theta(x) = \frac{\pi}{2} - d_{\mathbb{S}^2}(x, e_3) = \frac{\pi}{2} - \arccos(x \mid e_3) = \frac{\pi}{2} - \arccos(x_3) \in [-\frac{\pi}{2}, \frac{\pi}{2}],$$

which is the oriented angle of x from the *equator* $\{x \in \mathbb{S}^2 : x_3 = 0\}$. The *longitude* of $x \in \mathbb{S}^2 - \{\pm e_3\}$ is

$$\phi(x) = \operatorname{sign}(x_2) \arccos\left(\frac{(x_1, x_2, 0) \mid e_1)}{\|(x_1, x_2, 0)\|}\right) = \operatorname{sign}(x_2) \arccos\left(\frac{x_1}{\sqrt{x_1^2 + x_2^2}}\right) \in \left]-\pi, \pi\right],$$

where $\operatorname{sign}(t) = \frac{t}{|t|}$ for nonzero t and we set $\operatorname{sign}(0) = 1$.

The longitude is the oriented angle between x and the geodesic segment from the *north pole* e_3 to the *south pole* $-e_3$, called the 0-*meridian*.[2] Here we have chosen the value π for the longitude on the *international date line* which is the geodesic segment between the poles that passes through $-e_1$. More generally, the geodesic line between the poles determined by an equation $\phi = c$ is a *meridian* and the circle determined by an equation $\theta = c$ is a *parallel*.

The longitude and latitude of a point define a bijection $L \colon \mathbb{S}^2 - \{\pm e_3\} \to \left]-\pi, \pi\right] \times \left]\frac{\pi}{2}, \frac{\pi}{2}\right[$,

$$L(x) = (\phi(x), \theta(x)) .$$

The inverse of this map is given by

$$L^{-1}(\phi, \theta) = (\cos\phi \cos\theta, \sin\phi \cos\theta, \sin\theta) .$$

This map is good close to the equator but distances, areas and angles are badly distorted close to the poles.

Let $a \in \mathbb{R} - \{0\}$ and consider the *projection plane* $P_a = \{x \in \mathbb{E}^3 : x_3 = a\}$. For any $x \in \mathbb{S}^2$, let $S_0^a \colon \mathbb{S}^2 \to P_a$ be the map

$$S_0^a(x) = (1 - a)\,\frac{x - e_3}{1 - x_3} + e_3$$

that associates to x the unique point on P_a that lies on the affine line through e_3 and x. The *stereographic projection* $S^a \colon \mathbb{S}^2 - \{e_3\} \to \mathbb{E}^2$ is $\operatorname{pr}_3 \circ S_0^a$, where $\operatorname{pr}_3(y) = (y_1, y_2)$ is the orthogonal projection of $\mathbb{E}^3$ to $\mathbb{E}^2$ identified with the hyperplane $\mathbb{E}^2 \times \{0\}$:

$$S^a(x) = (1 - a)\Big(\frac{x_1}{1 - x_3}, \frac{x_2}{1 - x_3}\Big) .$$

Most often, one uses $a = 0$, which is the case where the projection plane passes through the origin, or $a = -1$, which is the case where the projection plane is tangent to the sphere at the south pole.

[2]This is the *Greenwich meridian* if we consider the Earth with its standard coordinates.

Length and area

The *(differential geometric) length* of a piecewise continuously differentiable path $\tau\colon I \to \mathbb{S}^2$ is

$$\ell(\tau) = \int_I \|\dot{\tau}\| \,,$$

where $\dot{\tau}(t)$ is the tangent (derivative) vector of the path for each $t \in I$.

Proposition 3.21. *Let $A, B \in \mathbb{S}^2$, $A \neq B$. Let j be a spherical segment that connects A and B. Then $\ell(j) \leqslant \ell(\tau)$ for all piecewise continuously differentiable paths τ.*

Proof. Using an isometry of $\mathbb{S}^2$, we can assume that A and B are contained in the 0-meridian. Using longitude-latitude coordinates, consider the continuous map proj defined by $\operatorname{proj}(\phi, \theta) = (0, \theta)$ whose image is contained in the 0-meridian. Clearly, $\ell(j) \leqslant \ell(\operatorname{proj} \circ \tau) \leqslant \ell(\tau)$. □

In the computation of the length of a path τ, the norm of the tangent vector $\dot{\tau}(t)$ is computed in the tangent plane $\tau(t)^\perp$ at $\tau(t)$. Using the coordinate maps, we get

The inner product of the tangent spaces can be used to define the area of a subset of the sphere. This gives the expressions

$$\operatorname{Area} A = \int_{L(A)} \cos\theta d\theta d\phi$$

in the longitude-latitude coordinates and

$$\operatorname{Area} A = \int_{S^0(A)} \frac{4\, dx_1 dx_2}{(1 + \|x\|^2)^2}$$

in the coordinates given by the stereographic projection.

Proposition 3.22. *The area of $\mathbb{S}^2$ is 4π.*

Let $0 < \alpha < \pi$. The area of the *(spherical) sector* $S_\alpha = \{x \in \mathbb{S}^2 : 0 \leqslant \phi(x) \leqslant \alpha\}$ and any of its isometric images is easily seen to be $\frac{\alpha}{2\pi} 4\pi = 2\alpha$.

Proposition 3.23 (Girard)**.** *The area of a triangle with angles α, β and γ is $\alpha + \beta + \gamma - \pi$.*

Proof. Let A, B and C be the vertices of the triangle. The antipodal points $-A$, $-B$ and $-C$ determine a triangle $(-A)(-B)(-C)$ that is isomorphic with ABC. The three great circles $\langle A, B\rangle \cap \mathbb{S}^2$, $\langle B, C\rangle \cap \mathbb{S}^2$ and $\langle C, A\rangle \cap \mathbb{S}^2$ determine six sectors with angles $\alpha, \alpha, \beta, \beta, \gamma, \gamma$ that cover the sphere. In the complement of the great circles, the triangles ABC and $(-A)(-B)(-C)$ are both covered by three sectors, other points are contained in one sector. Thus,

$$\begin{aligned} 4\pi &= \operatorname{Area} \mathbb{S}^2 = 2(\operatorname{Area} S_\alpha + \operatorname{Area} S_\beta + \operatorname{Area} S_\gamma) - 4 \operatorname{Area} ABC \\ &= 2(2\alpha + 2\beta + 2\gamma) - 4 \operatorname{Area} ABC \,, \end{aligned}$$

which gives the claim. □

Exercises

3.1. Prove Proposition 3.9(3).

3.2. Let H be a hyperplane in $\mathbb{S}^n$. Prove that $d(r_H(x), y) = d(x, y)$ for all $x \in \mathbb{S}^n$ and $y \in H$.

3.3. Let $\phi \in \mathrm{Isom}(\mathbb{S}^n) - \{\mathrm{id}\}$. Let H be a hyperplane such that $\phi|_H = \mathrm{id}|_H$. Prove that $\phi = r_H$.

3.4. Prove Corollary 3.12 for $n = 2$.

3.5. Prove Corollary 3.13.

3.6. Prove Proposition 3.15.

3.7. Prove Lemma 3.16.

Chapter 4

Hyperbolic space

In this chapter, we define hyperbolic space as a submanifold of Minkowski space with a metric that is analogous with the angle metric on the sphere. We will show that hyperbolic space is a uniquely geodesic metric space and that the orthogonal group of the Minkowski bilinear form is the group of isometries of hyperbolic space. The proof uses the hyperbolic law of cosines.

4.1 Minkowski space

In this section we introduce the indefinite Minkowski bilinear form in $\mathbb{R}^{n+1}$ and, in particular, the associated subset $\mathbb{H}^n$ that is used to define hyperbolic n-space in Section 4.3.

Let V and W be real vector spaces. A map $\Phi\colon V \times W \to \mathbb{R}$ is a *bilinear form*, if the maps $v \mapsto \Phi(v, w_0)$ and $v \mapsto \Phi(v_0, w)$ are linear for all $w_0 \in W$ and all $v_0 \in V$.

A bilinear form Φ is *nondegenerate* if

- $\Phi(x, y) = 0$ for all $y \in W$ only if $x = 0$, and
- $\Phi(x, y) = 0$ for all $x \in V$ only if $y = 0$.

If $W = V$, then Φ is *symmetric* if $\Phi(x, y) = \Phi(y, x)$ for all $x, y \in V$. It is

- *positive semidefinite* if $\Phi(x, x) \geqslant 0$ for all $x \in V$,
- *positive definite* if $\Phi(x, x) > 0$ for all $x \in V - \{0\}$,
- *negative (semi)definite* if $-\Phi$ is positive (semi)definite, and
- *indefinite* otherwise.

The function $q\colon V \to \mathbb{R}$, $q(x) = \Phi(x, x)$ is the *quadratic form* corresponding to a bilinear form $\Phi\colon V \times V \to \mathbb{R}$.

A positive definite symmetric bilinear form is often called an *inner product* or a *scalar product.*

If V is a real vector space with a symmetric bilinear form Φ, two vectors $u, v \in V$ are *orthogonal*, $u \perp v$, if $\Phi(u,v) = 0$. The *orthogonal complement* of $u \in V$ is

$$u^{\perp} = \{v \in V : u \perp v\}.$$

Let us consider the indefinite nondegenerate symmetric bilinear form $\langle \cdot \,|\, \cdot \rangle$ on $\mathbb{R}^{n+1}$ given by

$$\langle x \,|\, y \rangle = -x_0 y_0 + \sum_{i=1}^{n} x_i y_i = -x_0 y_0 + (\bar{x} \,|\, \bar{y}) = x^T J y,$$

where

$$J = J_{1,n} = \operatorname{diag}(-1, 1, \ldots, 1)$$

and $x = (x_0, x_1, \ldots, x_n) = (x_0, \bar{x})$.

The bilinear form $\langle \cdot \,|\, \cdot \rangle$ is the *Minkowski bilinear form*, and the pair

$$\mathbb{M}^{1,n} = \big(\mathbb{R}^{n+1}, \langle \cdot \,|\, \cdot \rangle\big)$$

is the $n+1$-dimensional *Minkowski space*.
A vector $x \in \mathbb{M}^{1,n} - \{0\}$ is

- *lightlike*[a] if $\langle x \,|\, x \rangle = 0$,
- *timelike* if $\langle x \,|\, x \rangle < 0$, and
- *spacelike* if $\langle x \,|\, x \rangle > 0$.

[a]Light-like vectors are also called *null-vectors*

The names for the three different types of vectors in Minkowski space come from Einstein's special theory of relativity, which lives in $\mathbb{M}^{1,3}$. Minkowski space has a number of geometrically significant subsets: The subset of *null-vectors* is the *light cone*

$$\mathscr{L}^n = \{x \in \mathbb{M}^{1,n} : \langle x \,|\, x \rangle = 0\}\,.$$

The smooth submanifold

$$\mathscr{L}^n_- = \{x \in \mathbb{M}^{1,n} : \langle x \,|\, x \rangle = -1\}$$

is a two-sheeted hyperboloid, and its upper sheet is

$$\mathbb{H}^n = \{x \in \mathbb{M}^{1,n} : \langle x \,|\, x \rangle = -1, \; x_0 > 0\}.$$

The smooth submanifold

$$\mathscr{L}^n_+ = \{x \in \mathbb{M}^{1,n} : \langle x \,|\, x \rangle = 1\}$$

is a one-sheeted hyperboloid.

The following is an important observation on time-like vectors.

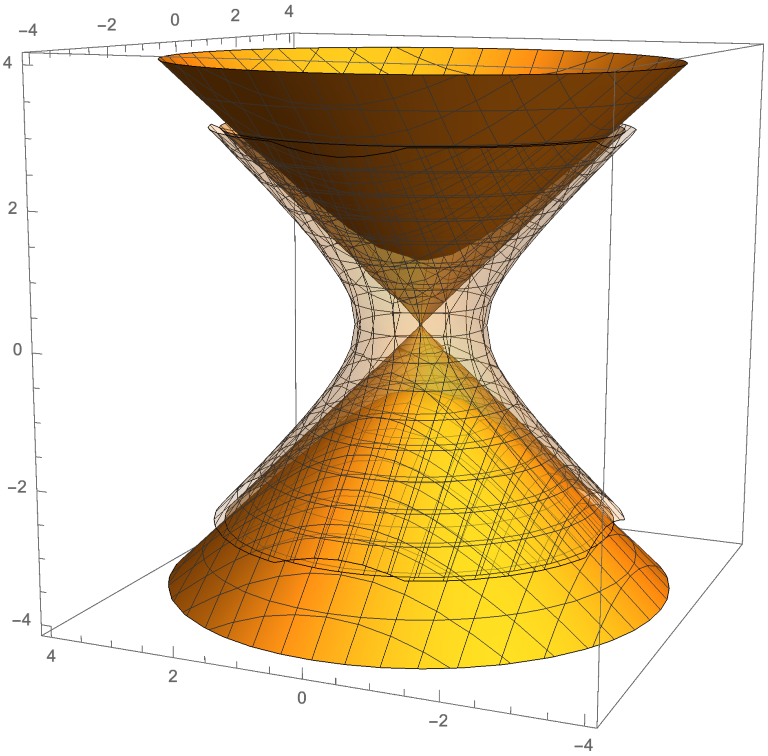


Figure 4.1 — The upper sheet of the two-sheeted hyperboloid with the lightcone and the one-sheeted hyperboloid.

Lemma 4.1. *If $u, v \in \mathbb{H}^n$, then $\langle u \,|\, v\rangle \leqslant -1$ with equality only if $u = v$.*

Proof. Using the Cauchy inequality for the Euclidean inner product in $\mathbb{R}^n$ for the first inequality and a simple calculation[1] for the second, we have

$$\begin{aligned}\langle u \,|\, v\rangle &= -u_0v_0 + \sum_{i=1}^n u_iv_i \leqslant -u_0v_0 + \sqrt{\sum_{i=1}^n u_i^2}\sqrt{\sum_{i=1}^n v_i^2} \\ &= -u_0v_0 + \sqrt{u_0^2-1}\sqrt{v_0^2-1} \leqslant -1\,.\end{aligned}$$

Cauchy's inequality is an equality if and only if u and v are parallel, and the final inequality is an equality if and only if $u_0 = v_0$. This implies the claim on equality. □

4.2 The orthogonal group of Minkowski space

The *orthogonal group* of the Minkowski bilinear form is

$$\begin{aligned}\mathrm{O}(1,n) &= \{A \in \mathrm{GL}_{n+1}(\mathbb{R}) : \langle Ax \,|\, Ay\rangle = \langle x \,|\, y\rangle \ \text{ for all } x, y \in \mathbb{M}^{1,n}\} \\ &= \{A \in \mathrm{GL}_{n+1}(\mathbb{R}) : {}^{T}\!AJ_{1,n}A = J_{1,n}\}\,.\end{aligned}$$

An element of $\mathrm{O}(1,n)$ is an *orthogonal transformation*.

[1] Manipulate the given inequality to remove the square roots etc.

Clearly, the linear action of $\mathrm{O}(1,n)$ on $\mathbb{M}^{1,n}$ preserves the light cone and the two-sheeted hyperboloid $\mathscr{L}^n$.

Let $A = (a_0, a_1, \ldots, a_n)$ be an $(n+1)\times(n+1)$-matrix A in terms of its column vectors $a_0, a_1, \ldots, a_n \in \mathbb{R}^{n+1}$. If $A \in \mathrm{O}(1,n)$, then $a_0 = A(e_0)$ for $e_0 = (1,0,\ldots,0) \in \mathbb{H}^n$. Thus $A(e_0) \in \mathbb{H}^n$ if and only if $A_{00} > 0$, and therefore the stabiliser in $\mathrm{O}(1,n)$ of the upper sheet $\mathbb{H}^n$ is

$$\begin{aligned}\mathrm{O}^+(1,n) &= \{A \in \mathrm{O}(1,n) : A\mathbb{H}^n = \mathbb{H}^n\} \\ &= \{A \in \mathrm{GL}_{n+1}(\mathbb{R}) : A_{00} > 0,\ \langle Ax \,|\, Ay\rangle = \langle x \,|\, y\rangle \text{ for all } x, y \in \mathbb{M}^{1,n}\} \\ &= \{A \in \mathrm{GL}_{n+1}(\mathbb{R}) : A_{00} > 0,\ {}^{T}\!AJ_{1,n}A = J_{1,n}\}\,.\end{aligned} \tag{4.1}$$

Let us check that the second of the three equalities in (4.1) holds: Let $A \in \mathrm{GL}_{n+1}(\mathbb{R})$ with $A_{00} > 0$ and $\langle Ax \,|\, Ay\rangle = \langle x \,|\, y\rangle$ for all $x, y \in \mathbb{M}^{1,n}$. The first and third properties are equivalent with $A \in \mathrm{O}(1,n)$ so it remains to check that $A\mathbb{H}^n = \mathbb{H}^n$. We know that $Ae_0 \in \mathbb{H}^n$. Linear automorphisms of $\mathbb{E}^{n+1}$ are continuous mappings and the image of a connected set under a continuous map is connected, so $\mathbb{H}^n$ is mapped into $\mathbb{H}^n$. Similarly, the lower half of the hyperboloid $\mathscr{L}^n$ is mapped into itself. Furthermore, the elements of $\mathrm{GL}_{n+1}(\mathbb{R})$ are linear bijections, so the restriction to $\mathbb{H}^n$ is a bijection of $\mathbb{H}^n$.

A basis $\{v_0, v_1, \ldots, v_n\}$ of $\mathbb{M}^{1,n}$ is *orthonormal* if the basis elements are pairwise orthogonal and if $\langle v_0 \,|\, v_0\rangle = -1$ and $\langle v_i \,|\, v_i\rangle = 1$ for all $i \in \{1, 2, \ldots, n\}$.

The following observation is proved in the same way as its Euclidean analog:

Lemma 4.2. *An $(n+1)\times(n+1)$-matrix $A = (a_0, a_1, \ldots, a_n)$ is in $\mathrm{O}(1,n)$ if and only if the vectors $a_0, a_1, \ldots, a_n$ form an orthonormal basis of $\mathbb{M}^{1,n}$. Furthermore, $A \in \mathrm{O}^+(1,n)$ if and only if $A \in \mathrm{O}(1,n)$ and $a_0 \in \mathbb{H}^n$.*

Proof. Exercise. □

Example 4.3. (1) Let $t \in \mathbb{R}$. The matrix

$$L_t = \begin{pmatrix} \cosh t & \sinh t & 0 \\ \sinh t & \cosh t & 0 \\ 0 & 0 & 1 \end{pmatrix} \in \mathrm{O}^+(1,2)$$

stabilizes any affine hyperplane

$$H_c = \{x \in \mathbb{M}^{1,2} : x_2 = c\}\,. \tag{4.2}$$

In particular, the path $t \mapsto L_t e_0 = (\cosh t, \sinh t, 0)$ parametrizes the hyperbola

$$\{x \in \mathbb{H}^2 : x_2 = 0\} = \mathbb{H}^2 \cap \{x \in \mathbb{M}^{1,2} : x_2 = 0\}\,.$$

(2) For any $\theta \in \mathbb{R}$, let $\widehat{R}_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \in \mathrm{O}(2)$, and let

$$R_\theta = \mathrm{diag}(1, \widehat{R}_\theta) = \begin{pmatrix} 1 & 0 \\ 0 & \widehat{R}(\theta) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \in \mathrm{O}^+(1,2)\,.$$

The mapping R_θ is a Euclidean rotation around the vertical axis by the angle θ. The rotation R_θ stabilizes each affine hyperplane

$$E_r = \{x \in \mathbb{M}^{1,2} : x_0 = r\}$$

for any $r \in \mathbb{R}$. Another important mapping that comes by extension from $\mathrm{O}(2)$ is given by the matrix $\mathrm{diag}(1, 1, -1)$, which is a reflection in the hyperplane H_0 defined in equation (4.2).

(3) The above examples can be generalized to higher dimensions:

- L_t is extended as the identity on the last coordinates to $\mathrm{diag}(L_t, I_{n-2}) \in \mathrm{O}^+(1, n)$.
- Any Euclidean orthogonal matrix $A \in \mathrm{O}(n)$ gives an isometry $\mathrm{diag}(1, A) \in \mathrm{O}^+(1, n)$.

Proposition 4.4. *The group* $\mathrm{O}^+(1, n)$ *acts transitively on* $\mathbb{H}^n$ *and on the one-sheeted hyperboloid* $\mathscr{L}^n_+$.

Proof. We use the notation of Example 4.3. If $x \in \mathbb{H}^n$, then $x = (\sqrt{\|\bar{x}\|^2 + 1}, \bar{x})$. There is some $\widehat{R}_\theta \in \mathrm{O}(n)$ such that $\widehat{R}_\theta \bar{x} = \|\bar{x}\| e_1$, and thus, $R_\theta(x) = (\sqrt{\|\bar{x}\|^2 + 1}, \|\bar{x}\| e_1)$. Furthermore,

$$L_{\operatorname{arsinh} \|\bar{x}\|} e_0 = (\sqrt{\|\bar{x}\|^2 + 1}, \|\bar{x}\| e_1)\,,$$

and we have $x = R_\theta^{-1} L_{\operatorname{arsinh} \|\bar{x}\|} e_0$. This implies that $\mathbb{H}^n$ is the $\mathrm{O}^+(1, n)$-orbit of e_0.

A similar proof shows that $\mathscr{L}^n_+$ is the $\mathrm{O}^+(1, n)$-orbit of $e_1 \in \mathbb{M}^{1,n}$, see Exercise 4.2. ☐

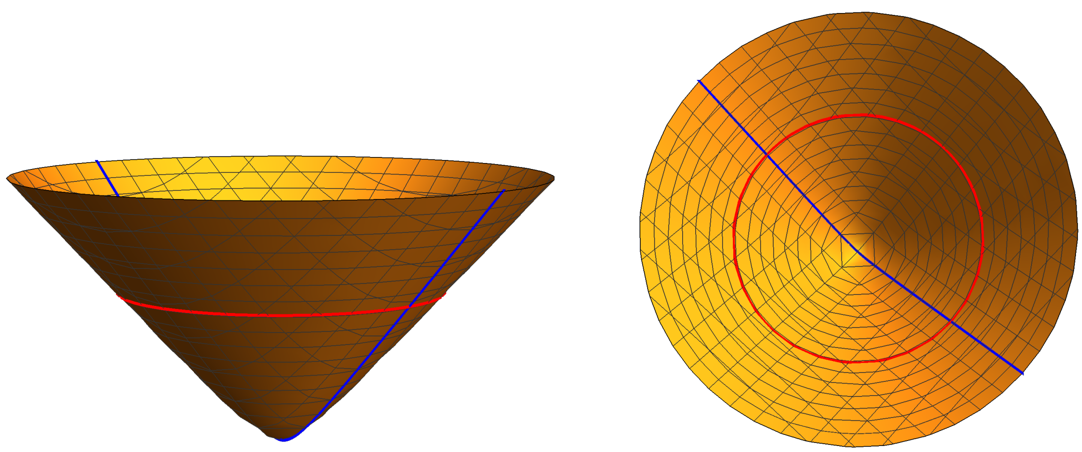


Figure 4.2 — The idea of the proof of Proposition 4.4 : R_θ moves the point x along the red circle to the blue curve and L_t moves the point along the blue curve to e_0. The hyperboloid is seen from the side and from the top.

The proofs of the following propositions demonstrate the use of a transitive group of transformations:

Proposition 4.5. *The restriction of the Minkowski bilinear form to the orthogonal complement of a timelike vector is positive definite.*[2]

[2]Naturally, the orthogonal complement is defined with respect to the Minkowski bilinear form.

Proof. Let $v \in \mathbb{M}^{1,n}$ be a timelike vector. We may assume that $v \in \mathbb{H}^n$. Proposition 4.4 implies the existence of an element $A \in \mathrm{O}^+(1,n)$ such that $Av = e_0$. The orthogonal complement of e_0 is the subspace $\{x \in \mathbb{M}^{1,n} : x_0 = 0\}$. The restriction of the Minkowski bilinear form to this subspace is the standard Euclidean inner product. By definition, $\langle A^{-1}u \,|\, A^{-1}u\rangle = \langle u \,|\, u\rangle > 0$ for all $u \in e_0^\perp$. □

Proposition 4.6. *For any $a \in \mathbb{H}^n$, the tangent space $T_a\mathbb{H}^n$ of $\mathbb{H}^n$ at a coincides with $a^\perp$.*

Proof. Let $p \in \mathbb{H}^n$. As the group $\mathrm{O}^+(1,n)$ acts transitively on $\mathbb{H}^n$ there is some $A \in \mathrm{O}^+(1,n)$ such that $Ae_0 = p$. As in Proposition 4.5, $Ae_0^\perp = p^\perp$. Considering the linear map A as a differentiable mapping of $\mathbb{R}^{n+1}$ to itself, its differential that coincides with A maps the tangent space at e_0 to the tangent spaces at p. Clearly,

$$T_{e_0}\mathbb{H}^n = \{x \in \mathbb{M}^{1,n} : x_0 = 0\} = e^\perp$$

and the same holds at p by the observations we just made. □

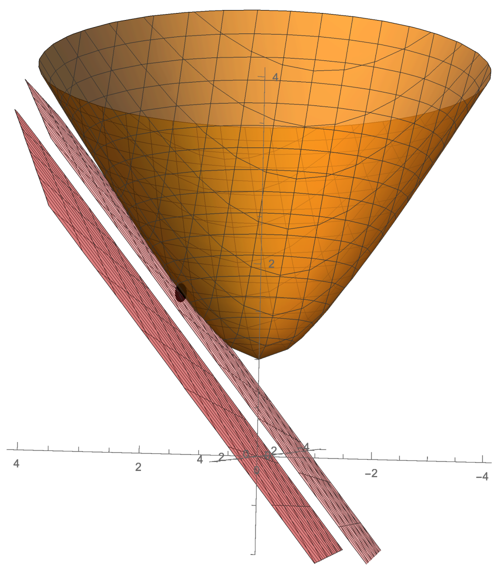


Figure 4.3 — The orthogonal complement $p^\perp$ of a point $p \in \mathbb{H}^2$ coincides with the tangent space $T_p(\mathbb{H}^2)$ as a vector subspace of $\mathbb{R}^3$. The figure also shows the affine tangent plane $p + p^\perp$ that is tangent to $\mathbb{H}^2$ at p. If we consider the standard Euclidean inner product in $\mathbb{R}^3$, the tangent plane coincides with the orthogonal complement only at e_0.

Propositions 4.5 and 4.6 imply that the restriction of the Minkowski bilinear form to each tangent space defines a Riemannian metric.

The *Riemannian metric* of $\mathbb{H}^n$ associates the inner product $\langle \cdot \mid \cdot \rangle|_{a^\perp}$ to all points $a \in \mathbb{H}^n$. The *norm* in $a^\perp$ is

$$|u| = \sqrt{\langle u \mid u \rangle}$$

for all $u \in a^\perp$.
The *angle* $\measuredangle(u, v)$ of any two vectors $u, v \in T_a\mathbb{H}^n = a^\perp - \{0\}$ is

$$\measuredangle(u, v) = \arccos\left(\frac{\langle u \mid v \rangle}{|u|\,|v|}\right).$$

We will not discuss Riemannian geometry in a formal manner. Hyperbolic space is an important example of a Riemannian manifold, and sometimes[3] hyperbolic metric is defined as a Riemannian metric. In that approach, hyperbolic metric appears as the path metric of the Riemannian metric.

The *Riemannian length* of a piecewise smooth path $\gamma\colon [a, b] \to \mathbb{H}^n$ is

$$\ell(\gamma) = \int_a^b \sqrt{\langle \dot\gamma(t) \mid \dot\gamma(t) \rangle}\, dt\,.$$

The length metric of the Riemannian metric of $\mathbb{H}^n$ is

$$d_{\mathrm{Riem}}(x, y) = \inf \ell(\gamma),$$

where the infimum is taken over all piecewise smooth paths that connect x to y.

In section 5.3, we will show that the Riemannian approach leads to the same hyperbolic metric as the one we will define in section 4.3. Riemannian geometry also provides a natural concept of volume in hyperbolic space, and we will discuss this in section 5.9.

4.3 Hyperbolic space

In this section, we define a metric on the upper sheet $\mathbb{H}^n$ using the Minkowski bilinear form analogously with the definition of the spherical metric in section 3.1.

The metric space $(\mathbb{H}^n, d)$, where

$$d(x, y) = \operatorname{arcosh}(-\langle x \mid y \rangle) \in [0, \infty[\,,$$

is the *hyperboloid model* of n-dimensional *(real) hyperbolic space*. The metric d is the *hyperbolic metric.*

We still need to show that the hyperbolic metric is a metric. The proof follows the same idea that was used to treat the angle metric for the sphere $\mathbb{S}^n$.

[3] See [And] or [Bea].

Let $a \in \mathbb{H}^n$, and let $u \in a^\perp$ such that $\langle u \,|\, u\rangle = 1$. [a] The mapping $j_{a,u} : \mathbb{R} \to \mathbb{H}^n$,

$$j_{a,u}(t) = a\cosh(t) + u\sinh(t)\,,$$

is the *hyperbolic line* through a in *direction* u. For any $T > 0$, the restriction $j_{a,u}|_{[0,T]}$ is a *hyperbolic segment.*

[a] Recall that the restriction of the Minkowski bilinear form to $a^\perp$ is positive definite by Corollary 4.5.

Lemma 4.7. *Let $a \in \mathbb{H}^n$ and $u \in a^\perp$.*
(1) The image of $j_{a,u}$ is contained in $\mathbb{H}^n$.
(2) For all $s, t \in \mathbb{R}$, we have

$$d(j_{a,u}(t), j_{a,u}(s)) = |s - t|\,. \tag{4.3}$$

(3) $A \circ j_{a,u} = j_{Aa,Au}$ for all $A \in \mathrm{O}^+(1,n)$.

Proof. Exercise 4.3. □

As in section 3.1 for the sphere, if we show that d is a metric, then Lemma 4.7 implies that $j_{a,u}$ is a geodesic line.

Lemma 4.8. *Let $p, q \in \mathbb{H}^n$ be two distinct points. Let*

$$u = \frac{q + \langle p \,|\, q\rangle p}{\sqrt{\langle p \,|\, q\rangle^2 - 1}}\,.$$

Then $j_{p,u}(0) = p$ and $j_{p,u}(\operatorname{arcosh}(-\langle p \,|\, q\rangle)) = q$.

Proof. Observe that Lemma 4.1 implies

$$\big\langle q + \langle p \,|\, q\rangle p \,\big|\, q + \langle p \,|\, q\rangle p\big\rangle = \langle p \,|\, q\rangle^2 - 1 > 0\,.$$

Thus, u is a unit tangent vector to the hyperboloid. The fact that $j_{p,u}(0) = p$ is immediate, and the other claim follows by noting that $\sinh(\operatorname{arcosh}(-\langle p \,|\, q\rangle)) = \sqrt{\langle p \,|\, q\rangle^2 - 1}$. □

Lemma 4.9. *For any $a \in \mathbb{H}^n$ and any $u \in a^\perp$, $j_{a,u}(\mathbb{R}) = \mathbb{H}^n \cap \langle a, u\rangle$. If a 2-plane T intersects $\mathbb{H}^n$, then $T \cap \mathbb{H}^n$ is the image of a hyperbolic line.*

Proof. Clearly, the image of $j_{a,u}$ is contained in the 2-plane $\langle a, u\rangle$. The fact the image of $j_{a,u}$ coincides with $\langle a, u\rangle \cap \mathbb{H}^n$ follows from the second statement of the Lemma that we prove below.

If $T = \langle e_0, e_1\rangle$, then $\mathbb{H}^n \cap T$ is a copy of the upper half of the hyperbola

$$\{x \in \mathbb{R}^2 : -x_0^2 - x_1^2 = -1\}\,,$$

and this intersection is parametrized by j_{e_0,e_1}. If $T = \langle e_0, v\rangle$ for any $v \in e_0^\perp$, then there is an element $B \in \mathrm{O}(n)$ such that $Be_1 = v$ and, consequently, an element $B' = \operatorname{diag}(1, B) \in \mathrm{O}^+(1,n)$ such that $Be_0 = e_0$ and $Be_1 = v$. Thus, $\mathbb{H}^n \cap T = B'(\mathbb{H}^n \cap \langle e_0, e_1\rangle)$ coincides with the image of the hyperbolic line $B' \circ j_{e_0,e_1} = j_{B'e_0,B'e_1} = j_{e_0,v}$, see Lemma 4.7.

If the plane T does not pass through e_0 but intersects $\mathbb{H}^n$, then Proposition 4.4 provides an element $A \in \mathrm{O}^+(1,n)$ such that $T = A(T_0)$ for some plane T_0 that intersects $\mathbb{H}^n$ at e_0. We saw above that this intersection is parametrized by a hyperbolic line $j_{e_0,v}$ for some $v \in e_0^\perp$. As above, we see that $\mathbb{H}^n \cap T$ is parametrized by $A \circ j_{e_0,v} = j_{Ae_0,A,v}$. □

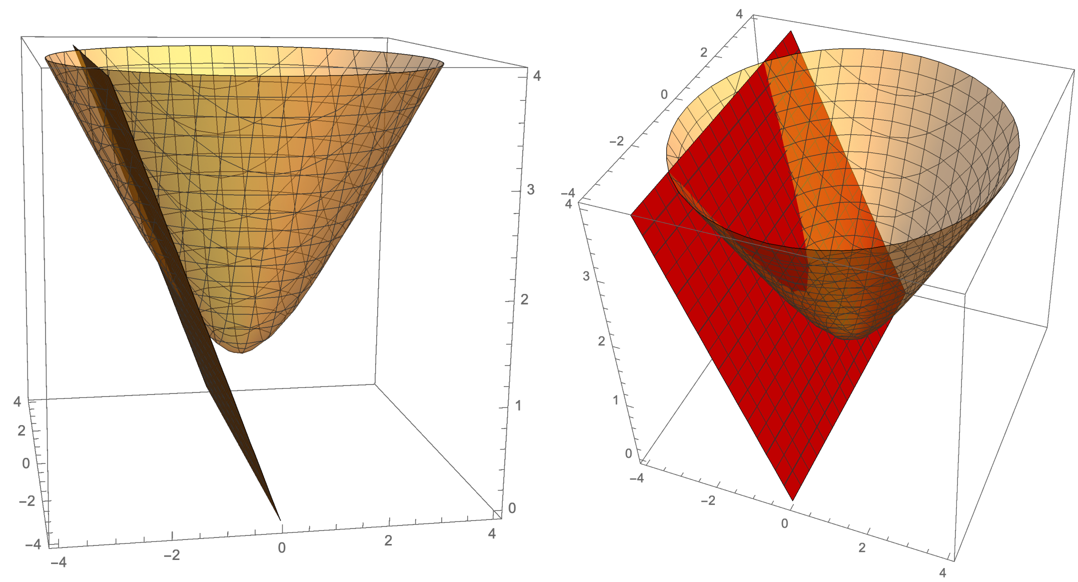


Figure 4.4 — A linear plane that intersects $\mathbb{H}^2$ seen from two different angles.

The fact that the hyperbolic metric is indeed a metric is proved in the same way as Proposition 3.3 in the spherical case. First we prove the law of cosines for triangles in hyperbolic space. As we cannot use a metric yet, we consider triangles whose sides are hyperbolic segments. The angles at the vertices are defined using the Riemannian metric of $\mathbb{H}^n$. We use the notation for triangles introduced in section 1.6.

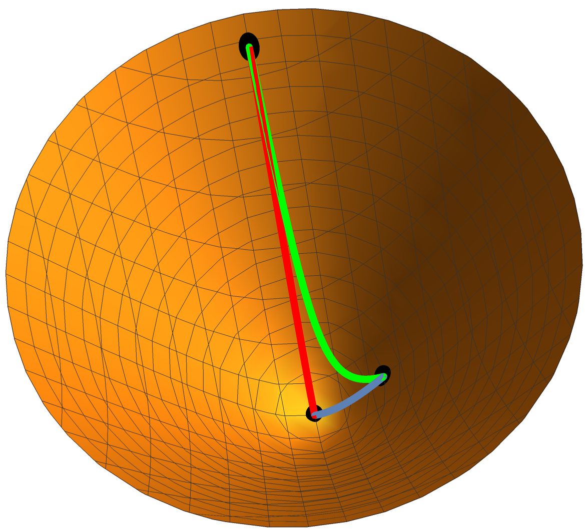


Figure 4.5 — A triangle in $\mathbb{H}^2$ with a vertice at e_0.

Proposition 4.10 (The first hyperbolic law of cosines)**.**

$$\cosh c = \cosh a \cosh b - \sinh a \sinh b \cos \gamma \ .$$

Proof. Let u and v be the initial tangent vectors of the hyperbolic segments from C to A and from C to B. As u and v are orthogonal to C, we have as in the spherical case,

$$\begin{aligned}\cosh c &= -\langle A \mid B\rangle = -\langle \cosh(b)C + \sinh(b)u \mid \cosh(a)C + \sinh(a)v\rangle \\ &= \cosh(a)\cosh(b) - \sinh(b)\sinh(a)\langle u \mid v\rangle\,.\end{aligned}$$ □

Theorem 4.11. *Hyperbolic space is a uniquely geodesic metric space. Hyperbolic lines are geodesic lines.*

Proof. To show that the hyperbolic metric is a metric, let $A, B, C \in \mathbb{H}^n$. Using the fixed notation for the hyperbolic triangle with vertices A, B and C, consider the strictly increasing function $f\colon [0,\pi] \to \mathbb{R}$,

$$f(\gamma) = \cosh a \cosh b - \sinh a \sinh b \cos \gamma,$$

that has a unique maximum at $\gamma = \pi$ with

$$\gamma(\pi) = \cosh a \cosh b + \sinh a \sinh b = \cosh(a+b)\,.$$

The first law of cosines implies that $\cosh c \leqslant \cosh(a+b)$, which yields the triangle inequality.

Now that we know that hyperbolic space is a metric space, hyperbolic lines are geodesic lines by Lemma 4.7(2). If A and B are distinct points in $\mathbb{H}^n$, there is a unique 2-plane T through them. Thus, there is exactly one image of a hyperbolic line through these points. Assume that there is a geodesic segment $k\colon [0, d(A,B)] \to \mathbb{H}^n$ such that $k(0) = A$, $k(d(A,B)) = B$ and the image of k is not contained in T. Let $C \in k([0, d(A,B)]) - T$ and consider the triangle with vertices A, B and C and sides the unique hyperbolic segments connecting A to B, B to C and C to A. As the function f is strictly increasing, equality is possible in the triangle inequality only when $\gamma = \pi$. This implies that the segments from B to C and from C to A are contained in a hyperbolic line. This hyperbolic line contains A and B and, therefore, the sides from B to C and from C to A are contained in the side from A to B, but this is a contradiction. Thus, $\mathbb{H}^n$ is uniquely geodesic. □

We will postpone the proof of the following important result until Section 5.3 where the details are simplified by a smart choice of coordinates.

Theorem 4.12. *Hyperbolic metric is the length metric of the Riemannian metric of hyperbolic space.*

4.4 Isometries of $\mathbb{H}^n$

Proposition 4.13. $\mathrm{O}^+(1,n)$ *acts transitively by isometries on* $\mathbb{H}^n$. *In particular,* $\mathrm{Isom}(\mathbb{H}^n)$ *acts transitively on* $\mathbb{H}^n$.

Proof. Transitivity of the action was proved in Proposition 4.4 so it remains to show that the elements of $\mathrm{O}^+(1,n)$ act as isometries. Let $g \in \mathrm{O}^+(1,n)$, and let $x, y \in \mathbb{H}^n$. By the definition of the hyperbolic metric and of $\mathrm{O}^+(1,n)$, we have

$$d(g(x), g(y)) = \operatorname{arcosh}(-\langle g(x) \mid g(y)\rangle) = \operatorname{arcosh}(-\langle x \mid y\rangle) = d(x,y)\,.$$ □

Example 4.14. (1) Let $t \in \mathbb{R}$. The matrix L_t of Example 4.3 acts on $\mathbb{H}^2$ as an isometry that preserves the intersection of $\mathbb{H}^2$ with any affine 2-plane $\{x \in \mathbb{M}^{1,2} : x_2 = c\}$. In particular, it stabilizes the geodesic line

$$\ell = \{x \in \mathbb{H}^3 : x_2 = 0\} .$$

For any point $p = (a, b, 0) \in \ell$, we have

$$d(L_t(p), p) = \operatorname{arcosh}(-\langle L_t p \,|\, p\rangle) = \operatorname{arcosh}((-a^2 + b^2)\cosh(t)) = |t| .$$

In chapter 5, we will see that all other points are moved a longer distance than $|t|$.
(2) If $r > 0$, then the set

$$\mathbb{H}^n \cap \{(\cosh r, \bar{x}) : \bar{x} \in \mathbb{R}^n\} = \{(\cosh r, \bar{x}) : \bar{x} \in \mathbb{R}^n\,,\ \|\bar{x}\| = \sinh r\}$$

is the sphere of radius r centered at the point $e_0 \in \mathbb{H}^n$. If $A \in \mathrm{O}(n)$, the isometry $\operatorname{diag}(1, A) \in \mathrm{O}^+(1, n)$ maps each sphere centered at e_0 to itself, and the subgroup $\{\operatorname{diag}(1, A) \in \mathrm{O}^+(1, n) : A \in \mathrm{O}(n)\} = \operatorname{Stab} e_0 < \operatorname{Isom} \mathbb{H}^n$ acts transitively on each such sphere.
(3) For each $v \in \mathscr{L}^2$ and $c < 0$, the set

$$\{x \in \mathbb{H}^2 : \langle v \,|\, x\rangle = c\}$$

is called a *horosphere* based at v. The mapping given by the matrix

$$N_s = \begin{pmatrix} 1 + \frac{s^2}{2} & -\frac{s^2}{2} & s \\ \frac{s^2}{2} & 1 - \frac{s^2}{2} & s \\ s & -s & 1 \end{pmatrix} \in \mathrm{O}^+(1, 2)$$

maps each horosphere based at $(1, 1, 0) \in \mathscr{L}^2$ to itself.
(4) Composing some number of the above mappings we obtain further examples of isometries of the hyperbolic plane. For example, if $p \in \mathbb{H}^2$, then there is some $\theta \in \mathbb{R}$ such that $R_\theta(p) \in \ell$. Now, $L^{-1}_{d(e_0,p)}(R_\theta(p)) = L_{-d(e_0,p)}(R_\theta(p)) = e_0$, and for any $\phi \in \mathbb{R}$, the mapping $S = R_{-\theta} \circ L_{d(e_0,p)} \circ R_\phi \circ L^{-1}_{d(e_0,p)} \circ R_\theta$ is an isometry that fixes p and maps each sphere centered at p to itself. The mapping S is conjugate[4] to R_ϕ in $\operatorname{Isom}(\mathbb{H}^n)$.

The isometries introduced above are classified according to the conic sections they correspond to. The mapping L_t and any of its conjugates in $\operatorname{Isom}(\mathbb{H}^n)$ is called *hyperbolic* because L_t maps each affine plane parallel to the (x_0, x_1)-plane in $\mathbb{M}^{1,2}$ to itself, and these planes intersect the light cone in hyperbola, except for the (x_0, x_1)-plane itself that intersects the lightcone in a pair of lines.

The mapping R_θ and any of its conjugates is called *elliptic* because R_θ preserves all horizontal hyperplanes in $\mathbb{M}^{1,2}$ and their intersections with $\mathscr{L}^2$, which are circles centered at points of the 0:th coordinate axis.

The mapping N_s and any of its conjugates is called *parabolic* because it preserves all affine hyperplanes $\{x \in \mathbb{M}^{1,2} : \langle v \,|\, x\rangle = c\}$, which intersect $\mathscr{L}^2$ in a parabola when $c < 0$.

As in the Euclidean and spherical geometries, we will now study a fundamental class of isometries, reflections in a hyperplane.

[4] If G is a group and $g, h \in G$, then the elements g and hgh^{-1} are *conjugate* elements in G.

If T is an $(m+1)$-dimensional linear subspace of $\mathbb{R}^{n+1}$ that intersects $\mathbb{H}^n$, then $T \cap \mathbb{H}^n$ is an m-dimensional *hyperbolic subspace* of $\mathbb{H}^n$. If $m = n-1$, then T is a *hyperplane*.

Proposition 4.15. *Let $1 \leqslant m < n$. Any two hyperbolic m-dimensional subspaces of $\mathbb{H}^n$ can be mapped to each other by isometries of $\mathbb{H}^n$.*

Proof. Exercise 4.4. □

Corollary 4.16. *If $2 \leqslant k \leqslant n$, then any k-dimensional hyperbolic subspace of $\mathbb{H}^n$ is isometric to $\mathbb{H}^k$.*

Proof. The hyperplane $\{x \in \mathbb{H}^n : x_{k+1} = x_{k+2} = \cdots = x_n = 0\}$ is clearly isometric to $\mathbb{H}^k$. The claim follows from Proposition 4.15. □

Any hyperplane T in $\mathbb{M}^{1,n}$ is of the form $T = u^\perp$ for some $u \in \mathbb{M}^{1,n} - \{0\}$ because the Minkowski bilinear form is nondegenerate. Let $H = u^\perp \cap \mathbb{H}^n$ be a hyperbolic hyperplane. Since H intersects $\mathbb{H}^n$, it contains a vector v for which $\langle v \,|\, v\rangle = -1$. Proposition 4.5 implies that $\langle u \,|\, u\rangle > 0$, and after normalising, we may assume that u is a unit vector.

Let $u \in \mathscr{L}^n_+$. The *reflection* in $H = u^\perp \cap \mathbb{H}^n$ is the map

$$r_H(x) = x - 2\langle x \,|\, u\rangle\, u\,. \tag{4.4}$$

Example 4.17. If $u_0 = 0$, then $\langle x \,|\, u\rangle = (x \,|\, u)$ for all $x \in \mathbb{M}^{1,n}$. This implies that the reflection in $u^\perp$ coincides with the Euclidean reflection in the hyperplane $u^\perp$ that contains e_0.

The proofs of the basic properties of reflections are natural modifications of those in the spherical case. Note that the expression (4.4) defines a mapping in Minkowski space, fixing the hyperplane $u^\perp$. The reflection in hyperbolic space is, in fact, the restriction of a reflection of Minkowski space.

Proposition 4.18. *Let H be a hyperbolic hyperplane. Then*
(0) r_H maps $\mathbb{H}^n$ into itself.
(1) $r_H \circ r_H$ is the identity.
(2) $r_H \in O^+(1,n)$.
(3) $d(r_H(x), y) = d(x, y)$ for all $x \in \mathbb{H}^n$ and all $y \in H$.
(4) The fixed point set of r_H is H

Proof. (0) Let $x \in \mathbb{H}^n$. Using bilinearity and symmetry of the Minkowski form and the fact that u is a unit vector, we get

$$\begin{aligned}\langle r_H(x) \,|\, r_H(x)\rangle &= \big\langle x - 2\langle x \,|\, u\rangle u \,\big|\, x - 2\langle x \,|\, u\rangle u\big\rangle \\ &= \langle x \,|\, x\rangle - 2\langle x \,|\, u\rangle\langle x \,|\, u\rangle - 2\langle x \,|\, u\rangle\langle u \,|\, x\rangle + 4\langle x \,|\, u\rangle\langle x \,|\, u\rangle\langle u \,|\, u\rangle \\ &= \langle x \,|\, x\rangle = -1\,.\end{aligned}$$

Thus, $r_H(x) \in \mathscr{L}^n_-$. Furthermore, for any $v \in H$,

$$r_H(v) = v - 2\langle v \,|\, u\rangle u = v,$$

so there are points in $\mathbb{H}^n$ which are mapped to $\mathbb{H}^n$. Since r_H is continuous and preserves the Minkowski form, $r_H(\mathbb{H}^n) \subset \mathbb{H}^n$.
(1) This easy computation is left as an exercise.
(2) Clearly, r_H is a linear mapping, and it is a bijection by (1). As in (0), we get

$$\langle r_H(x) \,|\, r_H(y) \rangle = \big\langle x - 2\langle x \,|\, u \rangle u \,\big|\, y - 2\langle y \,|\, u \rangle u \big\rangle = \langle x \,|\, y \rangle .$$

Thus, $r_H \in \mathrm{O}(1,n)$. Claim (0) gives $r_H \in \mathrm{O}^+(1,n)$.
(3) For any $x \in \mathbb{H}^n$ and all $y \in H$, we have

$$\langle r_H(x) \,|\, y \rangle = \langle x - 2\langle x \,|\, u \rangle u \,|\, y \rangle = \langle x \,|\, y \rangle - 2\langle x \,|\, u \rangle \langle u \,|\, y \rangle = \langle x \,|\, y \rangle,$$

where the final equality follows from the assumption $u \in H^\perp$.
(4) This follows immediately from (3) by taking $x = y \in H$. □

The *bisector* of two distinct points p and q in $\mathbb{H}^n$ is the hyperplane

$$\mathrm{bis}(p,q) = \{x \in \mathbb{H}^n : d(x,p) = d(x,q)\} .$$

Lemma 4.19. *If $p, q \in \mathbb{H}^n$, $p \neq q$, then* $\mathrm{bis}(p,q) = (p-q)^\perp \cap \mathbb{H}^n$.

Proof. Exercise 4.5. □

Proposition 4.20. *(1) For any $p, q \in \mathbb{H}^n$, the bisector* $\mathrm{bis}(p,q)$ *is a hyperbolic hyperplane.*
(2) If H is a hyperplane in $\mathbb{H}^n$ and $x, y \in \mathbb{H}^n - H$ with $r_H(x) = y$, then $H = \mathrm{bis}(x,y)$.
(3) If $p, q \in \mathbb{H}^n$, $p \neq q$, then $r_{\mathrm{bis}(p,q)}(p) = q$.
(4) Let $\phi \in \mathrm{Isom}(\mathbb{H}^n)$, $\phi \neq \mathrm{id}$. If $a \in \mathbb{H}^n$ with $\phi(a) \neq a$, then the fixed points of ϕ are contained in $\mathrm{bis}(a, \phi(a))$.
(5) Let $\phi \in \mathrm{Isom}(\mathbb{H}^n)$, $\phi \neq \mathrm{id}$. If H is a hyperplane such that $\phi|_H$ is the identity, then $\phi = r_H$.

Proof. (1) Lemma 4.1 implies that

$$\langle p - q \,|\, p - q \rangle = -2 - 2\langle p \,|\, q \rangle > 0 .$$

Let $\lambda > 0$ and $u \in \mathscr{L}^n_+$ such that $p - q = \lambda v$. Obviously, $(p-q)^\perp = v^\perp$. The second part of Proposition 4.4 implies that there is an element $A \in \mathrm{O}^+(1,n)$ such that $Av = e_1$. The orthogonal complement of e_1 is the hyperplane $\{x \in \mathbb{M}^{1,n} : x_1 = 0\}$ that contains e_0. The claim follows as A maps $\mathbb{H}^n$ to itself and $(Av)^\perp = A(v^\perp)$.
(2) follows from Proposition 4.18(3).
(3) Using the computation from (1) above, we have

$$2\langle p \,|\, p - q \rangle = 2(\langle p \,|\, p \rangle - \langle p \,|\, q \rangle) = -2 - 2\langle p \,|\, q \rangle = |p - q|^2 .$$

Thus,

$$r_{\mathrm{bis}(p,q)}(p) = p - 2\langle p \,|\, p - q \rangle \frac{p - q}{|p - q|^2} = q .$$

(4) If $\phi(b) = b$, then $d(a,b) = d(\phi(a), \phi(b)) = d(\phi(a), b)$, so that $b \in \mathrm{bis}(a, \phi(a))$.
(5) is an instructive exercise. □

Proposition 4.21. *Any two reflections in hyperbolic hyperplanes of* $\mathbb{H}^n$ *are conjugate in* $\operatorname{Isom}\mathbb{H}^n$.

Proof. Exercise 4.7. □

Next, we prove that all isometries of hyperbolic space are restrictions to $\mathbb{H}^n$ of linear automorphisms of $\mathbb{M}^{1,n}$:

Theorem 4.22. $\operatorname{Isom}(\mathbb{H}^n) = \mathrm{O}^+(1,n)$.

The idea of the proof is to show that each isometry of $\mathbb{H}^n$ is the composition of reflections in hyperbolic hyperplanes. Again, the proof follows the same ideas as in the Euclidean and spherical cases.

Proposition 4.23. *Let* $p_1, p_2, \ldots, p_k, q_1, q_2, \ldots, q_k \in \mathbb{H}^n$ *be points that satisfy*

$$d(p_i, p_j) = d(q_i, q_j)$$

for all $i, j \in \{1, 2, \ldots, k\}$. *Then, there is an isometry* $\phi \in \operatorname{Isom}(\mathbb{H}^n)$ *such that* $\phi(p_i) = q_i$ *for all* $i \in \{1, 2, \ldots, k\}$. *Furthermore, the isometry* ϕ *is the composition of at most* k *reflections in hyperplanes.*

Proof. Exercise 4.8. □

Note that Proposition 4.23 implies that if T and T' are two triangles in $\mathbb{H}^n$ with equal sides, then there is an isometry ϕ of $\mathbb{H}^n$ such that $\phi(T) = T'$.

Proof of Theorem 4.22. Let $\phi \in \operatorname{Isom}(\mathbb{H}^n)$. Let $\{a_0, a_1, \ldots, a_n\}$ be a set of points in $\mathbb{H}^n$ which is not contained in any proper hyperbolic subspace. This is achieved by choosing them so that they generate $\mathbb{M}^{1,n}$ as a vector space. Proposition 4.23 implies that there is an isometry $\phi_0 \in \mathrm{O}^+(1,n)$ such that $\phi_0(\phi(a_i)) = a_i$ for all $0 \leqslant i \leqslant n$. Since the set of fixed points of $\phi_0 \circ \phi$ contains the points $a_0, a_1, \ldots, a_n$, the fixed point set of $\phi_0 \circ \phi$ is not contained in a proper hyperbolic subspace. Proposition 4.20(4) implies that $\phi_0 \circ \phi$ is the identity map. Thus, $\phi = \phi_0^{-1}$. In particular, $\phi \in \mathrm{O}^+(1,n)$, which is all we needed to show. □

Corollary 4.24. *Any isometry of* $\mathbb{H}^n$ *can be represented as the composition of at most* $n+1$ *reflections.* □

Proposition 4.25. *The stabilizer of any point* $x \in \mathbb{H}^n$ *is isomorphic to* $\mathrm{O}(n)$.

Proof. Exercise 4.9. □

4.5 Triangles in $\mathbb{H}^n$

The law of cosines implies that a triangle in $\mathbb{E}^n$, $\mathbb{S}^n$ or $\mathbb{H}^n$ is uniquely determined up to an isometry of the space, if the lengths of the three sides are known. In Euclidean space, the three angles of a triangle do not determine the triangle uniquely. In $\mathbb{S}^n$ and $\mathbb{H}^n$ the angles determine a triangle uniquely. For $\mathbb{H}^n$, this is the content of

Proposition 4.26 (The second hyperbolic law of cosines)**.**

$$\cosh c = \frac{\cos\alpha\cos\beta + \cos\gamma}{\sin\alpha\sin\beta}\,.$$

Proof. This formula follows from the first law of cosines by a lengthy manipulation analogous to the proof of Proposition 3.5. See for example [Bea, p. 148–150]. □

The second law of cosines and Proposition 4.23 imply that if T and T' are two triangles in $\mathbb{H}^n$ with equal sides, then there is an isometry ϕ of $\mathbb{H}^n$ such that $\phi(T) = T'$.

Proposition 4.27 (The hyperbolic law of sines)**.**

$$\frac{\sinh a}{\sin\alpha} = \frac{\sinh b}{\sin\beta} = \frac{\sinh c}{\sin\gamma}\,.$$

Proof. The first law of cosines implies that

$$\left(\frac{\sinh c}{\sin\gamma}\right)^2 = \frac{\sinh^2 a\sinh^2 b\sinh^2 c}{2\cosh a\cosh b\cosh c - \cosh^2 a - \cosh^2 b - \cosh^2 c + 1}\,.$$

The claim follows because this expression is symmetric in a, b and c. □

The following two results on triangles will be useful later.

Proposition 4.28. *For any $0 < a, b, c$ for which $a + b > c$, $b + c > a$ and $c + a > b$, there is a triangle with side lengths a, b and c. Any two such triangles are isometric.*

Proof. The proof is analogous with that of Proposition 3.20 without the upper bound on the lengths. We use the hyperbolic law of cosines in the construction. If a triangle with the asserted properties exists, then the angle at C satisfies the cosine law. Therefore, we can compute what this angle needs to be if we know that

$$\left|\frac{\cosh a\cosh b - \cosh c}{\sinh a\sinh b}\right| < 1\,. \tag{4.5}$$

The inequality $c < a + b$ implies

$$\cosh c < \cosh(a + b) = \cosh a\cosh b + \sinh a\sinh b\,,$$

which gives

$$-1 < \frac{\cosh a\cosh b - \cosh c}{\sinh a\sinh b}\,.$$

The inequalities $b + c > a$ and $c + a > b$ give $|a - b| < c$, which implies

$$\cosh c > \cosh(a - b) = \cosh a\cosh b + \sinh a\sinh b\,,$$

and we get

$$\frac{\cosh a\cosh b - \cosh c}{\sinh a\sinh b} < 1\,.$$

Now we can place the sides of length a and b starting at C in the correct angle γ. The cosine law implies that the distance of the endpoints points A and B of these segments is c. There geodesic arc from A to B is therefore the side opposite to C of the desired length c.

The triangles are isometric by Proposition 4.23. □

Proposition 4.29. *Any triangle in $\mathbb{H}^n$ is contained in an isometrically embedded copy of $\mathbb{H}^2$ in $\mathbb{H}^n$.*

Proof. Any three points in the hyperboloid model $\mathbb{H}^n$ are contained in the intersection of $\mathbb{H}^n$ with a 3-dimensional linear subspace of $\mathbb{M}^{1,n}$, which is an isometrically embedded copy of the hyperbolic plane. The geodesic arc through any two of these points in is contained in the same hyperbolic 2-plane by Lemma 4.9. □

Using the hyperbolic law of cosines and the Taylor polynomials of hyperbolic functions at 0, $\cosh t = 1 + \frac{t^2}{2} + \mathrm{o}(t^2)$ and $\sinh t = t + \mathrm{o}(t)$, we see that if the sides of a triangle in hyperbolic space are short, then the sides satisfy the Euclidean law of cosines up to a small error.

Exercises

4.1. Prove Lemma 4.2.

4.2. Prove that $\mathscr{L}^2_+$ is the $\mathrm{O}^+(1,2)$-orbit of $e_1 \in \mathbb{M}^{1,2}$.[5]

4.3. Prove Lemma 4.7.

4.4. Prove Proposition 4.15.

4.5. Prove Lemma 4.19.

4.6. Prove Proposition 4.20(5).

4.7. Prove Proposition 4.21.[6]

4.8. Prove Proposition 4.23.[7]

4.9. Prove Proposition 4.25.[8]

[5]See Proposition 4.4.
[6]Use Proposition 4.15.
[7]The proof is formally exactly the same as that of Proposition 2.13.
[8]Follow the proof of Proposition 3.14. Assume that we know $\mathrm{Isom}\,\mathbb{H}^n = \mathrm{O}^+(1,n)$ and use transitivity.

Chapter 5

Models of hyperbolic space

The hyperboloid model of hyperbolic space introduced in chapter 4 model is used in many arithmetical applications and the closely related projective model has important generalizations to complex and quaternionic hyperbolic spaces.

In this chapter, we consider a number of other models for hyperbolic space. Hyperbolic space of dimension n is the class of all metric spaces isometric with the hyperboloid model $(\mathbb{H}^n, d)$, and we can use any model that is best suited for the geometric problem at hand. After this section we will often talk about the "upper halfplane model of $\mathbb{H}^2$" etc.

The underlying set of the Klein model and the Poincaré model is the unit ball in Euclidean space. Therefore, we introduce a special notation for this set:

$\mathbb{B}^n$ is the unit ball in $\mathbb{E}^n$.

In sections 5.2, 5.3 and 5.4, we use the geometric properties of inversions in spheres. We refer to Appendix A for details on inversions.

5.1 Klein's model

Each line in $\mathbb{M}^{1,n}$ through the origin which intersects the hyperboloid model $\mathbb{H}^n$, intersects it in exactly one point, and it also intersects the embedded copy $\{1\} \times \mathbb{B}^n$ in $\mathbb{M}^{1,n}$ of $\mathbb{B}^n$ in exactly one point. This correspondence determines a bijection $K \colon \mathbb{B}^n \to \mathbb{H}^n$, which has the explicit expression

$$K(x) = \frac{(1, x)}{\sqrt{1 - \|x\|^2}} .$$

The map K becomes an isometry when we define a metric on $\mathbb{B}^n$ by setting

$$d_K(x, y) = d(K(x), K(y)) = \operatorname{arcosh} \frac{1 - (x \,|\, y)}{\sqrt{1 - \|x\|^2}\sqrt{1 - \|y\|^2}} .$$

The metric space $(\mathbb{B}^n, d_K)$ is the *Klein model* of n-dimensional hyperbolic space.

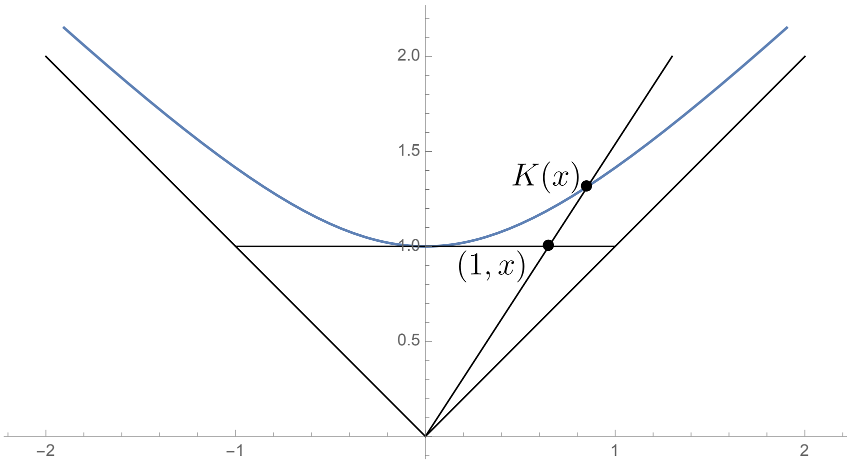


Figure 5.1 — The map K used in the construction of the Klein model.

Proposition 5.1. *The images of geodesic lines of the Klein model are the Euclidean open segments connecting two points in the Euclidean unit sphere.*

Proof. Geodesic lines in $\mathbb{H}^n$ are the intersections of $\mathbb{H}^n$ with 2-planes in $\mathbb{M}^{1,n}$ by Lemma 4.9. The intersection of such a plane with $\{1\} \times \mathbb{B}^n$ is the preimage under K of the geodesic line. Conversely, any line in $\{1\} \times \mathbb{R}^n$ is the intersection of a 2-plane with $\{1\} \times \mathbb{R}^n$. Such a plane intersects $\mathbb{H}^n$ in the image of a geodesic line if and only if the 2-plane intersects the Klein model. ☐

Corollary 5.2. *(1) For any two distinct points $a, b \in \mathbb{S}^{n-1} = \partial\mathbb{B}^n$, there is a unique image of a geodesic line $]a, b[$ in the Klein model.*

(2) If $x_0 \in \mathbb{B}^n$ and $b \in \partial\mathbb{B}^n$, there is a unique geodesic ray $\rho_{x_0,b} \colon [0, \infty[\to \mathbb{B}^n$ in the Klein model of $\mathbb{H}^n$ such that $\rho_{x_0,b}(0) = x_0$ and such that the Euclidean closure of the image $\rho_{x_0,b}([0, \infty[) = [x_0, b[$ is the Euclidean closed segment $[x_0, b]$. ☐

We call $]a, b[$ the *geodesic line with endpoints a and b in the Klein model of* $\mathbb{H}^n$.

If $\gamma \colon \mathbb{R} \to \mathbb{H}^n$ is a geodesic line and $T \in \mathbb{R}$, then the mapping $t \overset{\gamma_T}{\mapsto} \gamma(t - T)$ defined on $\mathbb{R}$ is a geodesic line such that $\gamma(\mathbb{R}) = \gamma_T(\mathbb{R})$.

Recall that in Euclidean plane geometry, two (geodesic) lines are *parallel* if they do not intersect. The *parallel axiom* states that through any point P in the Euclidean plane that is not contained in a line L, there is exactly one line that is parallel with L. It easy to see using the Klein model that the parallel axiom does not hold in $\mathbb{H}^2$, see Figure 5.2

5.2 Poincaré's ball model

Each affine line that passes through the point $(-1, 0) \in \mathbb{R} \times \mathbb{R}^n = \mathbb{M}^{1,n}$ which intersects $\mathbb{H}^n$, intersects it in exactly one point, and it also intersects the n-dimensional ball $\{0\} \times \mathbb{B}^n$ embedded in $\mathbb{M}^{1,n}$ in exactly one point. This correspondence determines a bijection $P \colon \mathbb{B}^n \to \mathbb{H}^n$,

$$P(x) = \frac{(1 + \|x\|^2, 2x)}{1 - \|x\|^2} \, .$$

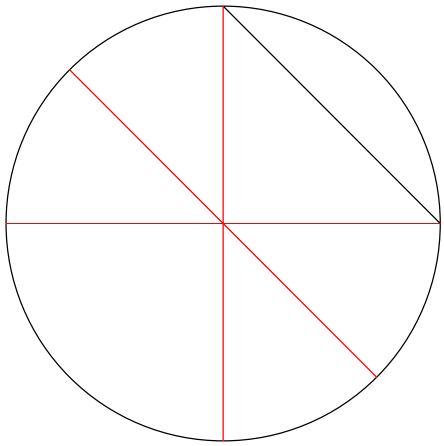


Figure 5.2 — Three red lines through the origin that are parallel in the Klein model with the line whose endpoints are $(0, 1)$ and $(1, 0)$.

This expression is found by computing for any $x \in \mathbb{B}^n$ that the point $y_t = (0, x) + t(1, x)$ on the line through the points $(0, x)$ and $(-1, 0)$ of $\mathbb{R} \times \mathbb{R}^n = \mathbb{M}^{1,n}$ is in $\mathbb{H}^n$ if and only if $t = \frac{1+\|x\|^2}{1-\|x\|^2}$.

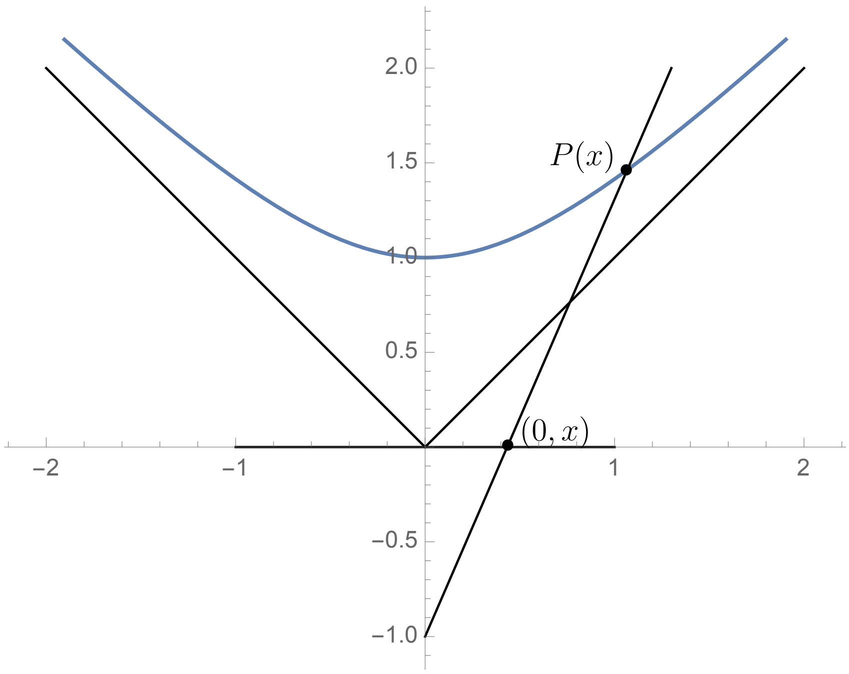


Figure 5.3 — The map P used in the construction of the Poincaré model.

The map P becomes an isometry when we define a metric on $\mathbb{B}^n$ by setting

$$d_P(x, y) = d(P(x), P(y)) = \operatorname{arcosh}\left(1 + 2\,\frac{\|x - y\|^2}{(1 - \|x\|^2)(1 - \|y\|^2)}\right).$$

The metric space $(\mathbb{B}^n, d_P)$ is the *Poincaré model* of n-dimensional hyperbolic space.

Lemma 5.3. *The hyperbolic ball of radius $r > 0$ centered at 0 in the Poincaré model coincides with the Euclidean ball of radius $\tanh \frac{r}{2}$ centered at 0. The Euclidean ball of radius $0 < R < 1$ centered at 0 coincides with the hyperbolic ball of radius $\log \frac{1+R}{1-R}$ centered at 0 in the Poincaré model.*

Proof. If $x \in \mathbb{B}^n$, we have

$$d_P(x, 0) = \operatorname{arcosh}\left(1 + 2\frac{\|x\|^2}{1 - \|x\|^2}\right) = \log\frac{1 + \|x\|}{1 - \|x\|}.$$

Both claims follow from this equation. □

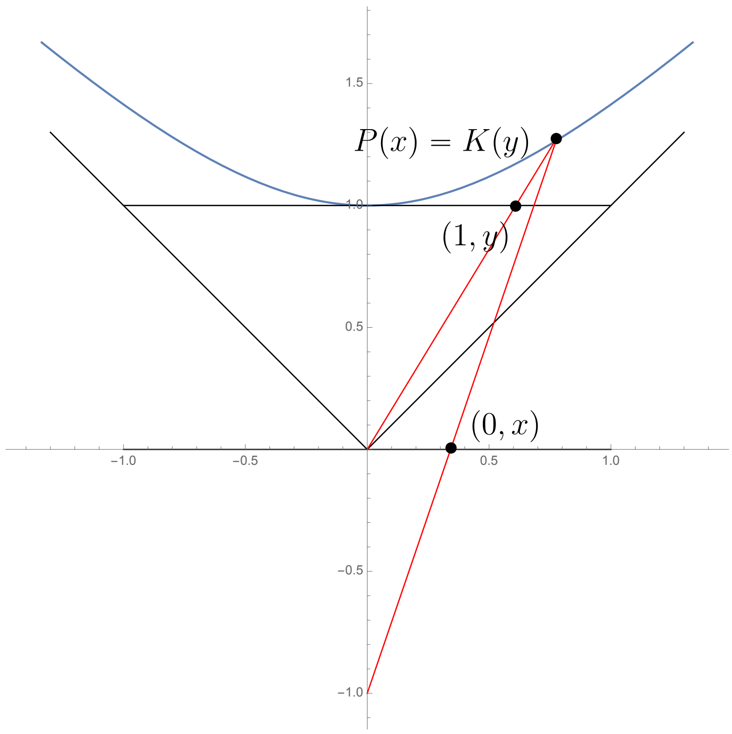


Figure 5.4 — The construction of the map h from the Poincaré model to the Klein model.

Proposition 5.4. *The images of geodesic lines of the Poincaré model are the intersections of the Euclidean unit ball with Euclidean circles and lines that are orthogonal to the unit sphere.*

Proof. The map $h = K^{-1} \circ P$ is an isometry between the Poincaré and Klein models. A computation[1] shows that

$$h(x) = \frac{2x}{1 + \|x\|^2}.$$

[1]This can be done by observing that h is a radial map and then solving the equation

$$\frac{(1, y)}{\sqrt{1 - y^2}} = \left(\frac{1 + x^2}{1 - x^2}, \frac{2x}{1 - x^2}\right)$$

with $0 \leqslant x, y < 1$.

The inversion $\iota_{(-1,0),2}$ in the sphere centered at $(-1,0) \in \mathbb{E}^1 \times \mathbb{E}^n$ of radius $\sqrt{2}$ maps $\{0\} \times \mathbb{E}^n \cup \{\infty\}$ to $\mathbb{S}^n$. It maps $\{0\} \times \mathbb{B}^n \cup \{\infty\}$ to the upper hemisphere of $\mathbb{S}^n$, fixing $\{0\} \times \mathbb{S}^{n-1}$. In coordinates,

$$\iota_{(-1,0),2}(x) = \left(\frac{1-\|x\|^2}{1+\|x\|^2}, \frac{2x}{1+\|x\|^2}\right),$$

so that if $\mathrm{pr}\colon \mathbb{E}^{n+1} = \mathbb{E}^1 \times \mathbb{E}^n \to \mathbb{E}^n$ is the Euclidean orthogonal projection on the second component of the product, we have

$$h = \mathrm{pr} \circ \iota_{(-1,0),2}\,.$$

The inversion $\iota_{(-1,0),2}$ maps any circle in $\{0\} \times \mathbb{B}^n$ orthogonal to $\{0\} \times \mathbb{S}^{n-1}$ to a circle on the unit sphere in $\mathbb{E}^{n+1}$ orthogonal to $\{0\} \times \mathbb{S}^{n-1}$. These circles are orthogonal to $\{0\} \times \mathbb{E}^n$, and they are exactly the intersections of the unit sphere with 2-planes parallel to the x_0-axis, and thus, pr maps them to the geodesic lines of the Klein model. As h is an isometry, the result follows. □

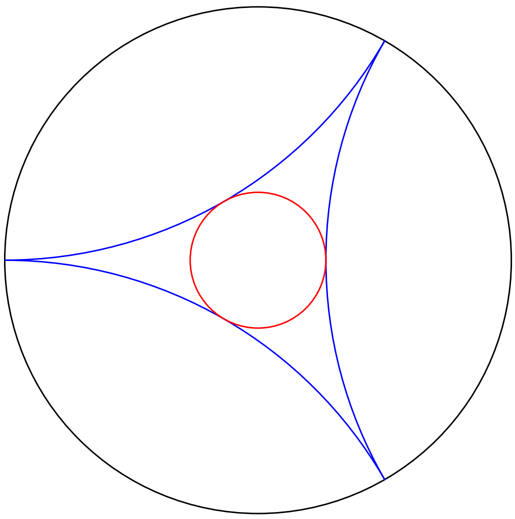


Figure 5.5 — Some geodesic lines and a ball in the Poincaré disk model of $\mathbb{H}^2$.

Note that the mapping h from the Klein model to the Poincaré model is the restriction of a homeomorphism of the Euclidean closure of $\mathbb{B}^n$ to itself. This extended mapping is the identity in the boundary of $\mathbb{B}^n$.

Corollary 5.5. *(1) For any two distinct points $a, b \in \mathbb{S}^{n-1} = \partial\mathbb{B}^n$, there is geodesic line $]a, b[$ in the Poincaré model that we call the* geodesic line with endpoints a and b in the Poincaré model of $\mathbb{H}^n$.

(2) If $x_0 \in \mathbb{B}^n$ and $b \in \partial\mathbb{B}^n$, there is a unique geodesic ray $\rho_{x_0,b}\colon [0,\infty[\ \to \mathbb{B}^n$ in the Poincaré model of $\mathbb{H}^n$ such that $\rho_{x_0,b}(0) = x_0$ and such that the Euclidean closure of the image $\rho_{x_0,b}([0,\infty[) = [x_0, b[$ is a closed Euclidean segment or a closed circular segment with one endpoint at b. □

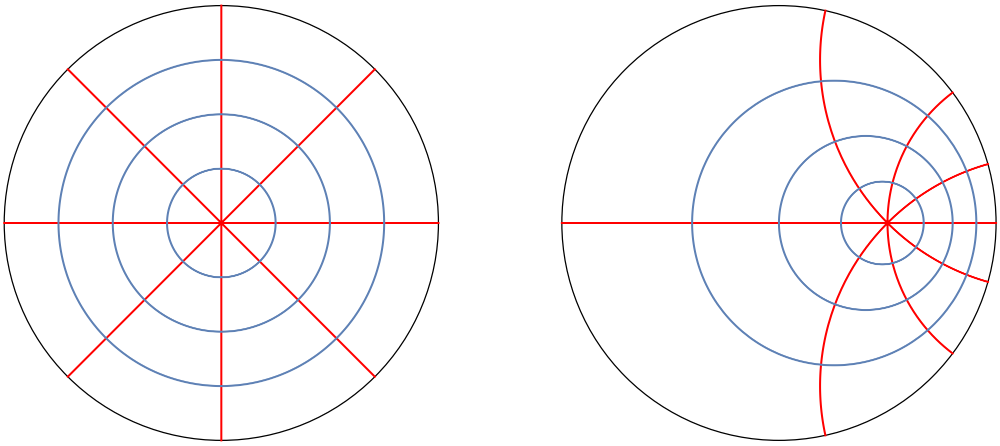


Figure 5.6 — Geodesic rays starting at 0 and at $(\frac{1}{2}, 0)$ with circles centered at the same points in the Poincaré disk model of $\mathbb{H}^2$.

Proposition 5.6. *The Riemannian metric of the ball model is* $\frac{4(\cdot\,|\,\cdot)}{(1-\|x\|^2)^2}$.

Proof. For all tangent vector $u \in T_x B(0,1)$, we have

$$DP(x)u = \left(\frac{4(x\,|\,u)}{(1-\|x\|^2)^2}, \frac{2u}{1-\|x\|^2} + \frac{4\,(x\,|\,u)x}{(1-\|x\|^2)^2}\right) \in \mathbb{M}^{1,n}.$$

Using this, for $u, v \in T_x\mathbb{B}^n$, we compute

$$\begin{aligned}\langle DP(x)u \,|\, DP(x)v\rangle &= -\frac{16(x\,|\,u)(x\,|\,v)}{(1-\|x\|^2)^4} + \frac{4(u\,|\,v)}{(1-\|x\|^2)^2} + \frac{16(x\,|\,u)(x\,|\,v)}{(1-\|x\|^2)^3} + \frac{16(x\,|\,u)(x\,|\,v)\|x\|^2}{(1-\|x\|^2)^4} \\ &= \frac{4(u\,|\,v)}{(1-\|x\|^2)^2}.\end{aligned}$$

□

Proposition 5.6 implies that the angles between tangent vectors of paths in the Poincaré model are the same as the angles measured in the ambient Euclidean space. The Klein model does not have this useful property. This is illustrated in Figure 5.7

5.3 The upper halfspace model

Let

$$\mathbb{R}^n{}_+ = \{x \in \mathbb{R}^n : x_n > 0\}$$

be the n-dimensional *upper halfspace*. Let $\iota_{-e_n,2}$ be the inversion in the sphere of center $-e_n \in \mathbb{E}^n$ of radius $\sqrt{2}$. The map

$$F = (\iota_{-e_n,2})|_{\mathbb{B}^n} : \mathbb{B}^n \to \mathbb{R}^n{}_+ \tag{5.1}$$

is a bijection, which becomes an isometry if we use the metric

$$d_{\mathbb{R}^n{}_+}(x,y) = d_P(F^{-1}(x), F^{-1}(y)) = \operatorname{arcosh}\left(1 + \frac{\|x-y\|^2}{2\,x_n y_n}\right) \tag{5.2}$$

in $\mathbb{R}^n{}_+$.

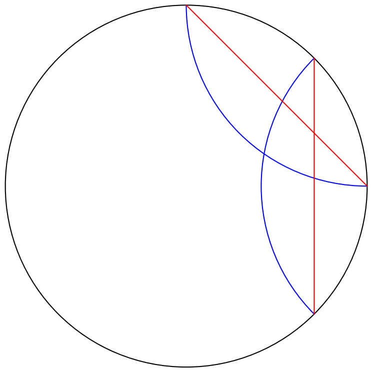


Figure 5.7 — The blue geodesic lines of the Poincaré model in this figure are the images of the red geodesic lines of the Klein model. The angles at the points of intersection are the same in hyperbolic plane but the angle in the ambient Euclidean space of the red lines is not the same as that of the blue circular segments.

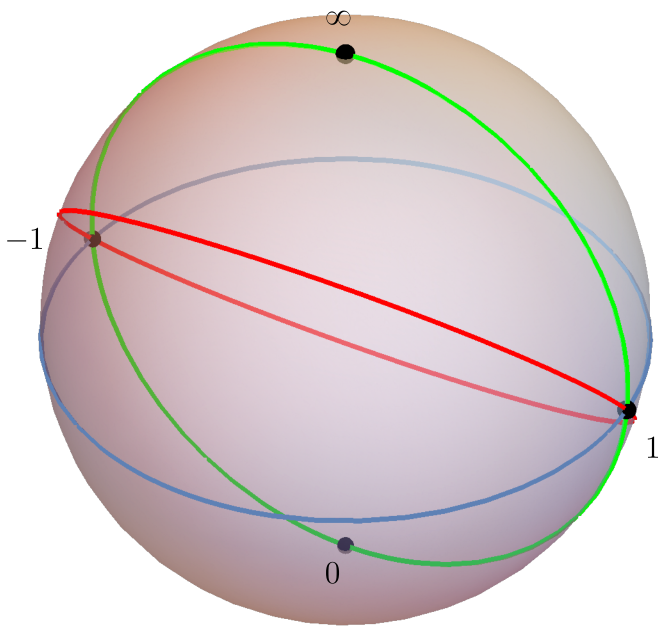


Figure 5.8 — The mapping F corresponds to the reflection in the red circle when $\widehat{\mathbb{E}}^2$ is identified with $\mathbb{S}^2$ by the stereographic projection. See section 3.3 and Appendix A.

The metric space $(\mathbb{R}^n{}_+, d_{\mathbb{R}^n{}_+})$ is the *upper halfspace model* of n-dimensional hyperbolic space.

Example 5.7. An elementary computation shows that if $x = (a, x_n)$ and $y = (a, y_n)$ for any $a \in \mathbb{R}^{n-1}$, then

$$d_{\mathbb{R}^n{}_+}(x, y) = \left| \log \frac{x_n}{y_n} \right| .$$

It is very common to identify the upper halfplane model of $\mathbb{H}^2$ with the upper halfplane in $\mathbb{C}$, and we will often do this, as in Example 5.8(2) below.

Example 5.8. (1) Let $n \geqslant 3$. The subspace $\{x \in \mathbb{R}^n_+ : x_2 = \cdots = x_{n-1} = 0\}$ with the metric induced from the upper halfplane model is an isometrically embedded copy of $\mathbb{H}^2$ in the upper halfspace model of $\mathbb{H}^n$.

(2) Let $0 < \phi < \pi$. Then the distance of the points i and $e^{i\phi}$ in the upper halfplane model is

$$d_{\mathbb{R}^2{}_+}(i, e^{i\phi}) = \operatorname{arcosh}\left(1 + \frac{\cos^2\phi + (1-\sin\phi)^2}{2\sin\phi}\right) = \operatorname{arcosh}\frac{1}{\sin\phi}\,.$$

Proposition 5.9. *The images of the geodesic lines of the upper halfspace model are the intersections of the upper halfspace with Euclidean circles and lines that are orthogonal to* $\mathbb{E}^{n-1} \times \{0\}$.

Proof. The inversion used in the definition of the upper halfspace model maps lines and circles to lines or circles and preserves angles. The claim follows from Proposition 5.4. □

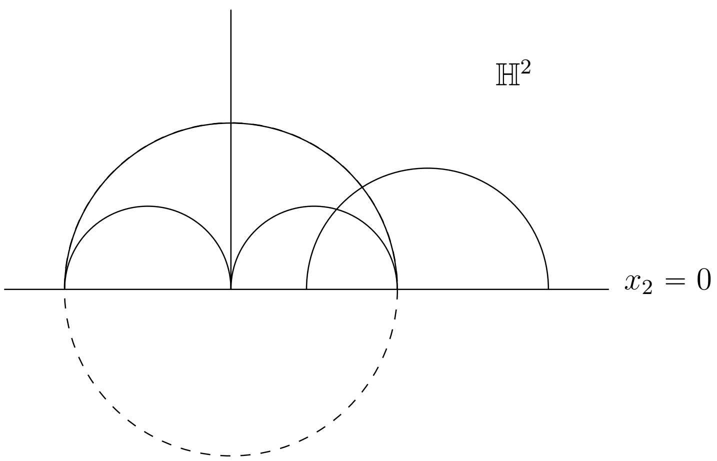


Figure 5.9 — Some geodesic lines in the upper halfplane model of $\mathbb{H}^2$.

Geodesic lines in the upper halfspace model are images under F of geodesic lines of the Poincaré model. If one of the endpoints of a geodesic line in the Poincaré model is $-e_n$, then F maps this geodesic line to a halfline orthogonal to $\mathbb{E}^{n-1} \times \{0\}$ at one end, and the other endpoint is mapped to $\infty \in \widehat{\mathbb{E}}^n$.

Corollary 5.10. *For any two distinct points* $a, b \in \mathbb{E}^{n-1} \times \{0\} \cup \{\infty\}$, *there is geodesic line* $]a, b[$ *in the upper halfspace model with endpoints* a *and* b. □

We have seen that the unit sphere in the Klein and Poincaré ball models and the set $\mathbb{E}^{n-1} \times \{0\} \cup \{\infty\} \subset \widehat{\mathbb{E}^n}$ in the upper halfspace model have a geometric meaning, and that there is a natural homeomorphism between these sets. In chapter **??**, we will see that these sets appear naturally as a geometrically defined *boundary at infinity* of $\mathbb{H}^n$, and we will use the notation $\partial_\infty\mathbb{H}^n$ for this set from now on.

In practical applications, it is good to remember that a circle is perpendicular to $\mathbb{E} \times \{0\} \subset \mathbb{E}^2$ if and only if its center is in $\mathbb{E} \times \{0\}$. In higher dimensions, this is no longer true.

The following lemma records the expressions of the geodesics in the upper halfspace.

Lemma 5.11. *Let $x \in \mathbb{R}^{n-1}$ and $y > 0$. The mapping $\gamma_{x,y} \colon \mathbb{R} \to \mathbb{R}^n_+$,*

$$\gamma_{x,y}(t) = (x, ye^t)$$

is a geodesic line in the upper halfspace model of $\mathbb{H}^n$ such that $\gamma_{x,y}(0) = (x, y)$ and with endpoints x an ∞. For any isometry g of the upper halfspace model, the mapping $g \circ \gamma_{x,y}$ is a geodesic line.

Proof. The mapping $\gamma_{x,y} \colon \mathbb{R} \to \mathbb{R}^n_+$ is a geodesic line by Example 5.7. □

Proposition 5.12. *The Riemannian metric of the upper halfspace model is $\frac{(\cdot\,|\,\cdot)}{x_n^2}$.*

Proof. The proof is similar to that of Proposition 5.6, using (the inverse of) the map F defined in equation (5.1) to transfer the Riemannian metric from the ball to the upper halfspace. Note that $F \circ F = \mathrm{id}$. As in the proof of Proposition 5.4, we compute

$$DF(x)u = -\frac{4(x + e_n)(x + e_n \,|\, u)}{\|x + e_n\|^4} + \frac{2u}{\|x + e_n\|^2}\,.$$

The claim follows because

$$(1 - \|F(x)\|^2)^2 = \frac{x_n^2}{(1 - \|x\|^2)^2}$$

and

$$\begin{aligned}
&(DF(x)u \,|\, DF(x)v)\\
&\quad= \frac{16\|x + e_n\|^2(x + e_n \,|\, u)(x + e_n \,|\, v)}{\|x + e_n\|^6} - \frac{16(x + e_n \,|\, u)(x + e_n \,|\, v)}{\|x + e_n\|^4} + \frac{4(u \,|\, v)}{\|x + e_n\|^4}\\
&\quad= \frac{4(u \,|\, v)}{\|x + e_n\|^4}\,.
\end{aligned}$$

□

Proposition 5.12 implies that the angles between tangent vectors of paths in the upper halfspace model are the same as the angles measured in the ambient Euclidean space.

Proof of Theorem 4.12. We will use the upper halfspace model to prove the result. Both quantities are invariant under isometries of hyperbolic space. Therefore, it is sufficient to show that the geodesic segment $[(0, 1), (0, T)]$ is the Riemannian geodesic segment from $(0, 1)$ to $(0, T)$ for any $T > 0$.

Let $\phi\colon [0, 1] \to \mathbb{H}^n$ be a piecewise smooth path such that $\phi(0) = (0, 1)$ and $\phi(1) = (0, T)$. [2] Let $p\colon \mathbb{H}^n \to [0, 1]$,

$$p(x, s) = (0, s)$$

[2] We can assume that all paths are defined on $[0, 1]$ because smooth reparametrization does not change the Riemannian length of a path, see for example [Pet, Section 5.3].

for all $x \in \mathbb{R}^{n-1}$ and $s > 0$, be the *horospherical projection* to the geodesic line $]0, \infty[$ that contains the points $(0,1)$ to $(0,T)$. Note that $Dp(x,s)u = u_n$ for all $(x,s) \in \mathbb{H}^n$ and all $u \in \mathbb{R}^n$. This implies that $\big|(p \circ \phi)^{\cdot}(\tau)\big| \leqslant |\dot{\phi}(\tau)|$ for all $\tau \in [0,1]$.[3] This gives the inequality we want:

$$\ell(\phi) = \int_0^1 \frac{|\dot{\phi}(\tau)|}{\phi_n(\tau)}\, d\tau \geqslant \int_0^1 \frac{|(p \circ \phi)^{\cdot}(\tau)|}{(p \circ \phi)_n(\tau)}\, d\tau \geqslant \log(p \circ \phi(1)) = \log T = d((0,1),(0,T))\,.$$

Note that the second inequality is strict if the mapping $t \mapsto \phi_n(t)$ is not monotonous.

If $\gamma_{0,1}$ is the geodesic line of Lemma 5.11, $\gamma_{0,1}(0) = (0,1)$, $\gamma_{0,1}(\log T) = (0,T)$ and

$$\ell(\gamma_{0,1}|_{[0,\log T]}) = \int_0^{\log T} \frac{|\dot{\gamma}(t)|}{\gamma_n(t)}\, dt = \int_0^{\log T} \frac{ye^t}{ye^t}\, dt = \log T\,.$$

This completes the proof. □

5.4 Isometries of the upper halfspace model

In the upper halfspace model, it is often convenient to move a geodesic line by an isometry such that the endpoints of the geodesic in the model are 0 and ∞. The following results on isometries allow to do that and a bit more. We illustrate the utility of the transitivity properties of the group of isometries in Proposition 5.16 and its corollaries, and in Lemma 5.22.

Let $b \in \mathbb{R}^{n-1} \times \{0\} \subset \mathbb{R}^n$. The mapping $T_b \colon \mathbb{R}^n_+ \to \mathbb{R}^n_+$,

$$T_b(x) = x + b\,,$$

is a *horizontal translation* by b.
Let $\lambda > 0$. The mapping $L_\lambda \colon \mathbb{R}^n_+ \to \mathbb{R}^n_+$,

$$L_\lambda(x) = \lambda x\,,$$

is a *dilation by factor* λ.
Let $Q_0 \in \mathrm{O}(n-1)$ and let us use the notation $x = (\bar{x}, x_n)$. The mapping $Q \colon \mathbb{R}^n_+ \to \mathbb{R}^n_+$,

$$Q(\bar{x}, x_n) = (Q_0(\bar{x}), x_n)\,,$$

is an *orthogonal mapping around the geodesic line* $]0,\infty[$.

Lemma 5.13. *Let $a, b \in \mathbb{R}^{n-1} \times \{0\} \subset \mathbb{R}^n$ and let $\lambda > 0$.*
(1) $T_b \circ \iota_{a,r^2} \circ T_{-b} = \iota_{a+b,r^2}$.
(2) $L_\lambda \circ \iota_{0,r^2} \circ L_{\frac{1}{\lambda}} = \iota_{0,(\lambda r)^2}$.

Proof. Exercise 5.5. □

Proposition 5.14. *The maps*

[3] $\dot{f}$ is the notation we use for the derivative vector of a path f.

- *T_b for any $b \in \mathbb{R}^{n-1} \times \{0\} \subset \mathbb{R}^n$,*
- *ι_{a,r^2}, for any $a \in \mathbb{R}^{n-1} \times \{0\} \subset \mathbb{R}^n$ and any $r > 0$,*
- *L_λ for any $\lambda > 0$, and*
- *Q for any $Q_0 \in \mathrm{O}(n-1)$*

are isometries of the upper halfspace model.

Proof. Let us consider the inversion in the Euclidean unit sphere. It preserves all affine rays from a, so it preserves the upper halfspace. To prove that its restriction to $\mathbb{H}^n$ is an isometry, equation (5.2) implies that it is enough to show that the expression $\frac{\|x-y\|^2}{x_n y_n}$ is invariant under the inversion. Let us compute:

$$\frac{\iota_{0,1}(x) - \iota_{0,1}(y)}{r^2} = \frac{x}{\|x\|^2} - \frac{y}{\|y\|^2} = \frac{x\|y\|^2 - y\|x\|^2}{\|x\|^2\|y\|^2} ,$$

which gives

$$\frac{\|\iota_{0,1}(x) - \iota_{0,1}(y)\|^2}{\iota_{0,1}(x)_n \iota_{0,1}(y)_n} = \frac{\frac{\|x\|^2\|y\|^4 - 2(x \mid y)\|x\|^2\|y\|^2 + \|x\|^4\|y\|^2}{\|x\|^4\|y\|^4}}{\frac{x_n y_n}{\|x\|^2\|y\|^2}} = \frac{\|x-y\|^2}{x_n y_n} .$$

The rest of the computations is done in Exercise 5.6. □

Corollary 5.15. *The subgroup of* $\mathrm{Isom}(\mathbb{H}^n)$ *generated by dilations fixing* 0 *and horizontal translations acts transitively on the upper halfspace model of* $\mathbb{H}^n$*.*

Proof. If x is in the upper half plane,

$$T_{-(x_1,x_2,\dots x_{n-1},0)}(x) = (0,\dots,x_n) = L_{x_n} e_n .$$

Thus,

$$x = T_{(x_1,x_2,\dots x_{n-1},0)} \circ L_{x_n} e_n . \qquad \square$$

We will now apply the transitivity of the action of the group of isometries and of suitable subgroups to geometric and topological questions.

Proposition 5.16. *Balls in the upper halfspace model and in the Poincaré ball model are Euclidean balls in the Euclidean space that contains the model.*

Proof. By Lemma 5.3, balls centered at the origin of the Poincaré ball model are Euclidean balls. The inversion that maps the ball model to the upper halfspace model is an isometry, and on the other hand it preserves generalized spheres. Thus, the images of the balls centered at the origin are hyperbolic and Euclidean balls. The hyperbolic center of these balls can be mapped to any other point in $\mathbb{H}^n$ by one of the isometries of Corollary 5.15. These mappings preserve spheres, which implies that all balls in the upper halfspace model are Euclidean balls. The rest of the claim follows by one more application of the inversion that maps the ball model to the upper halfspace model. □

Corollary 5.17. *Hyperbolic space* $\mathbb{H}^n$ *is homeomorphic with the open unit ball of* $\mathbb{E}^n$*.*

Proof. The identity map from the Poincaré model to $\mathbb{B}^n \subset \mathbb{E}^n$ with the induced metric is a homeomorphism by Proposition 5.16. □

Corollary 5.18. *Hyperbolic space $\mathbb{H}^n$ is a proper metric space.* □

Our study of the Klein, Poincaré and upper halfspace models of hyperbolic space suggest that it makes sense to *compactify* hyperbolic space by adding the boundary at infinity.

We will now consider $\mathbb{H}^n \cup \partial_\infty \mathbb{H}_n$ with the topology induced by the embedding of the Poincaré model and its boundary in Euclidean space.

We will see in Example **??** that this choice of topology is mathematically natural.

Proposition 5.19. *Let x_1, x_2, x_3 and y_1, y_2, y_3 be two triples of distinct points in the boundary at infinity of $\mathbb{H}^n$. There is an isometry of $\mathbb{H}^n$ which is the restriction of a homeomorphism g of $\mathbb{H}^n \cup \partial_\infty \mathbb{H}^n$ to itself such that $g(x_i) = y_i$ for all $i \in \{1, 2, 3\}$.*

Proof. Let us consider the question in the upper halfspace model. The mappings given in Proposition 5.14 are clearly continuous mappings of $\widehat{\mathbb{E}}^n$ to itself.

It suffices to show that we can use a combination of these isometries to map x_1, x_2, x_3 to $\infty, 0, (1, 0, \ldots, 0)$. If all points x_1, x_2, x_3 are finite, map x_1 by a translation T_{-x_1} to 0 and then by the inversion ι to ∞. Relabel $\iota \circ T_{-x_1}(x_2)$ and $\iota \circ T_{-x_1}(x_3)$ to x_2 and x_3. Map x_2 to 0 by a translation. This map keeps ∞ fixed. Map x_3 (again relabeled) to the unit sphere by a dilation and then to $(1, 0, \ldots, 0)$ by the extension of an orthogonal map of $\mathbb{E}^{n-1}$. These two maps fix ∞ and 0. □

Proposition 5.20. *Let $x, y \in \mathbb{H}^n$ and $a, b \in \partial_\infty \mathbb{H}^n$. There is an isometry of $\mathbb{H}^n$ which is the restriction of a homeomorphism g of $\mathbb{H}^n \cup \partial_\infty \mathbb{H}^n$ to itself such that $g(x) = y$ and $g(a) = b$.*

Proof. Exercise 5.7. □

In the proofs of Propositions 5.19 and 5.20, we used explicit isomorphisms of the upper half plane model that are restrictions of homeomorphic self-maps of $\mathbb{H}^n \cup \partial_\infty \mathbb{H}^n$. In fact, there is a result that generalizes this observation to all isometries:

Theorem 5.21. *The isometries of $\mathbb{H}^n$ are restrictions of homeomorphic self-maps of $\mathbb{H}^n \cup \partial_\infty \mathbb{H}^n$.*

Proof. We could prove this by showing that all reflections in hyperbolic hyperplanes have this property, and then using the fact that reflections generate $\operatorname{Isom} \mathbb{H}^n$. The proof relies on showing that in the upper halfplane model, reflections in hyperbolic hyperplanes are either conjugates of the map Q of Proposition 5.14 with Q_0 a hyperplane reflection in $\mathbb{E}^{n-1}$, or inversions. □

For any $r > 0$, the *r-neighbourhood* of any nonempty subset $A \subset \mathbb{H}^n$ is

$$\mathscr{N}_r(A) = \{x \in \mathbb{H}^n : d(x, A) < r\}.$$

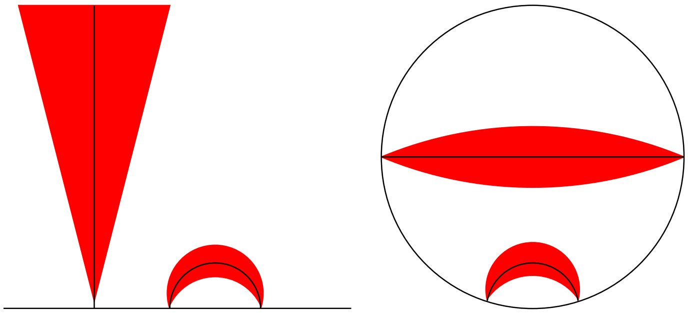


Figure 5.10 — Neighbourhoods of geodesic lines in the upper halfplane model and in the Poincaré ball model of $\mathbb{H}^2$.

Lemma 5.22. *Let* $L =]0, \infty[$ *in the upper halfspace model of* $\mathbb{H}^n$.
(1) $(0, \|x\|) \in \mathbb{R}^{n-1} \times \mathbb{R}_+$ *is the unique closest point to* $x \in \mathbb{R}^n_+$ *in* L.
(2) The r*-neighbourhood of* L *is the Euclidean infinite cone*[4]

$$\mathcal{N}_r(L) = \left\{ x \in \mathbb{R}^n_+ : \cos \measuredangle_0(L, x) > \frac{1}{\cosh r} \right\}.$$

Proof. (1) The function

$$\begin{aligned} t \mapsto \cosh d(x, \gamma_{0,\|x\|}(t)) &= 1 + \frac{x_1^2 + x_2^2 + \cdots + x_{n-1}^2 + (x_n - \|x\|e^t)^2}{2x_n\|x\|e^t} \\ &= \frac{2x_n\|x\|e^t + x_1^2 + x_2^2 + \cdots + x_{n-1}^2 + x_n^2 - 2x_n\|x\|e^t + \|x\|^2 e^{2t}}{2x_n\|x\|e^t} \\ &= \frac{\|x\|^2(1 + e^{2t})}{2x_n\|x\|e^t} = \frac{\|x\|}{x_n} \cosh t \end{aligned}$$

has a unique minimum at 0, and $\gamma_{0,\|x\|}(0) = \|x\| e_n$.
(2) Exercise 5.8. □

If L' is a geodesic line in the upper halfspace model, we can map it to L by a composition of the isometries used in Proposition 5.19. These isometries are conformal maps which map the set of spheres and hyperplanes in $\widehat{\mathbb{E}}^n$ to itself. It is easy to see that the neighbourhoods $\mathcal{N}_r(L')$ are infinite cones over Euclidean $(n-1)$-balls or shaped like n-dimensional bananas with opening angles at the endpoints given by Lemma 5.22, see Figure 5.10. As the isometry used to map the ball model to the upper halfspace model is an inversion, the r-neighbourhoods of geodesic lines in the ball model are bananas.

5.5 Möbius transformations and isometries of $\mathbb{H}^2$

The isometries of the upper halfplane model and the ball model of $\mathbb{H}^2$ can be described using 2×2- matrices and Möbius transformations.

[4] $\measuredangle_0(L, x)$ is the angle between the Euclidean ray L and the Euclidean ray from 0 through x.

We can use complex Möbius transformations to describe isometries of the hyperbolic plane and hyperbolic 3-space.

The *special linear group* with real and complex coefficients are

$$\mathrm{SL}_2(\mathbb{R}) = \{A \in \mathrm{GL}_2(\mathbb{R}) : \det A = 1\}$$

and

$$\mathrm{SL}_2(\mathbb{C}) = \{A \in \mathrm{GL}_2(\mathbb{C}) : \det A = 1\}\,,$$

and the *special unitary group* of signature $(1,1)$ is

$$\mathrm{SU}(1,1) = \{A \in \mathrm{SL}_2(\mathbb{C}) : A^* J A = J\} = \left\{ \begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix} : |a|^2 - |b|^2 = 1 \right\},$$

where $J = \mathrm{diag}(-1.1)$ and $A^* = {}^T\bar{A}$.

Recall from complex analysis[5] that any matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with $a, b, c, d \in \mathbb{C}$ and $\det A \neq 0$ determines a *Möbius transformation* $\text{Möb}\, A \colon \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$,

$$\text{Möb}\, A z = \frac{az + b}{cz + d}\,, \tag{5.3}$$

and that the mapping $\text{Möb} \colon \mathrm{SL}_2(\mathbb{C}) \to \mathrm{Homeo}(\widehat{\mathbb{C}})$ is an action by homeomorphisms with $\ker \text{Möb} = \{\pm I_2\}$.

Proposition 5.23. *Let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. The group $\mathrm{SL}_2(\mathbb{K})$ is generated by the elements*

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \textit{and} \quad T_b = \begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix},$$

with $\beta \in \mathbb{K}$.

Proof. Let $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathrm{SL}_2(\mathbb{K})$. Assume $\alpha, \beta, \gamma, \delta \in \mathbb{K}$ with $\gamma \neq 0$. Then, since $\alpha\delta - \beta\gamma = 1$, we have the following equation in $\mathrm{SL}_2(\mathbb{K})$:

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} 1 & \alpha\gamma^{-1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \gamma & 0 \\ 0 & \gamma^{-1} \end{pmatrix} \begin{pmatrix} 1 & \gamma^{-1}\delta \\ 0 & 1 \end{pmatrix}.$$

The claim now follows from the observation that

$$\begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix} = \begin{pmatrix} 1 & -\alpha \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -\alpha^{-1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -\alpha \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

The remaining case $\gamma = 0$ is easier. □

[5] See for example [Ahl, Section 3.3].

Let $f \in \text{Möb}(\widehat{\mathbb{E}}^1)$, $f(x) = \frac{ax+b}{cx+d}$ be a Möbius transformation. The Poincaré extension of f to $\mathbb{H}^2$ is $f\colon \mathbb{H}^2 \to \mathbb{H}^2$, $f(z) = \frac{az+b}{cz+d}$. This can be seen, for example, using the Proposition 5.23 and observing that the reflections in the Euclidean unit circle and the imaginary axis have simple expressions using complex numbers

$$\iota_{0,1}(z) = \frac{z}{|z|^2} = \frac{1}{\overline{z}}$$

and

$$R_{i\mathbb{R}}(z) = z - 2\,\text{Re}\, z = -\overline{z}\,.$$

Similarly, the Poincaré extension of any element of $\text{SU}(1,1)$ to the ball model is given by equation (5.3).

Proposition 5.24. *(1)* $\text{PSL}_2(\mathbb{R}) = \text{SL}_2(\mathbb{R})/\{\pm\,\text{id}\}$ *is the subgroup of index* 2 *in the isometry group of the upper halfplane model of* $\mathbb{H}^2$ *that consists of the orientation-preserving isometries.*

(2) $\text{PU}(1,1) = \text{SU}(1,1)/\{\pm\,\text{id}\}$ *is the subgroup of index* 2 *in the isometry group of the Poincaré disk model of* $\mathbb{H}^2$ *that consists of the orientation-preserving isometries.*

Proof. The Möbius transformations that are defined by elements of $\text{SL}_2(\mathbb{R})$ are analytic functions, thus they are orientation-preserving.

(1) It is easy to check that $\text{Möb}\, J$ and $\text{Möb}\, T_\beta$ for any $\beta \in \mathbb{R}$ are compositions of two reflections of the upper halfplane. Proposition 5.23 implies that $\text{Möb}\,\text{SL}_2(\mathbb{R})$ is a subgroup of the group of isometries.

See for example [And, Theorem 2.26] for the remaining parts of the claim.

(2) The isometry $\iota_{-i,2}$ between the Poincaré disk model and the upper halfplane model has the expression

$$\iota_{-i,2}(z) = \frac{-i\overline{z} + 1}{\overline{z} - i}\,,$$

and the reflection in the geodesic line $]0,\infty[$ is $z \mapsto -\overline{z}$. Their composition is the Möbius transformation $g\colon z \mapsto \frac{-iz-1}{z+i}$, and a computation shows that g gives a conjugacy between $\text{SL}_2(\mathbb{R})$ and $\text{SU}(1,1)$. □

Note that, in fact, we found all orientation-preserving isometries of $\mathbb{H}^2$ and $\mathbb{H}^3$ in our proof of Proposition 5.19.

Remarks 5.25. (1) The trace of a matrix is invariant under conjugation:

$$\text{tr}(BAB^{-1}) = \text{tr}\, A$$

for all $A, B \in \text{SL}_2(\mathbb{C})$. Since the kernel of the map from $\text{SL}_2(\mathbb{C})$ to $\text{Isom}(\mathbb{H}^n)$ is $\pm I_2$, the traces of the two matrices associates with an orientarion-preserving isometry differ by a sign, we can define a map $\text{tr}^2\colon \text{Isom}_+(\mathbb{H}^3) \to \mathbb{R}_+$. This map is invariant under conjugation, and it classifies the elements of $\text{PSL}_2(\mathbb{R})$ and $\text{PSL}_2(\mathbb{C})$ in three types. Items (2) to (4) below elaborate on the classification of $\text{PSL}_2(\mathbb{R})$.

(2) Using the representation of orientation-preserving isometries of $\mathbb{H}^2$ by Möbius transformations, it is straightforward to check that an orientation-preserving isometry A of $\mathbb{H}^2$

which is not the identity has one or two fixed points in $\mathbb{H}^2 \cup \partial_\infty \mathbb{H}^2$. More precisely, the fixed points of the transformation $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ are

$$\frac{a-d}{2c} \pm \frac{\sqrt{\operatorname{tr}^2 A - 4}}{2c} \,.$$

From this formula, we see that an isometry $A \in \mathrm{PSL}_2(\mathbb{R})$ has

- no fixed points in $\mathbb{H}^2$ and two fixed points in $\partial_\infty \mathbb{H}^2$ if $\operatorname{tr}^2 A > 4$,
- one fixed point in $\mathbb{H}^2$ and no fixed points in $\partial_\infty \mathbb{H}^2$ if $\operatorname{tr}^2 A \in [0, 4[$, and
- no fixed points in $\mathbb{H}^2$ and one fixed point in $\partial_\infty \mathbb{H}^2$ if $\operatorname{tr}^2 A = 4$.

(3) Using the above results, one can show that any two Möbius transformations $A, B \in \mathrm{SL}_2(\mathbb{R})$ with $\operatorname{tr}^2 A = \operatorname{tr}^2 B$ are conjugate in $\mathrm{Isom}(\mathbb{H}^2)$. For example, if $A \in \mathrm{SL}_2(\mathbb{R})$ with $\operatorname{tr}^2 A > 4$, then A has two fixed points, which we may assume are 0 and ∞. Now, the equations for fixed points and the determinant imply that $A = \operatorname{diag}(\lambda, \lambda^{-1})$, which implies that $\operatorname{tr}^2 A = (\lambda + 1/\lambda)^2$. Conjugating with the map $z \mapsto -1/z$, we may assume that $\lambda > 1$. Similarly, B is conjugate with $\operatorname{diag}(\lambda, \lambda^{-1})$. The other cases are proved in a similar way.

5.6 Triangles in $\mathbb{H}^n$ (part 2)

The Poincaré model and the upper halfspace model are very useful in many proofs for example because the angle between two tangent vectors is in these models is the same as the Euclidean angle. We use this property to prove the following facts on triangles in hyperbolic space.

Proposition 5.26. *(1) The sum of the angles of a nondegenerate triangle in hyperbolic space is strictly less than π.*

(2) For any $0 < \alpha, \beta, \gamma < \pi$ for which $\alpha + \beta + \gamma < \pi$, there is a triangle with angles α, β and γ. Any two such triangles are isometric.

Proof. By Proposition 4.29, it suffices to consider the hyperbolic plane.

(1) Let T be a triangle with vertices A, B and C. We may assume that one of the vertices A is the origin in the Poincaré disk model. Thus, two sides of the triangle are contained in two radii of the ball and the third one is contained in a circle which is orthogonal to the boundary of $\mathbb{B}^n$. Consider the Euclidean triangle with the same vertices as T. The angles β and γ are strictly smaller than the corresponding angles in the Euclidean triangle. This implies the result as the angles of an Euclidean triangle sum to π.

(2) The second hyperbolic law of cosines[6] implies that the angles of a triangle determine it up to isometry.

Let us consider the upper halfplane model of $\mathbb{H}^2$. Let $0 < r < 1$. At most one of the angles can be greater than or equal to $\frac{\pi}{2}$, and we may assume that $0 < \alpha, \beta < \frac{\pi}{2}$. The geodesic line contained in the Euclidean circle with center $\cos\alpha > 0$ and radius 1 intersects the geodesic line $]0, \infty[$ at an angle α, and the geodesic line contained in the

[6]Proposition4.26

Euclidean circle with center $-r\cos\beta < 0$ and radius r intersects $]0,\infty[$ at an angle β. When $\frac{1-\cos\alpha}{1+\cos\beta} < r < \frac{\sin\alpha}{\sin\beta}$, there are subsegments of these three geodesic lines that make up a triangle where the third angle grows from 0 to $\pi-\alpha-\beta$. $\square$

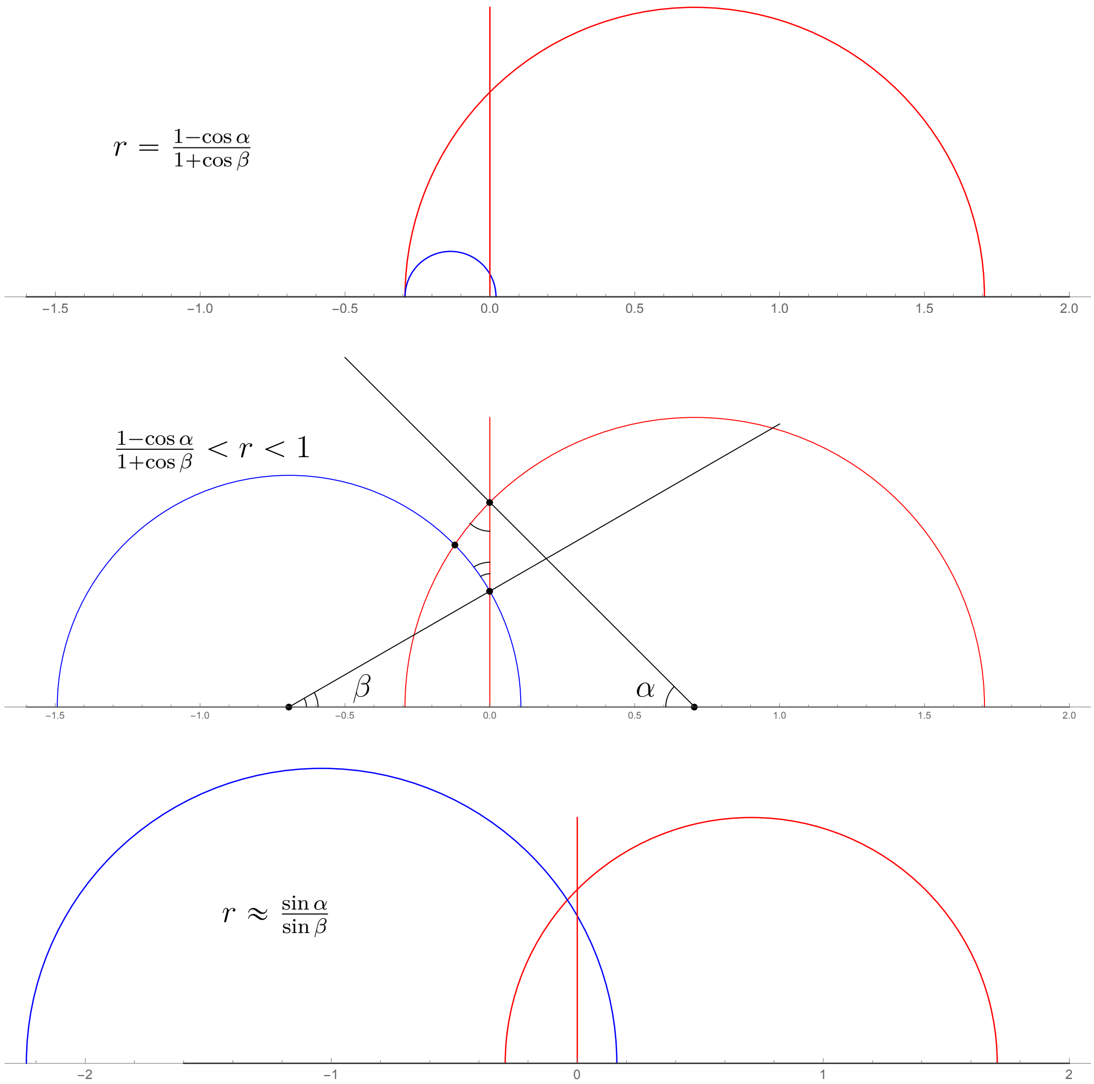


Figure 5.11 — The idea of the proof of Proposition 5.26(2). Here $\alpha = \frac{\pi}{4}$ and $\beta = \frac{\pi}{6}$.

5.7 Generalized triangles in $\mathbb{H}^n$.

We now extend the definition of triangles and allow some of the vertices to be points at infinity of $\mathbb{H}^n$:

A *(generalized) triangle* consists of three distinct points $A, B, C \in \mathbb{H}^n \cup \partial_\infty \mathbb{H}^n$, called the *vertices*, and of the geodesic arcs, rays or lines, called the *sides*, connecting the vertices. If all vertices of a triangle Δ are in $\partial_\infty \mathbb{H}^n$, then Δ is an *ideal triangle*.

Proposition 5.27. *(1) Any generalized triangle in $\mathbb{H}^n$ is contained in an isometrically embedded copy of $\mathbb{H}^2$ in $\mathbb{H}^n$.*

(2) If Δ and Δ' are ideal triangles in $\mathbb{H}^n$, there is an isometry $\gamma \in \operatorname{Isom} \mathbb{H}^n$ such that $\gamma(\Delta) = \Delta'$.

Proof. Exercise 5.9. □

Next, we prove an analog of the second law of cosines for a special kind of generalized triangles. Note that the first law of cosines cannot be generalized to this setting as the triangle in question has two infinitely long sides.

Proposition 5.28. *Let $A, B \in \mathbb{H}^n$ and let $C \in \partial_\infty \mathbb{H}^n$. Then*

$$\cosh c = \frac{1 + \cos\alpha \cos\beta}{\sin\alpha \sin\beta}\,. \tag{5.4}$$

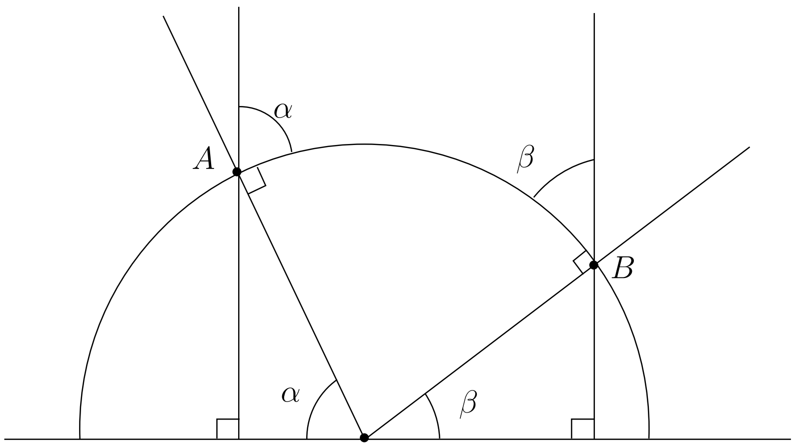


Figure 5.12 —

Proof. By proposition 5.27, it is enough to consider the hyperbolic plane. We use the upper halfplane model and normalize, using Proposition 5.19 with $x_1 = C$, x_2 and x_3 the endpoints of the geodesic line through A and B, and $y_1 = \infty$, $y_2 = -1$ and $y_3 = 1$, so that A and B are on the Euclidean unit circle and $C = \infty$.

Now, $A = (-\cos\alpha, \sin\alpha)$ and $B = (\cos\beta, \sin\beta)$. The result follows from equation (5.2), as

$$1 + \frac{\|A - B\|^2}{2A_2B_2} = 1 + \frac{(\cos\alpha + \cos\beta)^2 + (\sin\alpha - \sin\beta)^2}{2\sin\alpha\sin\beta} = \frac{1 + \cos\alpha\cos\beta}{\sin\alpha\sin\beta}\,. \qquad \square$$

The special case of equation (5.4) with $\beta = \frac{\pi}{2}$:

$$\cosh c = \frac{1}{\sin\alpha} \tag{5.5}$$

is known as the *angle of parallelism.* Another useful form of equation (5.5) is

$$c = \log\cot\frac{\alpha}{2}\,. \tag{5.6}$$

Note that equation (5.4) agrees with the second law of cosines if we define that

the angle at a vertex at infinity is 0.

From now on, we will use this convention.

5.8 Halfspaces and polytopes

Proposition 4.20 implies that hyperbolic hyperplanes are bisectors of two distinct points in $\mathbb{H}^n$. Using this, we can prove

Proposition 5.29. *Hyperplanes in the upper halfspace model are Euclidean hyperplanes orthogonal to the boundary at infinity or intersections with the upper halfspace of Euclidean spheres whose center is in the boundary at infinity.*

Proof. Let x, y be points in the upper halfplane model. Using equation 5.2, we see that the bisector of x and y consists of the solutions z in the upper halfspace of the equation

$$\frac{\|x-z\|}{x_n} = \frac{\|y-z\|}{y_n}. \tag{5.7}$$

If $x_n = y_n$, then equation 5.7 defines an affine plane in $\mathbb{E}^n$ that is orthogonal to the boundary at infinity because it is a translate of the orthogonal complement of the $x - y$ whose nth coordinate is 0.

If $x_n \neq y_n$, then equation 5.7 defines a sphere centered at $\frac{y_n}{x_n-y_n}x + \frac{x_n}{y_n-x_n}y$, which is in the boundary at infinity. ☐

The two connected components of the complement of a hyperplane P $\mathbb{H}^n$ are *open hyperbolic halfspaces*. Their closures in $\mathbb{H}^n$ are *closed hyperbolic halfspaces*.

Lemma 5.30. *Closed and open halfspaces are convex in* $\mathbb{H}^n$.

Proof. Exercise 5.10. ☐

If I is a finite or countable index set and $(H_i)_{i\in I}$ is a collection of closed halfplanes in $\mathbb{H}^n$ with nonempty intersection $P = \bigcap_{i\in I} H_i$ such that $(\partial H_i)_{i\in I}$ is a locally finite collection of hyperplanes,[a] then P is a *locally finite polytope* in $\mathbb{H}^n$.

In dimension $n = 2$, polytopes are *polygons* and in dimension $n = 3$, *polyhedra*.

[a] This means that for any compact $K \subset \mathbb{H}^n$, the set $\{i \in I : K \cap \partial H_i \neq \varnothing\}$ is finite.

Lemma 5.31. *Let X be a uniquely geodesic metric space. Let $K_\alpha \subset X$ be convex sets for all $\alpha \in A$. Then $\bigcap_{\alpha\in A} K_\alpha$ is convex or empty.*

Proof. Exercise 5.11. ☐

Proposition 5.32. *Polytopes in* $\mathbb{H}^n$ *are convex.* ☐

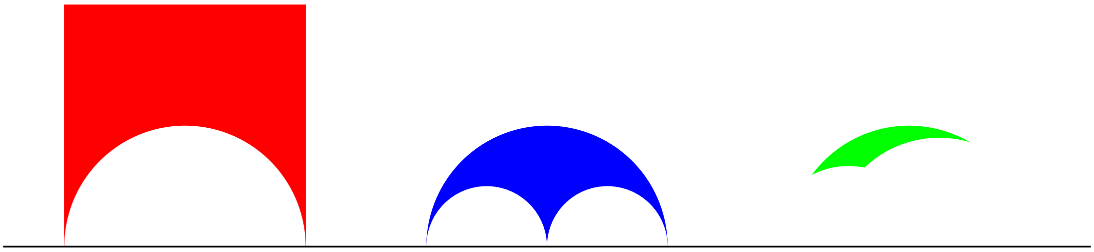


Figure 5.13 — Three polygons in the upper halfplane model of the hyperbolic plane.

5.9 Riemannian metrics, area and volume

The Riemannian metrics of the ball and upper halfspace models are conformal metrics: their expressions are a positive function times the Euclidean Riemannian metric of the underlying subset of $\mathbb{E}^n$.

The Riemannian structure defines a natural volume form and a volume measure on hyperbolic space: If V is for example an open subset of n-dimensional hyperbolic space, and λ_n is the n-dimensional Lebesgue measure, the volume of V is

$$\mathrm{Vol}(V) = \int_V \frac{2^n \, d\lambda_n(x)}{(1 - \|x\|^2)^n}$$

in the Poincaré ball model and

$$\mathrm{Vol}(V) = \int_V \frac{d\lambda_n(x)}{x_n^n}$$

in the upper halfspace model.

Proposition 5.33. *The volume of a ball in hyperbolic space is*

$$\mathrm{Vol}(B(x,r)) = \mathrm{Vol}(\mathbb{S}^{n-1}) \int_0^r \sinh^{n-1} t \, dt.$$

In the hyperbolic plane, we have

$$\mathrm{Vol}(B^2(x,r)) = 4\pi \sinh^2 \frac{r}{2}$$

for all $x \in \mathbb{H}^2$.

The length of a circle of radius r in $\mathbb{H}^2$ is $2\pi \sinh r$.

Proof. As the isometry group acts transitively, the volume of each ball of a fixed radius is the same. Thus, it suffices to consider balls centered at the origoin in the Poincaré ball model. Recall that the Euclidean radius of a ball of hyperbolic radius r centered at 0 in the Poincaré model is $\tanh \frac{r}{2}$. In order to compute the volume of the ball of radius r, recall that the Lebesgue measure is given in the spherical coordinates ($x \leftrightarrow (r, u)$) by

$d\lambda_n(x) = r^{n-1} d\,\mathrm{Vol}_{\mathbb{S}^{n-1}}(u)$, and thus, using a change of variables $s \leftrightarrow \tanh \frac{t}{2}$, we get

$$\begin{aligned}\mathrm{Vol}(\mathbb{B}(x,r)) = \mathrm{Vol}(\mathbb{B}(0,r)) &= \mathrm{Vol}(\mathbb{S}^{n-1}) \int_0^{\tanh\frac{r}{2}} \frac{2^n s^{n-1}}{(1-s^2)^n} ds \\ &= 2^{n-1}\,\mathrm{Vol}(\mathbb{S}^{n-1}) \int_0^r \sinh^{n-1}\frac{t}{2}\cosh^{n-1}\frac{t}{2}\,dt \\ &= \mathrm{Vol}(\mathbb{S}^{n-1}) \int_0^r \sinh^{n-1} t\,dt.\end{aligned}$$

The computation of the length of a circle is left as an execise. □

It is clear from the expression of the volume, that for all $x \in \mathbb{H}^n$, we have

$$\mathrm{Vol}(\mathbb{B}^n(x,r)) \sim \frac{\mathrm{Vol}(\mathbb{S}^n)}{2^{n-1}} e^{(n-1)r},$$

as $r \to \infty$. Thus, the volume of balls in hyperbolic space grows exponentially with the radius, much faster than in Euclidean space.

Proposition 5.34. *The area of the polygon in $\mathbb{H}^2$ bounded by a generalized triangle with angles α, β and γ is $\pi - (\alpha+\beta+\gamma)$.*

Proof. Any triangle T can be described as the difference of two triangles with one vertex at infinity. By the additivity of area and angles in the hyperbolic plane, we may restrict to this special case. Using Proposition 5.19, we can assume that that A and B are on the Euclidean unit circle and that the vertex C has been moved to infinity. Now, the area of T is

$$\int_T \frac{d\lambda_2(x)}{x_2^2} = \int_{-\cos(\alpha)}^{\cos\beta} \int_{\sqrt{1-x_1^2}}^{\infty} \frac{dx_1 dx_2}{x_2^2} = \int_{\cos(\pi-\alpha)}^{\cos\beta} \frac{dx_1}{\sqrt{1-x_1^2}} = \pi - \alpha - \beta\,.$$ □

Exercises

5.1. Fill in the details of the proof of Proposition 5.6.

5.2. Compute the radius of the red ball in Figure 5.5.

5.3. Prove that a ball in hyperbolic space has a unique center.

5.4. Compute the hyperbolic radius and center of the ball $\{z \in \mathbb{H}^2 : |z - ci| \leqslant 1\}$ for all $c > 1$ in the upper halfplane model of $\mathbb{H}^2$.[7]

5.5. Prove Lemma 5.13.

5.6. Complete the proof of Proposition 5.14.[8]

5.7. Prove Proposition 5.20.

5.8. Prove Lemma 5.22(2).

5.9. Prove Proposition 5.27.

5.10. Prove Lemma 5.30.

[7] We identify the upper halfplane model of $\mathbb{H}^2$ with the upper halfplane in $\mathbb{C}$.

[8] Use Lemma 5.13 for inversions.

5.11. Prove Lemma 5.31.

5.12. Prove that the length of a circle of radius r in $\mathbb{H}^2$ is $2\pi \sinh r$.

Appendix A

Inversive geometry

A.1 One-point compactification

Lemma A.1. *Let (X, τ) be a topological space and let ∞ be a point that is not an element of X. Let $\widehat{X} = X \cup \{\infty\}$ and let*

$$\tau_\infty = \{U \subset \widehat{X} : \infty \in U \text{ and } \widehat{X} - U \subset X \text{ is closed and compact}\}.$$

Then $\widehat{\tau} = \tau \cup \tau_\infty$ is a topology in $\widehat{X}$.

Proof. See the basic course in topology. □

Let X be a topological space that is not compact. The topological space $\widehat{X}$ is the *one point compactification* or the *Aleksandroff compactification* of X.

Theorem A.2. *Let (X, τ) be a topological space that is not compact. The one point compactification of X is compact and $(\overline{X})_{\widehat{\tau}} = \widehat{X}$. The topology of $\widehat{X}$ induces the original topology of X on X.*

Proof. Let $(U_\alpha)_{\alpha \in J}$ be an open cover of $\widehat{X}$. There is an index $\alpha_\infty \in J$ such that $\infty \in U_{\alpha_\infty}$. The sets $U_\alpha \cap X$ form an open cover of $X - U_{\alpha_\infty}$ in X. As $X - U_{\alpha_\infty}$ is compact in X, there is some finite $J_0 \subset J$ such that $\widehat{X} - U_{\alpha_\infty} \subset \bigcup_{\alpha \in J_0} U_\alpha$. The finite collection $(U_\alpha)_{\alpha \in J_0 \cup \{\alpha_\infty\}}$ is a cover of $\widehat{X}$.

The subset X is dense in $\widehat{X}$ because, by definition, every open neighbourhood of ∞ intersects X. The topology $\widehat{\tau}$ induces the topology τ in X because τ consists, by definition of elements of τ and of sets formed as the union of an element of τ and $\{\infty\}$. □

Example A.3. *The stereographic projection $\mathscr{S} : \mathbb{S}^n - \{e_3\} \to \mathbb{E}^n = \mathbb{E}^n \times \{0\} \subset \mathbb{E}^{n+1}$*[1]is the mapping

$$\mathscr{S}(x) = \frac{(x_1, x_2, \dots, x_n)}{1 - x_{n+1}}.$$

[1]from the north pole to the level of the equator

It is a homeomorphism that maps each point $x \in \mathbb{S}^n - \{e_{n+1}\}$ to the unique point in $\mathbb{E}^n$ (thought of as the hyperplane $\mathbb{E}^2 \times \{0\}$ in $\mathbb{E}^3$)on the affine line through e_{n+1} and x. Setting $\mathscr{S}(e_{n+1}) = \infty$ we obtain a homeomorphism $\mathscr{S}\colon \mathbb{S}^n \to \widehat{\mathbb{E}^n}$.

The one-point compactification of the Euclidean plane appears in complex analysis as the *Riemann sphere* $\mathbb{C} \cup \{\infty\}$. For example, the mapping $z \mapsto \frac{1}{z}$ becomes a self-homeomorphism of the Riemann sphere if we set $0 \mapsto \infty$ and $\infty \mapsto 0$.

A.2 Inversions

In this short section, we review some basic material on inversions.

Let $c \in \mathbb{E}^n$ and let $\alpha \in \mathbb{R} - \{0\}$. The mapping $\iota_{c,\alpha}\colon \mathbb{E}^n - \{c\} \to \mathbb{E}^n - \{c\}$,

$$\iota_{c,\alpha}(x) = c + \alpha \frac{x - c}{\|x - c\|^2},$$

is an *α-inversion* with a *pole* at c. The number α is called the *power* of the inversion.

Example A.4. In the complex plane,

$$\iota_{0,1}(z) = \frac{z}{|z|^2} = \frac{1}{\bar{z}}.$$

Clearly, for all $x \in \mathbb{E}^n - \{c\}$, we have

$$(x - c \,|\, \iota_{c,\alpha}(x) - c) = \alpha$$

and $\iota_{c,\alpha} \circ \iota_{c,\alpha} = \mathrm{id}|_{\mathbb{E}^n - \{0\}}$. If $\alpha > 0$, then the restriction of $\iota_{c,\alpha}$ to the sphere of center c and radius $\sqrt{\alpha}$ is the identity. The points x and $\iota(x)$ are on the same ray starting at c, and they satisfy

$$\|x - c\| \|\iota_{c,r^2}(x) - c\| = r^2.$$

Let $c \in \mathbb{E}^n$ and $r > 0$. The mapping ι_{c,r^2} is the *inversion in the sphere* of radius r centered at c.

We extend the definition of an inversion $\iota_{c,r}$ to the one-point compactification $\widehat{\mathbb{E}}^n$ of $\mathbb{E}^n$ by setting $\iota_{c,\alpha}(c) = \infty$ and $\iota_{c,\alpha}(\infty) = c$.

Example A.5. $\iota_{e_{n+1},2}|_{\mathbb{S}^n} = \mathscr{S}\colon \mathbb{S}^n \to \widehat{\mathbb{E}}^n$.

Spheres and hyperplanes in $\mathbb{E}^n$ are *generalized hyperplanes*.

Proposition A.6. *Let $c \in \mathbb{E}^n$ and let $\alpha \in \mathbb{R} - \{0\}$. The inversion $\iota_{c,\alpha}$ maps*

(1) the affine subspaces that contain c to themselves,

(2) spheres passing through c to affine hyperplanes that do not contain c,

(3) affine hyperplanes that do not contain c to spheres passing through c, and

(4) spheres that do not pass through c to spheres that do not pass through c.

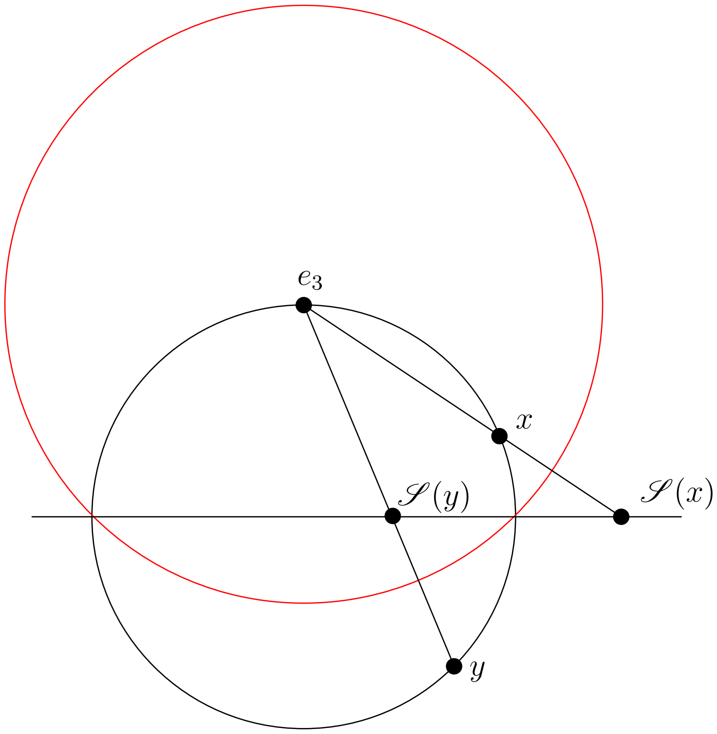


Figure A.1 — Stereographic projection is the restriction to the sphere of an inversion whose center is the north pole.

Proof. (1) is clear from the expression of the inversion.
(2) Clearly, it is enough to consider the case $c = 0$. For any $a \in \mathbb{E}^n - \{0\}$, the sphere $\partial B(a, \|a\|)$ passes through 0 and

$$\partial B(a, \|a\|) = \{x \in \mathbb{E}^n : \|x\|^2 = 2(x \,|\, a)\}.$$

This implies that for any $x \in \partial B(a, \|a\|)$, we have $i_{0,\alpha}(x) = \frac{\alpha x}{2(x \,|\, a)}$, and this gives $(i(x) \,|\, a) = \frac{\alpha}{2}$. Thus,

$$i_{0,\alpha}(\partial B(a, \|a\|)) = \left\{y \in \mathbb{E}^n : (y \,|\, a) = \frac{\alpha}{2}\right\},$$

which is a hyperplane.
(3) follows from (2) and the fact that $i_{0,\alpha}^2 = \mathrm{id}\,|_{\mathbb{E}^n - \{0\}}$.
(4) Consider the sphere $\partial B(a, \rho)$ with $\rho \neq \|a\|$. If $x_1, x_2 \in \partial B(a, \rho)$ are on a line L (through 0), then $\frac{x_1 + x_2}{2}$ is the orthogonal projection of a on L, and we have

$$\|x_1 + x_2\|^2 + \|x_1 + x_2 - 2a\|^2 = 4\|a\|^2$$

and

$$\|x_1 - x_2\|^2 + \|x_1 + x_2 - 2a\|^2 = 4\|\rho^2\|^2.$$

Thus,

$$(x_1 \,|\, x_2) = \|a\|^2 - \rho^2,$$

and therefore $x_2 = \iota_{0,\|a\|^2 - \rho^2}(x_1)$, and we have $= \iota_{0,\|a\|^2 - \rho^2}(\partial B(a, \rho)) = \partial B(a, \rho)$. A simple computation shows that for any $\alpha, \beta \in \mathbb{R} - \{0\}$, we have $\iota_\alpha \circ \iota_\beta(x) = \frac{\alpha}{\beta} x$ for all $x \neq 0$, so we get

$$\iota_{0,\alpha} = \frac{\alpha}{\|a\|^2 - \rho^2} \iota_{0,\|a\|^2-\rho^2} ,$$

which implies $\iota_{0,\alpha}(\partial B(a,\rho)) = (\partial B(a,\rho))$. □

Let D be an open subset of $\mathbb{E}^n$. A mapping $F\colon D \to \mathbb{E}^n$ is *locally conformal*, if it preserves the angles between tangent vectors. Clearly, any mapping whose differential at any point is the composition of an orthogonal transformation and a dilation is locally conformal. A homeomorphism which is a locally conformal map is called a *conformal mapping*. Sometimes one wants to be more precise and say that mappings which preserve angles and orientation are (directly) conformal and those that preserve angles but reverse the orientation are indirectly conformal.

Proposition A.7. *Let $c \in \mathbb{E}^n$ and let $\alpha \in \mathbb{R} - \{0\}$. The inversion $\iota_{c,\alpha}$ is conformal.*

Proof. Observe that $\iota_{c,\alpha} = T_c \circ \iota_{0,\alpha} \circ T_{-c}$. Translations and dilation by α are clearly conformal mappings so it suffices to prove the claim for the standard inversion $\iota_{0,1}$. Note that

$$D\iota_{0,1}(x) = \frac{1}{\|x\|^2} I_n - \frac{2}{\|x\|^4} x\,{}^T\!x ,$$

where ${}^T\!x$ is the transpose of x when x is a column vector. Observe that ${}^T\!D\iota_{0,1}(x) = D\iota_{0,1}(x)$ and that

$$D\iota_{0,1}(x)^2 = \frac{1}{\|x\|^2} I_3 - \frac{4}{\|x\|^6} x\,{}^T\!x + \frac{4}{\|x\|^8} x\,{}^T\!x x\,{}^T\!x = \frac{1}{\|x\|^2} I_n .$$

Thus, $D\iota_{0,1}(x)$ is a multiple of an orthogonal matrix. □

Exercises

A.1. Fill in the details Example A.5.

Part II

Negatively curved spaces

Chapter 6

Gromov-hyperbolic spaces

Gromov-hyperbolic spaces form a class of geodesic metric spaces[1]where some geometric features are similar to hyperbolic space. There are several equivalent definitions of Gromov-hyperbolicity in the literature, most of which formalize the idea that triangles are thin or slim in these spaces in a controlled way. In this chapter, we introduce Gromov-hyperbolic spaces in the same way as they are defined in [BH] and the introduction of [GdlH]. We will also discuss the definition used by [BS], and we will show that these definitions give the same class of Gromov hyperbolic spaces.

6.1 The Rips condition and δ-hyperbolic spaces

The first definition captures a feature of triangles in hyperbolic space.

Let X be a geodesic metric space and let $\delta > 0$. A triangle Δ *satisfies the Rips condition*[a] *with constant* δ if any side of Δ is contained in the union of the closed δ-neighbourhoods of the other two.

[a] or is δ-*slim* as in [BH]

Proposition 6.1. *All triangles in* $\mathbb{H}^n$ *satisfy the Rips condition with constant* $\log(1+\sqrt{2})$.

Proof. By Proposition 4.29, it suffices to consider $\mathbb{H}^2$. Let x, y and z be the vertices of a nondegenerate triangle in the upper halfplane model of the hyperbolic plane. Using the transitivity properties of the isometry group,[2] we may assume that the geodesic line containing the edge $[x, y]$ is $]-1, 1[$, which is the intersection of the Euclidean unit circle with the upper halfplane. Furthermore, using reflections in the imaginary axis and the Euclidean unit circle, we may assume that Re $x <$ Re y and that the Euclidean distance of z from 0 is greater than 1. Using an isometry $\iota_{-1,2} \circ L_t \circ \iota_{-1,2}$ with an appropriate $t \in \mathbb{R}$, we may assume that z is in the imaginary axis as in Figure 6.1.

[1] See [Väi] for a treatment of the theory with weaker assumptions.
[2] See Proposition 5.19.

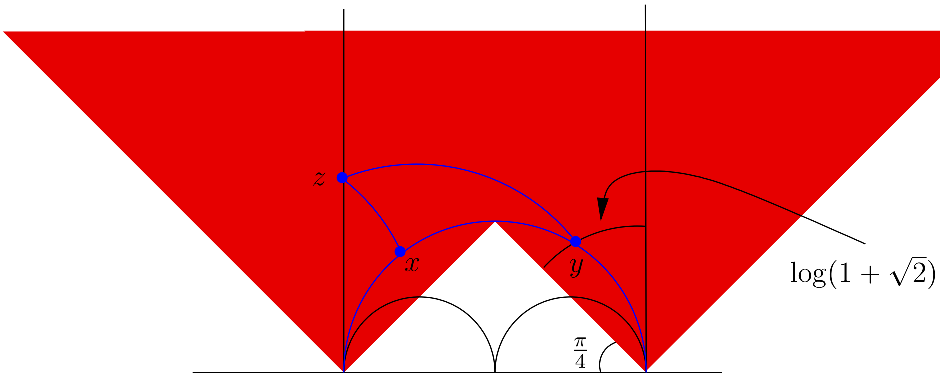


Figure 6.1 — The ideas of Example 6.2.

Let us show that $[x,y] \subset \overline{\mathcal{N}}_{\log(1+\sqrt{2})}([x,z]) \cup \overline{\mathcal{N}}_{\log(1+\sqrt{2})}([y,z])$, using the ideal triangle with vertices at -1, 1 and ∞. If $p \in [x,y] \subset \mathbb{H}^2$, then the shortest geodesic segment from $]-1,\infty[$ to p passes through $[x,z] \cup [z,y]$, and similarly for the shortest geodesic segment from $]-1,\infty[$ to p. It is easy to check with the help of Lemma 5.22 that $]-1,1[$ is contained in the union of the closed $\log(1+\sqrt{2})$-neighbourhoods[3] of the geodesic lines $]-1,\infty[$ and $]-1,\infty[$. Thus, the distance from p to $[x,z] \cup [z,y]$ is at most $\log(1+\sqrt{2})$. □

Let X be a geodesic metric space. If all triangles in X satisfy the Rips condition with constant δ, then X is a *δ-hyperbolic space*.

If X is δ-hyperbolic for some $\delta > 0$, then X is A *Gromov hyperbolic space*.

Example 6.2. (1) We showed in Proposition 6.1 that $\mathbb{H}^n$ is $\log(1+\sqrt{2})$-hyperbolic.

(2) $\mathbb{E}^n$ is not a hyperbolic space if $n \geqslant 2$. If Δ is a non-degenerate triangle in $\mathbb{E}^n$, the midpoint of any one of the sides is at a positive finite distance s from the union of the two others. If $k > 0$, the image of Δ under the homothety (stretch map) $x \mapsto kx$ is a triangle where the corresponding distance is ks. Letting k grow to ∞ proves the claim.

(3) If X is a a geodesic metric space such that the diameter $\operatorname{diam} X$ of X is finite, then X is $\operatorname{diam} X$-hyperbolic. We are not interested in spaces like this.

(4) Any $\mathbb{R}$-tree is 0-*hyperbolic*: Let X be an $\mathbb{R}$-tree and let $x,y,z \in X$. If $[x,y] \cap [x,z] = \{x\}$, then $[x,y] \cup [x,z]$ is an arc with endpoints y and z. Thus, it is the unique arc that joins y to z, in particular, $[x,y] \cup [x,z] = [y,z]$. If $[x,y] \cap [x,z] = [x,w]$ for some $w \neq x$, then $[w,y] \cap [w,z] = \{w\}$ and $[y,z] = [y,w] \cap [w,z] \subset [x,y] \cup [x,z]$.

In particular, $\mathbb{E}^1$ is Gromov-hyperbolic.

(5) The *bi-infinite simplicial*[4] *ladder* is Gromov-hyperbolic. See Figure 6.3.

[3] $\operatorname{arcosh} \frac{1}{\cos \frac{\pi}{4}} = \log(1+\sqrt{2})$.

[4] Recall from section 1.5 that this means we have a metric graph with constant edge length 1.

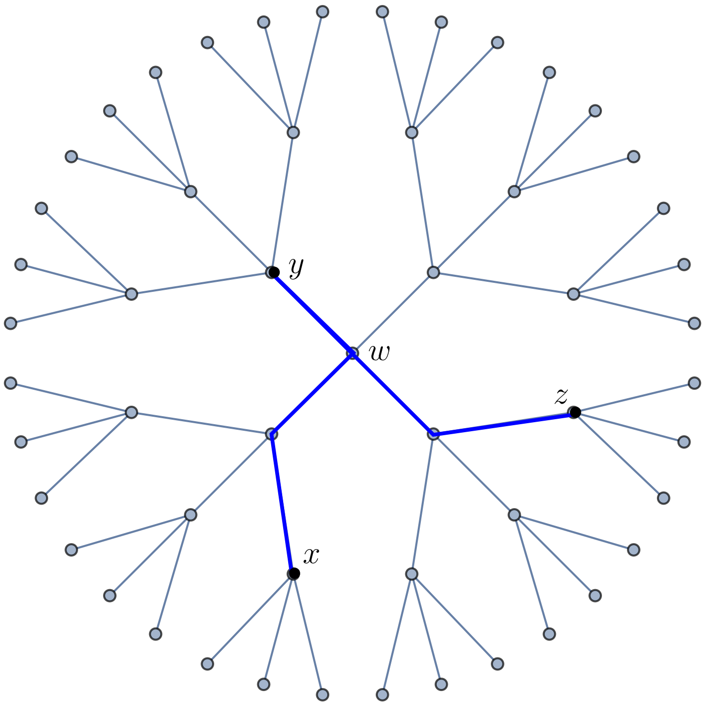


Figure 6.2 — A triangle with vertices x, y and z in a tree.

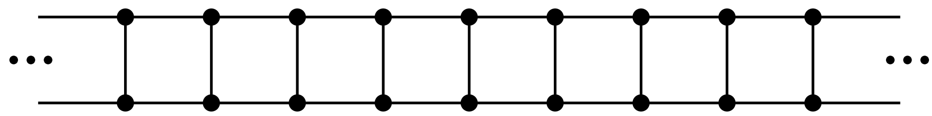


Figure 6.3 — The bi-infinite simplicial ladder.

6.2 The Gromov product and thin triangles

Let X be a metric space and let $x, y, z \in X$. There is a unique triple of positive numbers $r_x, r_y, r_z > 0$ such that

$$\begin{cases} r_x + r_y = d(x,y) \\ r_x + r_z = d(x,z) \\ r_y + r_z = d(y,z)\,. \end{cases} \tag{6.1}$$

Let X be a metric space and let $x, y, z \in X$. The *Gromov product* of y and z *with respect to* x is

$$(y \mid z)_x = \frac{1}{2}\big(d(x,y) + d(x,z) - d(y,z)\big)\,.$$

It is easy to check that the triple $(r_x, r_y, r_z) = ((y \mid z)_x, (x \mid z)_y, (x \mid y)_z)$ is a solution of the linear system (6.1). The triangle inequality implies that the Gromov product is nonnegative.

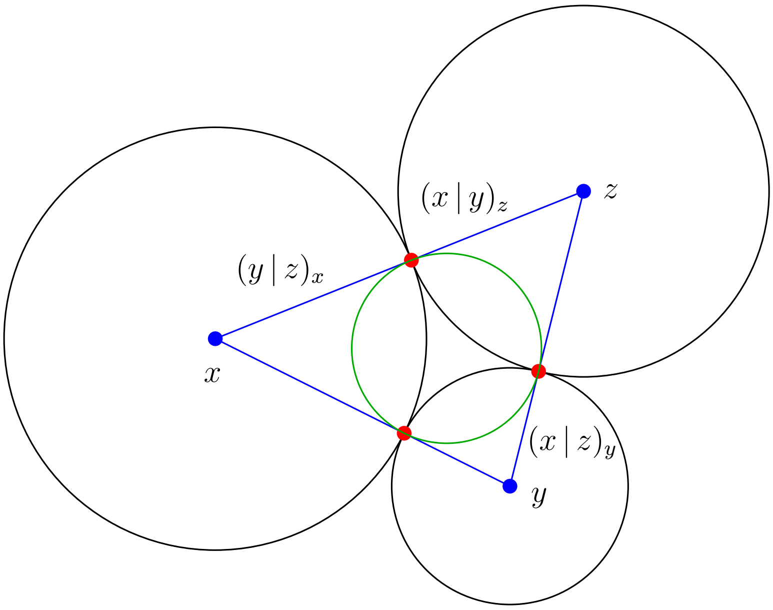


Figure 6.4 — The geometric meaning of the solution of the system (6.1) in the Euclidean plane. The green circle is the incircle of the triangle with endpoints x, y and z. It is the unique circle inside the triangle that is tangent to all the sides. The points of tangency are exactly the internal points of the triangle.

A metric tree with three sides and four vertices such that one vertex has degree 3 and three vertices have degree 1 is a *tripod*.

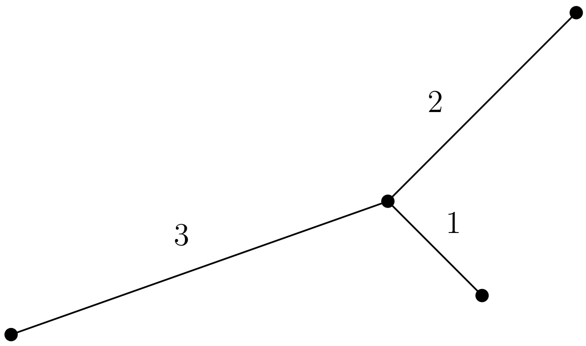


Figure 6.5 — The tripod T_Δ of a triangle Δ with side lengths 3, 4 and 5.

Lemma 6.3. *Let X be a geodesic metric space and let $x, y, y', z, z' \in X$ such that $y' \in [x, y]$ and $z' \in [x, z]$. Then*

$$(y' \mid z')_x \leqslant (y \mid z)_x \,.$$

Proof. Exercise 6.1. □

Lemma 6.4. *Let X be a geodesic metric space and let Δ be a triangle with vertices x, y, z. Let T_Δ be the tripod with side lengths $(y \,|\, z)_x$, $(x \,|\, z)_y$ and $(x \,|\, y)_z$. There is mapping $f_\Delta \colon \Delta \to T_\Delta$ such that the restriction of f_Δ to any side of Δ is an isometry.*

Proof. This is clear because the Gromov products of the vertices give the solution of the system of equations (6.1). □

In many statements and proofs starting from Lemma 6.5, the notation $[a, b]$ means *some or any* geodesic segment with endpoints a and b in places where the actual choice of the possible geodesic segments is not important.

Lemma 6.5. *Let X be a geodesic metric space. Let Δ be a triangle with vertices $x, y, z \in X$. Then*

$$(y \,|\, z)_x \leqslant d(x, [y, z]) .$$

Proof. Let $w \in [y, z]$ be a closest point to x. By Lemma 6.4, there is a point $\widetilde{w} \in [x, y] \cup [x, z]$ such that $f_\Delta(w) = f_\Delta(\widetilde{w})$. We may assume that $\widetilde{w} \in [x, y]$. Note that $d(y, \widetilde{w}) = d(y, w)$ and, as $w \in [y, z]$, $(y \,|\, z)_x \leqslant d(x, \widetilde{w})$. Thus,

$$(y \,|\, z)_x \leqslant d(x, \widetilde{w}) = d(x, y) - d(y, \widetilde{w}) = d(x, y) - d(y, w) \leqslant d(x, w) = d(x, [y, z]) . \quad \square$$

Let X be a geodesic metric space and let $\delta > 0$. A triangle Δ in X is δ-*thin* if $d(a, b) \leqslant \delta$ for all $b \in f_\Delta^{-1}(f_\Delta(a))$ and all $a \in \Delta$.

Lemma 6.6. *Let X be a geodesic metric space. If Δ is a δ-thin triangle with vertices $x, y, z \in X$. Then*

$$(y \,|\, z)_x \leqslant d(x, [y, z]) \leqslant (y \,|\, z)_x + \delta .$$

Proof. The first inequality holds by Lemma 6.5. To prove the second, let v_0 be the central vertex of T_Δ, and let $a \in f_\Delta^{-1}(v_0) \cap [x, y]$ and $b \in f_\Delta^{-1}(v_0) \cap [y, z]$. By assumption, we get

$$d(x, [y, z]) \leqslant d(x, a) + d(a, b) \leqslant (y \,|\, z)_x + \delta . \quad \square$$

Proposition 6.7. *A δ-thin triangle satisfies the Rips condition with constant δ.*

Proof. Exercise 6.2. □

Proposition 6.8. *Let X be a δ-hyperbolic space. Then all triangles in X are 4δ-thin.*

Proof. Assume that there is a triangle Δ with vertices $x, y, z \in X$ that is not 4δ-thin. Then (changing the names of the vertices if necessary) there are points $u \in [x, y]$ and $v \in [x, z]$ such that $f_\Delta(u) = f_\Delta(v)$ and $d(u, v) > 4\delta$. By continuity and as we are assuming a strict inequality $d(u, v) > 4\delta$, we may choose the points u and v such that

$$d(x, u) = d(x, v) < (y \,|\, z)_x . \tag{6.2}$$

Lemma 6.5 applied to triangles with vertices x, u and v, and with vertices y, u and v implies that

$$d(v, [x, y]) = \min\big(d\big(v, [x, u]\big), d\big(v, [u, y]\big)\big) \geqslant \min\big((x \,|\, u)_v, (y \,|\, u)_v\big) . \tag{6.3}$$

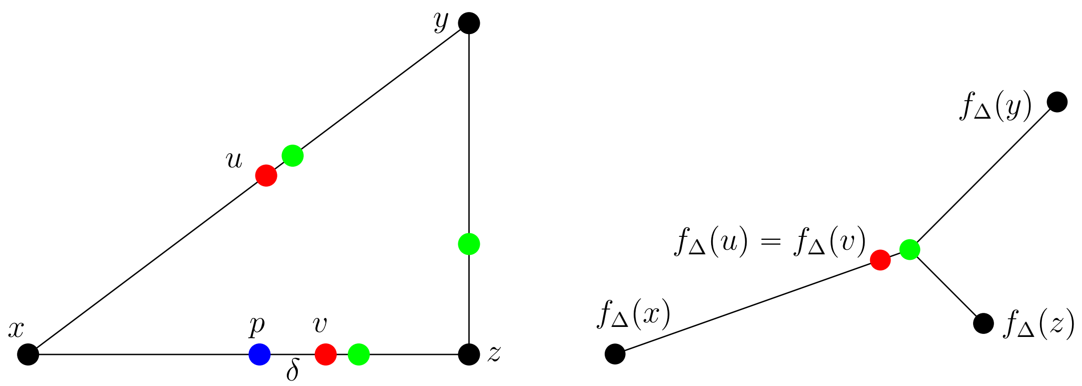


Figure 6.6 — The choice of u and v.

Furthermore, using the assumption that $d(x,u) = d(x,v)$,

$$2(x\,|\,u)_v = d(x,v) + d(u,v) - d(x,u) = d(u,v)$$

and

$$\begin{aligned}2(y\,|\,u)_v &= d(y,v) + d(u,v) - d(y,u)\\ &= d(y,v) + d(u,v) - (d(y,x) - d(x,u))\\ &= d(u,v) + (d(y,v) + d(x,v) - d(y,x))\\ &= d(u,v) + 2(x\,|\,y)_v \geqslant d(u,v)\end{aligned}$$

Combining these observations with the inequality (6.3), we get

$$d(v,[x,y]) \geqslant \frac{1}{2}d(u,v) > 2\delta.$$

In particular, $d(x,v) > 2\delta$ and there is a unique point $p \in [x,v]$ with $d(p,v) = \delta$ and

$$d(p,[x,y]) > \delta\,. \tag{6.4}$$

It remains to estimate the distance from p to $[y,z]$: Lemma 6.5 and the inequality (6.2) imply

$$\begin{aligned}d(p,[y,z]) &\geqslant d(x,[y,z]) - d(p,x) \geqslant (y\,|\,z)_x - d(p,x)\\ &> d(x,v) - d(x,p) = d(p,v) = \delta\,.\end{aligned} \tag{6.5}$$

The inequalities (6.4) and (6.5) show that the triangle Δ does not satisfy the Rips condition with constant δ. □

6.3 The 4-point condition

A metric space X satisfies the 4-*point condition* with parameter δ, if

$$(x\,|\,z)_w \geqslant \min\big((x\,|\,y)_w, (y\,|\,z)_w\big) - \delta$$

for all $x, y, z, w \in X$.

Lemma 6.9. *A metric space X satisfies the 4-point condition with parameter δ if and only if*

$$d(x,z)+d(y,w) \leqslant \max\big(d(x,y)+d(z,w)\,,\, d(x,w)+d(z,y)\big)+2\delta$$

for all $x,y,z,w \in X$.

Proof. Exercise 6.4 □

Example 6.10. If $x=0$, $y=(r,0)$, $z=(r,r)$ and $w=(0,r)$ for $r>0$ in the Euclidean plane, then $d(x,z)+d(y,w)=2\sqrt{2}r$ and $d(x,y)+d(z,w)=d(x,w)+d(z,y)=2r$, and there is no $\delta>0$ such that the 4-point condition would hold for all r.

Proposition 6.11. *If X satisfies the 4-point condition with constant δ, then all triangles in X are 4δ-thin.*

Proof. Let Δ be a triangle with vertices $x,y,z \in X$. Let $u \in [x,y]$ and $v \in [x,z]$ such that $f_\Delta(u)=f_\Delta(v)$. By assumption, we have $d(x,y) \leqslant (y\,|\,z)_x$, and $(u\,|\,y)_x=(v\,|\,z)_x=d(x,u)$.

The 4-point condition gives

$$(u\,|\,v)_x \geqslant \min((u\,|\,z)_x,(z\,|\,v)_x)-\delta=\min((u\,|\,z)_x,d(x,u))-\delta$$

and

$$(u\,|\,z)_x \geqslant \min((u\,|\,y)_x,(y\,|\,z)_x)-\delta\,.$$

Combining these two inequalities, we have

$$(u\,|\,v)_x \geqslant d(x,u)-2\delta. \tag{6.6}$$

On the other hand,

$$(u\,|\,v)_x=\frac{d(u,x)+d(v,x)-d(u,v)}{2}=d(u,x)-\frac{d(u,v)}{2}\,. \tag{6.7}$$

Combining inequality (6.6) and equation (6.7), we have the claim

□

Proposition 6.12. *If all triangles in X are δ-thin, then X satisfies the 4-point condition with constant 2δ.*

Proof. Let $x,y,z,w \in X$, and let us prove that the 4-point condition with parameter δ holds for these points. There is nothing to prove unless

$$\min\big((x\,|\,y)_w,(y\,|\,z)_w\big) > (x\,|\,z)_w\,, \tag{6.8}$$

so we will assume that inequality (6.8) holds.

Let $x' \in [w,x]$, $y' \in [w,y]$, $z' \in [w,z]$ such that

$$d(w,x')=d(w,y')=d(w,z')=\min\big((x\,|\,y)_w,(y\,|\,z)_w\big)\,.$$

As the triangles Δ_{xy} with vertices w,x,y and Δ_{yz} with vertices w,y,z are δ-thin, we have $d(x',y'),d(y',z') \leqslant \delta$, so that

$$d(x',z') \leqslant 2\delta\,. \tag{6.9}$$

By inequality 6.8, there are points $p_x, p_z \in [x, z]$ such that $f_{\Delta_{xz}}(p_x) = f_{\Delta_{xz}}(x')$ and $f_{\Delta_{xz}}(p_z) = f_{\Delta_{xz}}(z')$. In particular, $d(p_x, x'), d(p_z, z') \leqslant \delta$. Note that using the definitions of the various points, we have

$$d(x, p_x) = d(x, x') = d(x, w) - d(w, x') = d(x, w) - \min\big((x \,|\, y)_w, (y \,|\, z)_w\big)\,,$$

and

$$d(z, p_z) = d(z, z') = d(z, w) - d(w, z') = d(z, w) - \min\big((x \,|\, y)_w, (y \,|\, z)_w\big)\,.$$

Thus,

$$\begin{aligned} d(p_x, p_z) &= d(x, z) - d(x, p_x) - d(p_z, z) \\ &= d(x, z) - d(x, w) - d(z, w) + 2\min\big((x \,|\, y)_w, (y \,|\, z)_w\big) \\ &= 2(\min\big((x \,|\, y)_w, (y \,|\, z)_w\big) - (x \,|\, z)_w) \end{aligned}$$

Therefore, using the triangle inequality in the beginning,

$$\begin{aligned} d(x', z') &\geqslant d(p_x, p_z) - 2\delta \\ &= 2\big(\min\big((x \,|\, y)_w, (y \,|\, z)_w\big) - (x \,|\, z)_w - \delta\big) \end{aligned}$$

Combining this with inequality (6.9) gives the claim. □

6.4 Approximation of paths by geodesics

In this section, we prove a technical result that is useful in section **??**. The proof makes strong use of δ-hyperbolicity.

Proposition 6.13. *Let X be a δ-hyperbolic space. Let $\gamma\colon [0,1] \to X$ be a rectifiable path*[5] *and let $j\colon [0, d(\gamma(0), \gamma(1)] \to X$ be a geodesic segment such that $j(0) = \gamma(0)$ and $j(1) = \gamma(1)$. For any $t \in [0, d(\gamma(0), \gamma(1))]$,*

$$d(j(t), \gamma([0,1])) \leqslant \delta \log_2 \ell(\gamma) + 1\,. \tag{6.10}$$

Proof. The inequality (6.10) is satisfied trivially if $\ell(\gamma) \leqslant 1$. We assume that $\ell(\gamma) \geqslant 1$ and that γ is parametrized proportional to arclength.[6]

Let $N \in \mathbb{N}$ such that $\frac{\ell(\gamma)}{2} \leqslant 2^N \leqslant \ell(\gamma)$. Let $t \in [0, d(\gamma(0), \gamma(1))]$. Let Δ_1 be a triangle with vertices $\gamma(0)$, $\gamma(1)$ and $\gamma(\frac{1}{2})$ such that one of the sides is the image of the geodesic segment j. As X is δ-hyperbolic,

$$\gamma(t) \in \overline{\mathscr{N}}_\delta\big([\gamma(0), \gamma(\frac{1}{2})]\big) \cup \overline{\mathscr{N}}_\delta\big([\gamma(\frac{1}{2}), \gamma(1)]\big)\,.$$

Thus, there is a point $y_1 \in \big[\gamma(0), \gamma(\frac{1}{2})\big] \cup \big[\gamma(\frac{1}{2}), \gamma(1)\big]$ such that $d(j(t), y_1) \leqslant \delta$. If $y_1 \in \big[\gamma(0), \gamma(\frac{1}{2})\big]$, let Δ_2 be a triangle with vertices $\gamma(0)$, $\gamma(\frac{1}{4})$ and $\gamma(\frac{1}{2})$. Otherwise, let Δ_2 be the triangle with vertices $\gamma(\frac{1}{2})$, $\gamma(\frac{3}{4})$ and $\gamma(1)$.

[5] A path γ is *rectifiable* if $\ell(\gamma) < \infty$.

[6] See [BH, Proposition I.1.20] and the remarks after it for a proof that we can make the second assumption without loss of generality.

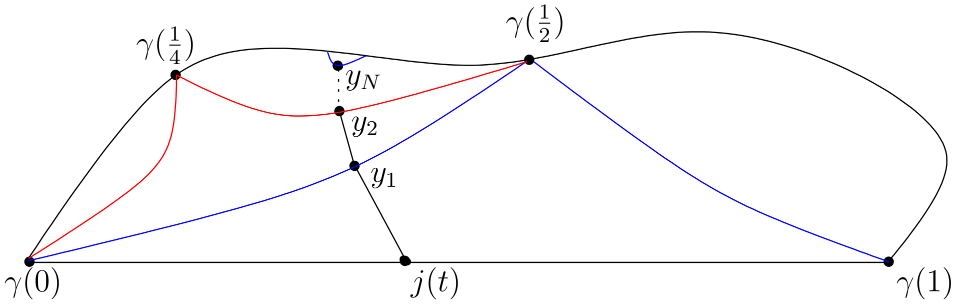


Figure 6.7 — The idea of the proof of Proposition 6.13.

Assume that we are in the first case. Then, using δ-hyperbolicity as above, there is a point $y_2 \in \big[\gamma(0), \gamma(\frac{1}{4})\big] \cup \big[\gamma(\frac{1}{4}), \gamma(\frac{1}{2})\big]$ such that $d(y_1, y_2) \leqslant \delta$. We continue inductively, and construct a finite sequence of points $y_1, y_2, \ldots, y_N$ such that $d(y_k, y_{k+1}) \leqslant \delta$ for all $1 \leqslant k \leqslant N-1$. Note that, by construction, $y_N \in \big[\gamma(\frac{k}{2^N}), \gamma(\frac{k+1}{2^N})\big]$ for some $0 \leqslant k \leqslant 2^N - 1$, and therefore, $d(y_N, \gamma([0,1])) \leqslant \frac{\ell(\gamma)}{2^{N+1}} \leqslant 1$. The triangle inequality gives the estimate

$$d\big(j(t), \gamma([0,1])\big) \leqslant N\delta + 1 \leqslant \log_2 \ell(\gamma) + 1\,. \qquad \square$$

Example 6.14. In the Euclidean plane, the distance from the center of a half-circle to the half-circle grows linearly with the radius. Therefore, the inequality (6.10) cannot be satisfied for the geodesic segment $[-r, r]$ and a parametrization $\gamma(t) = re^{i\pi t}$ of the half-circle.

Exercises

6.1. Prove Lemma 6.3.

6.2. Prove Proposition 6.7.

6.3. Let Δ be a triangle in a geodesic space X and let $\delta \geqslant 0$. Prove that Δ is δ-thin if and only if

$$d(u, v) \leqslant d\big(f_\Delta(u), f_\Delta(v)\big) + \delta$$

for all $u, v \in \Delta$.

6.4. Prove Lemma 6.9.

6.5. Let T be a simplicial tree. Let $x_0 \in T$ and let $\rho_1, \rho_2 \colon [0, \infty[\; \to T$ be geodesic rays such that $\rho_1(0) = \rho_2(0) = x_0$ and $\rho_1 \neq \rho_2$. Prove that the limit $\lim_{t\to\infty} (\rho_1(t) \,|\, \rho_2(t))_{x_0}$ exists.[7]

6.6. Let $\rho_1, \rho_2 \colon [0, \infty[\; \to \mathbb{H}^2$ be geodesic rays such that $\rho_1(0) = \rho_1(0) = 0$ in the Poincaré disk model and $\rho_1 \neq \rho_2$. Prove that $(\rho_1(t) \,|\, \rho_2(t))_0$ is bounded.[8]

6.7. Let $\rho_1, \rho_2 \colon [0, \infty[\; \to \mathbb{E}^2$ be geodesic rays such that $\rho_1(0) = \rho_1(0) = 0$ and $\rho_1 \neq -\rho_2$. Prove that $(\rho_1(t) \,|\, \rho_2(t))_0$ is not bounded.

[7] Prove that the function $t \mapsto (\rho_1(t) \,|\, \rho_2(t))_{x_0}$ is constant for large t.

[8] Lemma 6.6 and Proposition 4.26 can be useful.

Bibliography

[Ahl] L. V. Ahlfors. *Complex analysis*. McGraw-Hill Book Co., New York, third edition, 1978.

[And] J. W. Anderson. *Hyperbolic geometry*. Springer Undergraduate Mathematics Series. Springer-Verlag London Ltd., London, second edition, 2005.

[Bea] A. F. Beardon. *The geometry of discrete groups*, volume 91 of *Graduate Texts in Mathematics*. Springer-Verlag, 1983.

[Bou1] N. Bourbaki. *General topology. Chapters 1–4*. Elements of Mathematics. Springer-Verlag, Berlin, 1998.

[Bou2] N. Bourbaki. *General topology. Chapters 5–10*. Elements of Mathematics (Berlin). Springer-Verlag, Berlin, 1998.

[BH] M. R. Bridson and A. Haefliger. *Metric spaces of non-positive curvature*, volume 319 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, 1999.

[BS] S. Buyalo and V. Schroeder. *Elements of asymptotic geometry*. EMS Monographs in Mathematics. European Mathematical Society (EMS), Zürich, 2007.

[GdlH] É. Ghys and P. de la Harpe, editors. *Sur les groupes hyperboliques d'après Mikhael Gromov*, volume 83 of *Progress in Mathematics*. Birkhäuser Boston, Inc., Boston, MA, 1990.

[Lee] J. M. Lee. *Riemannian manifolds*, volume 176 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1997. An introduction to curvature.

[Mun] J. R. Munkres. *Topology*. Prentice Hall, Inc., second edition, 2000.

[Pet] P. Petersen. *Riemannian geometry*, volume 171 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2006.

[Ser] J.-P. Serre. *Trees*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2003.

[Str] K. R. Stromberg. *Introduction to classical real analysis*. Wadsworth International Mathematics Series. Wadsworth International, Belmont, CA, 1981.

[Väi] J. Väisälä. Gromov hyperbolic spaces. *Expo. Math.*, 23(3):187–231, 2005.