# Hyperbolic geometry, surfaces, and 3-manifolds

Bruno Martelli

Dipartimento di Matematica "Tonelli", Largo Pontecorvo 5, 56127 Pisa, Italy

*E-mail address*: `martelli at dm dot unipi dot it`

*version: september 30, 2013*

# Contents

# Introduction

The aim of this book is to introduce hyperbolic geometry and its applications to two- and three-manifolds topology. Chapter 1 includes all the preliminaries we will need, all the material in the subsequent chapters is self-contained.

The book is still incomplete and all references are missing. Most of the topics presented here have their origin in Thurston's notes and papers and are of course already covered by other books, which we have shamelessly and widely used. These include *Lectures on hyperbolic geometry* by Benedetti and Petronio, *Foundations of hyperbolic manifolds* by Ratcliffe, *Travaux de Thurston sur les surfaces* by Fathi, Laudenbach and Poenaru, and *A primer on the mapping class group* by Farb and Margalit. For the theory of currents we have consulted Bonahon's original papers and McMullen's *Teichmüller theory notes.*

CHAPTER 1

# Preliminaries

We expose in this chapter a quick overview of the basic differential topology and geometry that we will use in this book.

## 1. Differential topology

**1.1. Differentiable manifolds.** A *topological manifold* of dimension $n$ is a paracompact Hausdorff topological space $M$ locally homeomorphic to $\mathbb{R}^n$. In other words, there is a covering $\{U_i\}$ of $M$ consisting of open sets $U_i$ homeomorphic to open sets $V_i$ in $\mathbb{R}^n$.

Topological manifolds are difficult to investigate, their definition is too general and allows to directly define and prove only few things. Even the notion of dimension is non-trivial: to prove that an open set of $\mathbb{R}^k$ is not homeomorphic to an open set of $\mathbb{R}^h$ for different $k$ and $h$ we need to use non-trivial constructions like homology. It is also difficult to treat topological subspaces: for instance, the *Alexander horned sphere* shown in Fig. 1 is a

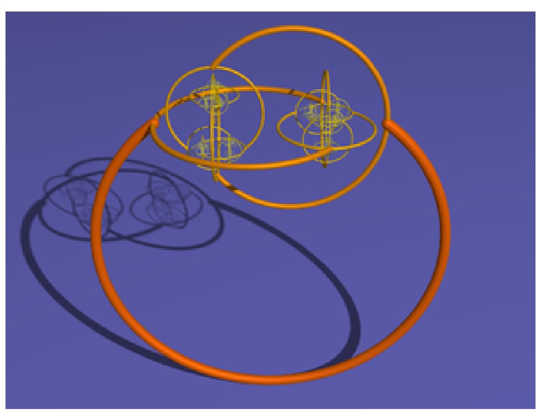


FIGURE 1. The *Alexander horned sphere* is a subset of $\mathbb{R}^3$ homeomorphic to the 2-sphere $S^2$. It divides $\mathbb{R}^3$ into two connected components, none of which is homeomorphic to an open ball. It was constructed by Alexander as a counterexample to a natural three-dimensional generalization of Jordan's curve theorem. The natural generalization would be the following: does every 2-sphere in $\mathbb{R}^3$ bound a ball? If the 2-sphere is only topological, the answer is negative as this counterexample shows. If the sphere is a differentiable submanifold, the answer is however positive as proved by Alexander himself.

subspace of $\mathbb{R}^3$ topologically homeomorphic to a 2-sphere. It is a complicate object that has many points that are not "smooth" and that cannot be "smoothened" in any reasonable way.

We need to define some "smoother" objects, and for that purpose we can luckily invoke the powerful multivariable infinitesimal calculus. For this purpose we introduce the notions of *chart* and *atlas*. Let $U \subset \mathbb{R}^n$ be an open set: a map $f\colon U \to \mathbb{R}^k$ is *smooth* if it is $C^\infty$, *i.e.* it has partial derivatives of any order.

Definition 1.1. Let $M$ be a topological manifold. A *chart* is a fixed homeomorphism $\varphi_i\colon U_i \to V_i$ between an open set $U_i$ of $M$ and an open set $V_i$ of $\mathbb{R}^n$. An *atlas* is a set of charts $\big\{(U_i, \varphi_i)\big\}$ such that the open sets $U_i$ cover $M$.

If $U_i \cap U_j \neq \emptyset$ there is a *transition map* $\varphi_{ji} = \varphi_j \circ \varphi_i^{-1}$ that sends homeomorphically the open set $\varphi_i(U_i \cap U_j)$ onto the open set $\varphi_j(U_i \cap U_j)$. Since these two open sets are in $\mathbb{R}^n$, it makes sense to require $\varphi_{ij}$ to be smooth. A *differentiable atlas* is an atlas where the transition maps are all smooth.

Definition 1.2. A *differentiable manifold* is a topological manifold equipped with a differentiable atlas.

We will often use the word *manifold* to indicate a differentiable manifold. The integer $n$ is the *dimension* of the manifold. We have defined the objects, so we now turn to their morphisms.

Definition 1.3. A map $f\colon M \to M'$ between differentiable manifolds is *smooth* if it is smooth when read locally through charts. This means that for every $p \in M$ and any two charts $(U_i, \varphi_i)$ of $M$ and $(U'_j, \varphi'_j)$ of $N$ with $p \in U_i$ and $f(p) \in U'_j$, the composition $\varphi'_j \circ f \circ \varphi_i^{-1}$ is a smooth map from $V_i$ to $V'_j$.

A *diffeomorphism* is a smooth map $f\colon M \to M'$ that admits a smooth inverse $g\colon M' \to M$.

A *curve* in $M$ is a smooth map $\gamma\colon I \to M$ defined on some interval $I$ of the real line, which may be bounded or unbounded.

Definition 1.4. A differentiable manifold is *oriented* if it is equipped with an orientable atlas, *i.e.* an atlas where all transition functions are orientation-preserving (that is, the determinant of their differential at any point is positive).

A manifold which can be oriented is called *orientable*.

**1.2. Tangent space.** Let $M$ be a differentiable manifold of dimension $n$. We may define for every point $p \in M$ a $n$-dimensional vector space $T_pM$ called the *tangent space*.

The space $T_p$ may be defined briefly as the set of all curves $\gamma\colon ]-a, a[\to M$ such that $f(0) = p$ and $a > 0$ is arbitrary, considered up to some equivalence relation. The relation is the following: we identify two curves that,

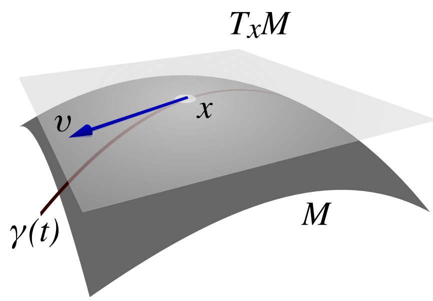



FIGURE 2. The tangent space in $x$ may be defined as the set of all curves $\gamma$ with $\gamma(0) = x$ seen up to an equivalent relation that identifies two curves having (in some chart) the same tangent vector at $x$. This condition is chart-independent.

read on some chart $(U_i, \varphi_i)$, have the same tangent vector at $\varphi_i(p)$. The definition does not depend on the chart chosen.

A chart identifies $T_pM$ with the usual tangent space at $\varphi_i(p)$ in the open set $V_i = \varphi_i(U_i)$, which is simply $\mathbb{R}^n$. Two distinct charts $\varphi_i$ and $\varphi_j$ provide different identifications with $\mathbb{R}^n$, which differ by a linear isomorphism: the differential $d\varphi_{ji}$ of the transition map $\varphi_{ij}$. The structure of $T_p$ as a vector space is then well-defined, while its identification with $\mathbb{R}^n$ is not.

Every smooth map $f: M \to N$ between differentiable manifolds induce at each point $p \in M$ a linear map $df_p: T_pM \to T_{f(p)}N$ between tangent spaces in the following simple way: the curve $\gamma$ is sent to the curve $f \circ \gamma$.

DEFINITION 1.5. A smooth map $f: M \to N$ is a *local diffeomorphism* at a point $p \in M$ if there are two open sets $U \subset M$ e $V \subset N$ containing respectively $p$ and $f(p)$ such that $f|_U: U \to V$ is a diffeomorphism.

The inverse function theorem in $\mathbb{R}^n$ implies easily the following fact, that shows the importance of the notion of tangent space.

THEOREM 1.6. *Let $f: M \to N$ be a smooth map between manifolds of the same dimension. The map is a local diffeomorphism at $p \in M$ if and only if the differential $df_p: T_pM \to T_{f(p)}N$ is invertible.*

In the theorem a condition satisfied at a single point (differential invertible at $p$) implies a local property (local diffeomorphism). Later, we will see that in riemannian geometry a condition satisfied at a single point may even imply a global property.

If $\gamma: I \to M$ is a curve, its *velocity* $\gamma'(t)$ in $t \in I$ is the tangent vector $\gamma'(t) = d\gamma_t(1)$. Here “1” means the vector 1 in the tangent space $T_tI = \mathbb{R}$. We note that the velocity is a vector and not a number: the modulus of a tangent vector is not defined in a differentiable manifold (because the tangent space is just a real vector space, without a norm).

**1.3. Differentiable submanifolds.** Let $N$ be a differentiable manifold of dimension $n$.

DEFINITION 1.7. A *differentiable submanifold* $M \subset N$ of dimension $m \leqslant n$ is a subset such that for every $p \in M$ there is an open set $U \subset M$ and a diffeomorphism $\varphi\colon U \to V$ onto an open set $V \subset \mathbb{R}^n$ that sends $U \cap M$ onto $V \cap L$ where $L$ is a linear subspace of dimension $m$.

The pairs $\{U \cap M, \varphi|_{U\cap M}\}$ form an atlas for $M$, which then inherits a structure of $m$-dimensional differentiable manifold. At every point $p \in M$ the tangent space $T_pM$ is a linear subspace of $T_pN$.

**1.4. Fiber bundles.** The following notion is fundamental in differential topology.

DEFINITION 1.8. A *smooth fiber bundle* is a smooth map

$$\pi\colon E \longrightarrow M$$

such that every fiber $\pi^{-1}(p)$ is diffeomorphic to a fixed manifold $F$ and $\pi$ looks locally like a projection. This means that $M$ is covered by open sets $U_i$ equipped with diffeomorphisms $\psi_i\colon U_i \times F \to \pi^{-1}(U_i)$ such that $\pi \circ \psi_i$ is the projection on the first factor.

The manifolds $E$ and $B$ are called the *total* and *base manifold*, respectively. The manifold $F$ is the *fiber* of the bundle. A *section* of the bundle is a smooth map $s\colon B \to E$ such that $\pi \circ s = \mathrm{id}_B$.

A *smooth vector bundle* is a smooth fiber bundle where every fiber $\pi^{-1}(p)$ has the structure of a $n$-dimensional vector space which varies smoothly with $p$. This smoothness condition is realized formally by requiring that $F = \mathbb{R}^n$ and $\psi(p, \cdot)\colon F \to \pi^{-1}(p)$ be an isomorphism of vector spaces for every $\psi$ as above.

The *zero-section* of a smooth vector bundle is the section $s\colon B \to E$ that sends $p$ to $s(p) = 0$, the zero in the vector space $\pi^{-1}(p)$. The image $s(B)$ of the zero-section is typically identified with $B$ via $s$.

Two vector bundles $\pi\colon E \to B$ and $\pi'\colon E' \to B$ are *isomorphic* if there is a diffeomorphism $\psi\colon E \to E'$ such that $\pi = \pi' \circ \psi$, which restricts to an isomorphism of vector spaces on each fiber.

As every manifold here is differentiable, likewise every bundle will be smooth and we will hence often drop this word.

**1.5. Tangent and normal bundle.** Let $M$ be a differentiable manifold of dimension $n$. The union of all tangent spaces

$$TM = \bigcup_{p\in M} T_pM$$

is naturally a differentiable manifold of double dimension $2n$, called the *tangent bundle*. The tangent bundle $TM$ is naturally a vector bundle over $M$, the fiber over $p \in M$ being the tangent space $T_pM$.

Let $M \subset N$ be a smooth submanifold of $N$. The *normal space* at a point $p \in M$ is the quotient vector space $\nu_p M = T_pN/_{T_pM}$. The *normal bundle* $\nu M$ is the union

$$\nu M = \bigcup_{p \in M} \nu_p M$$

and is also naturally a smooth fiber bundle over $M$. The normal bundle is not canonically contained in $TN$ like the tangent bundle, but (even more usefully) it may be embedded directly in $N$, as we will soon see.

**1.6. Immersion and embedding.** A smooth map $f\colon M \to N$ between manifolds is an *immersion* if its differential is everywhere injective: note that this does not imply that $f$ is injective. The map is an *embedding* if it is a diffeomorphism onto its image: this means that $f$ is injective, its image is a submanifold, and $f\colon M \to f(M)$ is a diffeomorphism.

PROPOSITION 1.9. *If $M$ is compact, an embedding is an injective immersion.*

**1.7. Homotopy and isotopy.** Let $X$ and $Y$ be topological spaces. We recall that a *homotopy* between two continuous maps $\varphi, \psi\colon X \to Y$ is a map $F\colon X \times [0,1] \to Y$ such that $F_0 = \varphi$ and $F_1 = \psi$, where $F_t = F(\cdot, t)$. A homotopy is an *isotopy* if every map $F_t$ is injective.

An *ambient isotopy* on a topological space $X$ is an isotopy between $\mathrm{id}_X$ and some other homeomorphism $\varphi\colon X \to X$. When $X$ is a smooth manifold we tacitly suppose that $\varphi$ is a diffeomorphism. The following theorem says that isotopy implies ambient isotopy under mild assumptions. The *support* of an ambient isotopy is the closure of the set of points that are effectively moved.

THEOREM 1.10. *Let $f, g\colon M \to N$ be two smooth isotopic embeddings of manifolds. If $M$ is compact there is an ambient isotopy relating them supported on a compact subset of $N$.*

**1.8. Tubolar neighborhood.** Let $M \subset N$ be a differentiable submanifold. A *tubolar neighborhoods* of $M$ is an open subset $U \subset N$ such that there is a diffeomorphism $\nu M \to U$ sending the zero-section onto $M$ via the identity map.

THEOREM 1.11. *Let $M \subset N$ be a closed differentiable submanifold. A tubolar neighborhood for $M$ exists and is unique up to an ambient isotopy fixing $M$ pointwise.*

Vector bundles are hence useful (among other things) to understand neighborhoods of submanifolds. Since we will be interested essentially in manifolds of dimension at most 3, two very simple cases will be enough for us.

PROPOSITION 1.12. *A connected closed manifold $M$ has a unique orientable line bundle $E \to M$ up to isomorphism.*

The orientable line bundle on $M$ is a product $M \times \mathbb{R}$ precisely when $M$ is also orientable. If $M$ is not orientable, the unique orientable line bundle is indicated by $M \tilde{\times} \mathbb{R}$.

Proposition 1.13. *For every $n$, there are exactly two vector bundles of dimension $n$ over $S^1$ up to isomorphism, one of which is orientable.*

Again, the orientable vector bundle is just $S^1 \times \mathbb{R}^n$ and the non-orientable one is denoted by $S^1 \tilde{\times} \mathbb{R}^n$. These simple facts allow to fully understand the neighborhoods of curves in surfaces, and of curves and surfaces inside orientable 3-manifolds.

**1.9. Manifolds with boundary.** A *differentiable manifold $M$ with boundary* is a topological space with charts on a fixed half-space of $\mathbb{R}^n$ instead of $\mathbb{R}^n$, forming a smooth atlas. The points corresponding to the boundary of the half-space form a subset of $M$ denoted by $\partial M$ and called *boundary*. The boundary of a $n$-manifold is naturally a $(n-1)$-dimensional manifold without boundary. The *interior* of $M$ is $M \setminus \partial M$.

We can define the tangent space $T_xM$ of a point $x \in \partial M$ as the set of all curves in $M$ starting from $x$, with the same equivalence relation as above. The space $T_xM$ is naturally a semi-vector space, limited by a hyperplane naturally identified with $T_xM$. Most of the notions introduced for manifolds extend in an appropriate way to manifolds with boundary.

The most important manifold with boundary is certainly the *disc*

$$D^n = \big\{x \bigm| \|x\| \leqslant 1\big\} \subset \mathbb{R}^n.$$

More generally, a *disc* in a $n$-manifold $N$ is a submanifold $D \subset N$ with boundary, diffeomorphic to $D^n$. Since a disc is in fact a (closed) tubolar neighborhood of any point in its interior, the uniqueness of tubular neighborhoods imply the following.

Theorem 1.14. *Let $N$ be a connected manifold. Two discs $D, D' \subset N$ contained in the interior of $N$ are always related by an ambient isotopy.*

A *boundary component* $N$ of $M$ is a connected component of $\partial M$. A *collar* for $N$ is an open neighborhood diffeomorphic to $N \times [0, 1)$. As for tubular neighborhoods, every compact boundary component has a collar, unique up to ambient isotopy.

**1.10. Cut and paste.** If $M \subset N$ is an orientable $(n-1)$-manifold in an orientable $n$-manifold, it has a tubular neighborhood diffeomorphic to $M \times \mathbb{R}$. The operation of *cutting* $N$ along $M$ consists of the removal of the portion $M \times (-1, 1)$. The resulting manifold has two new boundary components $M \times \{-1\}$ and $M \times \{1\}$, both diffeomorphic to $M$. By the uniqueness of the tubular neighborhood, the cut manifold depends (up to diffeomorphisms) only on $M \subset N$.

Conversely, let $M$ and $N$ be two $n$-manifolds with boundary, and let $\varphi\colon \partial M \to \partial N$ be a diffeomorphism. It is possible to *glue* $M$ and $N$ along $\varphi$ and obtain a new $n$-manifold as follows.

A naïf approach would consist in taking the topological space $M \sqcup N$ and identify $x$ with $\varphi(x)$ for all $x \in M$. The resulting quotient space is indeed a topological manifold, but the construction of a smooth atlas is not immediate. A quicker method consists of taking two collars $\partial M \times [0,1)$ and $\partial N \times [0,1)$ of the boundaries and then consider the topological space

$$(M \setminus \partial M) \sqcup (N \setminus \partial N).$$

Now we identify the points $(x,t)$ and $(\varphi(x), 1-t)$ of the open collars, for all $x \in \partial M$ and all $t \in (0,1)$. Having now identified two *open* subsets of $M \setminus \partial M$ and $N \setminus \partial N$, a differentiable atlas for the new object is immediately derived from the atlases of $M$ and $N$.

**1.11. Transversality.** Let $f\colon M \to N$ be a smooth map between manifolds and $X \subset N$ be a submanifold. We say that $f$ is *transverse* to $X$ if for any $p \in f^{-1}X$ the following condition holds:

$$\mathrm{Im}\,(df_p) + T_{f(x)}X = T_{f(x)}N.$$

The maps transverse to a fixed $X$ are generic, that is they form an open dense subset in the space of all continuous maps from $X$ to $Y$, with respect to some topology. In particular the following holds.

THEOREM 1.15. *Let $f\colon M \to N$ be a continuous map and $d$ a distance on $N$ compatible with the topology. For every $\varepsilon > 0$ there is a smooth map $g$ transverse to $X$, homotopic to $f$, with $d(f(p), g(p)) < \varepsilon$ for all $p \in M$.*

## 2. Riemannian geometry

**2.1. Metric tensor.** In a differentiable manifold, a tangent space at every point is defined. However, many natural geometric notions are not defined, such as distance between points, angle between tangent vectors, length of tangent vectors and volume. Luckily, to obtain these geometric notions it suffices to introduce a single object, the metric tensor.

A *metric tensor* for $M$ is the datum of a scalar product on each tangent space $T_p$ of $M$, which varies smoothly on $p$: on a chart the scalar product may be expressed as a matrix, and we require that its coefficients vary smoothly on $p$.

DEFINITION 2.1. A *riemannian manifold* is a differentiable manifold equipped with a metric tensor which is positive definite at every point. Typically we denote it as a pair $(M, g)$, where $M$ is the manifold and $g$ is the tensor.

We introduce immediately two fundamental examples.

EXAMPLE 2.2. The *euclidean space* is the manifold $\mathbb{R}^n$ equipped with the euclidean metric tensor $g(x,y) = \sum_{i=1}^n x_i y_i$ at every tangent space $T_p = \mathbb{R}^n$.

EXAMPLE 2.3. Every differential submanifold $N$ in a riemannian manifold $M$ is also riemannian: it suffices to restrict for every $p \in N$ the metric tensor on $T_pM$ on the linear subspace $T_pN$.

In particular, the *sphere*

$$S^n = \left\{ x \in \mathbb{R}^{n+1} \mid \|x\| = 1 \right\}$$

is a submanifold of $\mathbb{R}^{n+1}$ and is hence riemannian.

The metric tensor $g$ defines in particular a norm for every tangent vector, and an angle between tangent vectors with the same basepoint. The velocity $\gamma'(t)$ of a curve $\gamma\colon I \to M$ at time $t \in I$ now has a module $|\gamma'(t)| \geqslant 0$ called *speed*, and two curves that meet at a point with non-zero velocities form a well-defined angle. The *length* of $\gamma$ may be defined as

$$L(\gamma) = \int_I |\gamma'(t)| dt$$

and can be finite or infinite. A *reparametrization* of $\gamma$ is the curve $\eta\colon J \to M$ obtained as $\eta = \gamma \circ \varphi$ where $\varphi\colon J \to I$ is a diffeomorphism of intervals. The length is invariant under reparametrization, that is $L(\gamma) = L(\eta)$.

**2.2. Distance, geodesics, volume.** Let $(M, g)$ be a connected riemannian manifold. The curves in $M$ now have a length and hence may be used to define a *distance* on $M$.

DEFINITION 2.4. The distance $d(p, q)$ between two points $p$ and $q$ is defined as

$$d(p, q) = \inf_\gamma L(\gamma)$$

where $\gamma$ varies among all curves $\gamma\colon [0, 1] \to M$ with $\gamma(0) = p$ and $\gamma(1) = q$.

The manifold $M$ equipped with the distance $d$ is a metric space (which induces on $M$ the same topology of $M$).

DEFINITION 2.5. A *geodetic* is a curve $\gamma\colon I \to M$ having constant speed $k$ and locally realizes the distance. This means that for any $t_0 \in I$ there is a $\epsilon > 0$ such that $d(\gamma(t), \gamma(t')) = L(\gamma|_{[t,t']}) = k|t - t'|$ for any $t, t' \in [t - \epsilon, t + \epsilon]$.

Note that with this definition the constant curve $\gamma(t) = p_0$ is a geodetic with constant speed $k = 0$. Such a geodesic is called *trivial*. A curve that realizes the distance locally may not realize them globally.

EXAMPLE 2.6. The non-trivial geodesics in euclidean space $\mathbb{R}^n$ are affine lines run at constant speed.

The non-trivial geodesics in the sphere $S^n$ are Le geodetiche non banali nella sfera $S^n$ are portions of great circles, run at constant speed.

If the differentiable manifold $M$ is oriented, the metric tensor also induces a *volume form*.

Briefly, the best method to define a volume in a $n$-manifold $M$ is to construct an appropriate *$n$-form.* A $n$-form $\omega$ is an alternating multilinear form

$$\omega_p : \underbrace{T_p \times \ldots \times T_p}_{n} \to \mathbb{R}$$

at each point $p \in M$, which varies smoothly with $p$. The alternating condition means that if we swap two vectors the result changes by a sign. Up to rescaling there exists only one $\omega_p$ which fulfills this condition: after identifying $T_p$ with $\mathbb{R}^n$ this is just the determinant.

The $n$-forms are useful because they can be integrated: it makes sense to write

$$\int_D \omega$$

on any open set $D$. A *volume form* on an oriented manifold $M$ is a $n$-form $\omega$ such that $\omega_p(v_1, \ldots, v_n) > 0$ for each positive basis $v_1, \ldots, v_n$ of $T_p$ and for every $p \in M$.

The metric tensor defines a volume form as follows: we simply set $\omega_p(e_1, \ldots, e_n) = 1$ on each positive orthonormal basis $e_1, \ldots, e_n$. With this definition every open set $D$ of $M$ has a well-defined *volume*

$$\mathrm{Vol}(D) = \int_D \omega$$

which is a positive number or infinity. If $D$ has compact closure the volume is necessarily finite. In particular, a compact riemannian manifold $M$ has finite volume $\mathrm{Vol}(M)$.

**2.3. Exponential map.** Let $(M, g)$ be a riemannian manifold. A geodetic $\gamma : I \to M$ is *maximal* if it cannot be extended to a geodesic on a strictly bigger domain $J \supset I$. Maximal geodesics are determined by some first-order conditions:

THEOREM 2.7. *Let $p \in M$ be a point and $v \in T_pM$ a tangent vector. There exists a unique maximal geodesic $\gamma : I \to M$ with $\gamma(0) = p$ and $\gamma'(0) = v$. The interval $I$ is open and contains* $0$.

This important fact has many applications. For instance, it allows to define the following notion.

DEFINITION 2.8. Let $p \in M$ be a point. The *exponential map* in $p$ is the map

$$\exp_p \colon U_p \to M$$

defined on a subset $U_p \subset T_p$ containing the origin as follows.

A vector $v \in T_p$ determines a maximal geodesic $\gamma_v \colon I_v \to M$ with $\gamma_v(0) = p$ and $\gamma_v'(0) = v$. Let $U$ be the set of vectors $v$ with $1 \in I_v$. For these vectors $v$ we define $\exp_p(v) = \gamma_v(1)$.

THEOREM 2.9. *The set $U_p$ is an open set containing the origin. The differential of the exponential map* $\exp_p$ *at the origin is the identity and hence* $\exp_p$ *is a local diffeomorphism at the origin.*

Via the exponential map, an open set of $T_p$ can be used as a chart near $p$: we recover here the intuitive idea that tangent space approximates the manifold near $p$.

**2.4. Injectivity radius.** The maximum radius where the exponential map is a diffeomorphism is called *injectivity radius*.

DEFINITION 2.10. The *injectivity radius* $\mathrm{inj}_p M$ of $M$ at a point $p$ is defined as follows:

$$\mathrm{inj}_p M = \sup \big\{ r > 0 \bigm| \exp_p|_{B_0(r)} \text{ is a diffeomorphism onto its image} \big\}.$$

Here $B_0(r)$ is the open ball with center 0 and radius $r$ in tangent space $T_p$. The injectivity radius is always positive by Theorem 2.9. For every $r$ smaller than the injectivity radius the exponential map transforms the ball of radius $r$ in $T_p$ into the ball of radius $r$ in $M$. That is, the following equality holds:

$$\exp_p(B_0(r)) = B_p(r)$$

and the ball $B_p(r)$ is indeed diffeomorphic to an open ball in $\mathbb{R}^n$. When $r$ is big this may not be true: for instance if $M$ is compact there is a $R > 0$ such that $B_p(R) = M$.

The injectivity radius $\mathrm{inj}_p(M)$ varies continuously with respect to $p \in M$; the injectivity radius $\mathrm{inj}(M)$ of $M$ is defined as

$$\mathrm{inj}(M) = \inf_{p \in M} \mathrm{inj}_p M.$$

PROPOSITION 2.11. *A compact riemannian manifold has positive injectivity radius.*

PROOF. The injectivity radius $\mathrm{inj}_p M$ is positive and varies continuously with $p$. □

Finally we note the following. A *closed curve* is a curve $\gamma\colon [a,b] \to M$ with $\gamma(a) = \gamma(b)$.

PROPOSITION 2.12. *Let $M$ be a riemannian manifold. A closed curve in $M$ of length smaller than* $2 \cdot \mathrm{inj}(M)$ *is homotopically trivial.*

PROOF. Set $x = \gamma(a) = \gamma(b)$. Since $\gamma$ is shorter than $2 \cdot \mathrm{inj}(M)$, it cannot escape the ball $B_x(r)$ for some $r < \mathrm{inj}(M) \leqslant \mathrm{inj}_x M$. This ball is diffeomorphic to a ball in $\mathbb{R}^n$, hence in particular it is contractible, so $\gamma$ is homotopically trivial. □

**2.5. Completeness.** A riemannian manifold $(M, g)$ is also a metric space, which can be complete or not. For instance, a compact riemannian manifold is always complete. On the other hand, by removing a point from a riemannian manifold we always get a non-complete space. Non-compact manifolds like $\mathbb{R}^n$ typically admit both complete and non-complete riemannian structures.

The completeness of a riemannian manifold may be expressed in various ways:

THEOREM 2.13 (Hopf-Rinow). *Let $(M, g)$ be a connected riemannian manifold. The following are equivalent:*

1. *$M$ is complete,*
2. *a subset of $M$ is compact if and only if it is closed and bounded,*
3. *every geodesic can be extended on the whole $\mathbb{R}$.*

*If $M$ is complete any two points $p, q \in M$ are joined by a* minimizing *geodesic $\gamma$,* i.e. *a curve such that $L(\gamma) = d(p, q)$.*

Note that (3) holds if and only if the exponential map is defined on the whole tangent space $T_p$ for all $p \in M$.

**2.6. Curvature.** The *curvature* of a riemannian manifold $(M, g)$ is a complicate object, typically defined from a *connection* $\nabla$ called *Levi-Civita connection*. The connection produces a tensor called *Riemann tensor* that records all the informations about the curvature of $M$.

We do not introduce this concepts because they are too powerful for the kind of spaces that we will encounter here: in hyperbolic geometry the manifolds have "constant curvature" and the full Riemann tensor is not necessary. It suffices to introduce the *sectional curvature* in a geometric way.

If $M$ has dimension 2, that is it is a surface, all the notions of curvature simplify and reduce to a unique quantity called *gaussian curvature*. If $M$ is contained in $\mathbb{R}^3$ the gaussian curvature is defined as the product of its two principal curvatures. If $M$ is abstract the principal curvatures however make no sense and hence we must take a different perspective.

We have seen in the previous section that on a riemannian manifold $(M, g)$, for every $p \in M$ there is an $\varepsilon > 0$ such that the ball $B_p(\varepsilon)$ centered in $p$ with radius $\varepsilon$ is really diffeomorphic to an open ball in $\mathbb{R}^n$.

The volume of this ball $B_p(\epsilon)$ is not necessarily equal to the volume of a euclidean ball of the same radius: it may be bigger or smaller, and this discrepancy is a measure of the curvature of $(M, g)$ at $p$.

DEFINITION 2.14. Let $(M, g)$ be a surface. The *gaussian curvature* at a point $p$ is defined as

$$K = \lim_{\epsilon \to 0} \left( \left( \pi\epsilon^2 - \mathrm{Vol}(B_p(\epsilon)) \right) \cdot \frac{12}{\pi\epsilon^4} \right).$$

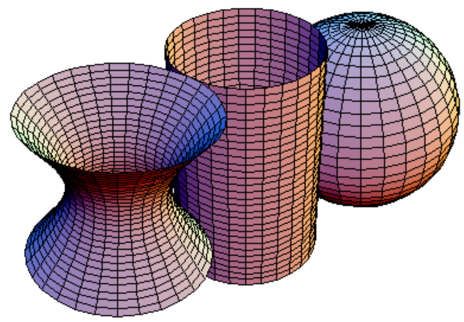


FIGURE 3. Three surfaces in space (hyperboloid of one sheet, cylinder, sphere) whose gaussian curvature is respectively negative, null, and positive at each point. The curvature on the sphere is constant, while the curvature on the hyperboloid varies: as we will see, a complete surface in $\mathbb{R}^3$ cannot have constant negative curvature.

In other words, the following formula holds:

$$\mathrm{Vol}(B_p(\epsilon)) = \pi\epsilon^2 - \frac{\pi\epsilon^4}{12}K + o(\epsilon^4).$$

The coefficient $\pi/12$ normalizes $K$ so that the curvature of a sphere of radius $R$ is $1/R^2$. We note in particular that $K$ is positive (negative) if $B_p(\epsilon)$ has smaller (bigger) area than the usual euclidean area.

If $(M, g)$ has dimension $n \geqslant 3$ we may still define a curvature by evaluating the difference between $\mathrm{Vol}(B_p(\epsilon))$ and the volume of a euclidean ball: we obtain a number called *scalar curvature*. gaussiana. The scalar curvature in dimension $\geqslant 3$ is however only a poor description of the curvature of the manifold, and one usually looks for some more refined notion which contains more geometric informations. The curvature of $(M, g)$ is typically encoded by one of the following two objects: the *Riemann tensor* or the *sectional curvature*. These objects are quite different but actually contain the same amount of information. We introduce here the sectional curvature.

DEFINITION 2.15. Let $(M, g)$ be a riemannian manifold. Let $p \in M$ be a point and $W \subset T_pM$ a 2-dimensional vector subspace. By Theorem 2.9 there exists an open set $U_p \subset T_pM$ containing the origin where $\exp_p$ is a diffeomorphism onto its image. In particular $S = \exp_p(U_p \cap W)$ is a small smooth surface in $M$ passing through $p$, with tangent plane $W$. As a submanifold it has a riemannian structure induced by $g$.

The *sectional curvature* of $(M, g)$ along $(p, W)$ is defined as the gaussian curvature of $S$ in $p$.

The sectional curvature is hence a number assigned to every pair $(p, W)$ where $p \in M$ is a point and $W \subset T_pM$ is a 2-dimensional vector space.

DEFINITION 2.16. A riemannian manifold $(M, g)$ has *constant sectional curvature* $K$ if the sectional curvature assigned to every $p \in M$ and every 2-dimensional vector space $W \subset T_pM$ is always $K$.

Remark 2.17. On a riemannian manifold $(M, g)$ one may *rescale* the metric of some factor $\lambda > 0$ by substituting $g$ with the tensor $\lambda g$. At every point the scalar product is rescaled by $\lambda$. Lengths of curves are rescaled by $\sqrt{\lambda}$ and volumes are rescaled by $\lambda^{\frac{n}{2}}$. The sectional curvature is rescaled by $1/\lambda$.

By rescaling the metric it is hence possible to transform a riemannian manifold with constant sectional curvature $K$ into one with constant sectional curvature $-1$, $0$, or $1$.

Example 2.18. Euclidean space $\mathbb{R}^n$ has constant curvature zero. A sphere of radius $R$ has constant curvature $1/R^2$.

**2.7. Isometries.** Every honest category has its morphisms. Riemannian manifolds are so rigid, that in fact one typically introduces only isomorphisms: these are called *isometries.*

Definition 2.19. A diffeomorphism $f : M \to N$ between two riemannian manifolds $(M, g)$ e $(N, h)$ is an *isometry* if it preserves the scalar product. That is, the equality

$$\langle v, w\rangle = \langle df_p(v), df_p(w)\rangle$$

holds for all $p \in M$ and every pair of vectors $v, w \in T_pM$. The symbols $\langle , \rangle$ indicate the scalar products in $T_x$ and $T_{f(x)}$.

As we said, isometries are extremely rigid. These are determined by their first-order behavior at any single point.

Theorem 2.20. *Let $f, g : M \to N$ be two isometries between two connected riemannian manifolds. If there is a point $p \in M$ such that $f(p) = g(p)$ and $df_p = dg_p$, then $f = g$ everywhere.*

Proof. Let us show that the subset $S \subset M$ of the points $p$ such that $f(p) = g(p)$ and $df_p = dg_p$ is open and closed.

The locus where two functions coincide is typically closed, and this holds also here (to prove it, take a chart). We prove that it is open: pick $p \in S$. By Theorem 2.9 there is an open neighborhood $U_p \subset T_pM$ of the origin where the exponential map is a diffeomorphism onto its image. We show that the open set $\exp_p(U_p)$ is entirely contained in $S$.

A point $x \in \exp_p(U_p)$ is the image $x = \exp(v)$ of a vector $v \in U_p$ and hence $x = \gamma(1)$ for the geodetic $\gamma$ determined by the data $\gamma(0) = p, \gamma'(0) = v$. The maps $f$ and $g$ are isometries and hence send geodesics to geodesics: here $f \circ \gamma$ and $g \circ \gamma$ are geodesics starting from $f(p) = g(p)$ with the same initial velocities and thus they coincide. This implies that $f(x) = g(x)$. Since $f$ and $g$ coincide on the open set $\exp_p(U_p)$, also their differentials do. $\square$

The isometries $f : M \to M$ from a manifold $M$ to itself form the *isometry group* of $M$, denoted $\mathrm{Isom}(M)$.

**2.8. Riemannian manifolds with boundary.** Many geometric notions in riemannian geometry extend easily to manifolds $M$ with boundary: a metric tensor on $M$ is a positive definite scalar product on each (semi-)space $T_x$ that varies smoothly in $x \in M$. The boundary $\partial M$ of a riemannian manifold $M$ is naturally itself a riemannian manifold without boundary.

The exponential map and the injectivity radius $\mathrm{inj}_x M$ of a boundary point $x \in \partial M$ are still defined as in Section 2.3, taking into account that the tangent space $T_x$ is actually only a half-space.

## 3. Measure theory

We will use some basic measure theory only in two points in this book.

**3.1. Borel measure.** A *Borel set* in a topological space $X$ is any set obtained from open sets through the operations of countable union, countable intersection, and relative complement. Let $\mathscr{F}$ denote the set of all Borel sets. A *Borel measure* on $X$ is a function $\mu\colon \mathscr{F} \to [0, +\infty]$ which is additive on any countable collection of disjoint sets.

The measure is *locally finite* if every point has a neighborhood of finite measure and is *trivial* if $\mu(S) = 0$ for all $S \in \mathscr{F}$.

Exercise 3.1. If $\mu$ is a locally finite Borel measure then $\mu(X) < +\infty$ for any compact Borel set $K \subset X$.

Example 3.2. Let $D \subset X$ be a discrete set. The *Dirac measure* $\delta_D$ concentrated in $D$ is the measure

$$\delta_D(S) = \#(S \cap D).$$

Since $D$ is discrete $\delta_D$ is locally finite.

The *support* of a measure is the set of all points $x \in X$ such that $\mu(U) > 0$ for any open set $U$ containing $x$. The support is a closed subset of $X$. The measure is *fully supported* if its support is $X$. The support of $\delta_D$ is of course $D$. A measure can be defined using local data by the following.

Proposition 3.3. *Let $\{U_i\}_{i\in I}$ be a countable, locally finite open covering of $X$ and for any $i \in I$ let $\mu_i$ be a locally finite Borel measure on $U_i$. If $\mu_i|_{U_i\cap U_j} = \mu_j|_{U_i\cap U_j}$ for all $i, j \in I$ there is a unique locally finite Borel measure $\mu$ on $X$ whose restriction to $U_i$ is $\mu_i$ for all $i$.*

Proof. For every finite subset $J \subset I$ define $X_J = \big(\cap_{j\in J} U_j\big) \setminus \big(\cup_{i\in I\setminus J} U_i\big)$. The sets $X_J$ form a countable partition of $X$ into Borel sets and every $X_J$ is equipped with a measure $\mu_J = \mu_j|_{X_j}$ for any $j \in J$. Define $\mu$ by setting

$$\mu(S) = \sum_{j\in J} \mu(S \cap X_j)$$

on any Borel $S \subset X$. $\square$

When $X$ is a reasonable space some hypothesis may be dropped.

Proposition 3.4. *If $X$ is paracompact and separable, Proposition 3.3 holds for any open covering $\{U_i\}_{i\in I}$.*

Proof. By paracompactness and separability he open covering $\{U_i\}$ has a refinement that is locally finite and countable: apply Proposition 3.3 to get a unique measure $\mu$. To prove that indeed $\mu|_{U_i} = \mu_i$ apply Proposition 3.3 again to the covering of $U_i$ given by the refinement. □

If a group $G$ acts on a set $X$ we say that a measure $\mu$ is *$G$-invariant* if $\mu(g(A)) = \mu(A)$ for any Borel set $A$ and any $g \in G$.

**3.2. Topology on the measure space.** In what follows we suppose for simplicity that $X$ is a finite-dimensional topological manifold, although everything is valid in a wider generality. We indicate by $\mathscr{M}(X)$ the space of all locally finite Borel measures on $X$ and by $C_c(X)$ the space of all continuous functions $M \to \mathbb{R}$ with compact support: the space $C_c(X)$ is not a Banach space, but is a topological vector space.

Recall that the *topological dual* of a topological vector space $V$ is the vector space $V^*$ formed by all continuous linear functionals $V \to \mathbb{R}$. A measure $\mu \in \mathscr{M}(X)$ acts like a continuous functional on $C_c(X)$ as follows

$$\mu\colon f \longmapsto \int_\mu f$$

and hence defines an element of $C_c^*(X)$. A functional in $C_c^*(X)$ is *positive* if it assumes non-negative values on non-negative functions.

Theorem 3.5 (Riesz representation). *The space $\mathscr{M}(X)$ may be identified in this way to the subset in $C_c(X)^*$ of all positive functionals.*

The space $\mathscr{M}(X)$ in $C_c(X)^*$ is closed with respect to sum and product with a positive scalar.

Definition 3.6. Let $V$ be a real topological vector space. Every vector $v \in V$ defines a functional in $V^*$ as $f \mapsto f(v)$. The *weak-* topology* on $V^*$ is the weakest topology among those where these functionals are continuous.

We give $C_c(X)^*$ the weak-* topology. With this topology a sequence of measures $\mu_i$ converges to $\mu$ if and only if $\int_{\mu_i} f \to \int_\mu f$ for any $f \in C_c(X)$. This type of weak convergence is usually denoted with the symbol $\mu_i \rightharpoonup \mu$.

Exercise 3.7. Let $x_n$ be a sequence of points in $X$ that tends to $x \in X$: hence $\delta_{x_n} \rightharpoonup \delta_x$.

**3.3. Lie groups.** A *Lie group* is a smooth manifold $G$ which is also a group, such that the operations

$$G \times G \to G, \quad (a,b) \mapsto ab$$
$$G \to G, \quad a \mapsto a^{-1}$$

are smooth.

A non-trivial group $G$ is è *simple* if it has no normal subgroups except $G$ and $\{e\}$. The definition on Lie groups is a bit different.

DEFINITION 3.8. A Lie group $G$ is *simple* if it is connected, non abelian, and has no connected normal groups except $G$ and $\{e\}$.

**3.4. Haar measures.** Let $G$ be a Lie group.

DEFINITION 3.9. A *left-invariant Haar measure* on $G$ is a locally finite fully supported Borel measure $\mu$ on $G$, invariant by the left action of $G$.

THEOREM 3.10 (Haar theorem). *A Lie group $G$ has a left-invariant Haar measure, unique up to rescaling.*

A right-invariant Haar measure is defined analogously and is also unique up to rescaling. The group $G$ is *unimodular* if a left-invariant Haar measure is also right-invariant.

If $\mu$ is right-invariant and $g \in G$ is an element, the measure $\mu^g(A) = \mu(g^{-1}A)$ is also right-invariant, and by uniqueness $\mu^g = \lambda_g \mu$ for some $\lambda_g > 0$. The *modular function* $g \mapsto \lambda_g$ is a continuous homomorphism $\lambda\colon G \to \mathbb{R}_{>0}$. The group $G$ is unimodular if and only if its modular function is trivial.

PROPOSITION 3.11. *Compact, abelian, discrete, and simple groups are unimodular.*

PROOF. If $G$ is compact every continuous homomorphisms to $\mathbb{R}_{>0}$ is trivial. If $G$ is simple, the normal subgroup $\ker \lambda$ is trivial or $G$, and the first case is excluded because $G$ is not abelian. If $G$ is abelian, left- and right-measures obviously coincide. If $G$ is discrete every singleton has the same measure and hence left- and right- measures coincide. □

EXAMPLE 3.12. The group $\mathrm{Aff}(\mathbb{R}) = \{x \mapsto ax + b \mid a \in \mathbb{R}^*, b \in \mathbb{R}\}$ of affine transformations in $\mathbb{R}$ is not unimodular.

## 4. Cell complexes and handle decompositions

**4.1. Cell complex.** Recall that a *finite cell complex* of dimension $k$ (briefly, a $k$-complex) is a topological space obtained iteratively in the following manner:

- a 0-complex $X^0$ is a finite number of points,
- a $k$-complex $X^k$ is obtained from a $(k-1)$-complex $X^{k-1}$ by attaching finitely many $k$-*cells*, that is copies of $D^k$ glued along continuous maps $\varphi\colon \partial D^k \to X^{k-1}$.

The subset $X^i \subset X^k$ is a closed subset called the $i$-*skeleton*, for all $i < k$.

PROPOSITION 4.1. *The inclusion map $i\colon X^i \hookrightarrow X$ induces an isomorphism $i_*\pi_j(X^i) \to \pi_j(X)$ for all $i$.*

PROOF. Maps $S^j \to X$ and homotopies between them can be homotoped away from cells of dimension $\geqslant j+2$. □

In particular, the space $X$ is connected if and only if $X^1$ is, and its fundamental group of $X$ is captured by $X^2$.

Recall that a *finite presentation* of a group $G$ is a description of $G$ as

$$\langle g_1, \ldots, g_k \mid r_1, \ldots, r_s \rangle$$

where $g_1, \ldots, g_k \in G$ are the *generators* and $r_1, \ldots, r_s$ are words in $g_i^{\pm 1}$ called *relations*, such that

$$G \cong F(g_i)/_{N(r_j)}$$

where $F(g_i)$ is the free group generated by the $g_i$'s and $N(r_j) \triangleleft F(g_i)$ is the *normalizer* of the $r_j$'s, the smallest normal subgroup containing them.

A presentation for the fundamental group of $X$ can be constructed as follows. If $x_0 \in X^0$, we fix a maximal tree $T \subset X^1$ containing $x_0$ and give the $k$ arcs in $X^1 \setminus T$ some arbitrary orientations. These arcs determine some generators $g_1, \ldots, g_k \in \pi_1(X, x_0)$. The boundary of a 2-cell makes a circular path in $X^1$: every time it crosses an arc $g_i$ in some direction (entering from one side and exiting from the other) we write the corresponding letter $g_i^{\pm 1}$ and get a word. The $s$ two-cells produce $s$ relations. We have constructed a presentation for $\pi_1(X)$.

THEOREM 4.2. *Every differentiable compact $n$-manifold may be realized topologically as a finite $n$-complex.*

**4.2. Aspherical cell-complexes.** A finite cell complex is locally contractible and hence has a universal covering $\widetilde{X}$; if $\widetilde{X}$ is contractible the complex $X$ is called *aspherical*.

THEOREM 4.3. *Let $X, Y$ be connected finite cellular complexes with basepoints $x_0 \in X^0$, $y_0 \in Y^0$ and $f\colon \pi_1(X, x_0) \to \pi_1(Y, y_0)$ a homomorphism. If $Y$ is aspherical there is a continuous map $F\colon (X, x_0) \to (Y, y_0)$ that induces $f$, unique up to homotopy.*

PROOF. We construct $f$ iteratively on $X^i$, starting from $i = 1$. Let $T$ be a maximal tree in $X^1$. The oriented 1-cells $g_1, \ldots, g_k$ in $X^1 \setminus T$ define generators in $\pi_1(X, x_0)$: we construct $F$ by sending each $g_i$ to any loop in $Y$ representing $f(g_i)$.

The map $F$ sends homotopically trivial loops in $X^1$ to homotopically trivial loops in $Y$ and hence extends to $X^2$. Since $\widetilde{Y}$ is aspherical, the higher homotopy groups $\pi_i(Y)$ with $i \geqslant 2$ vanish and hence $F$ extends to $X^3, \ldots, X^k = X$ step by step.

We prove that $F\colon (X, x_0) \to (Y, y_0)$ is unique up to homotopy. Take another $F'$ that realizes $f$, and construct a homotopy $F \sim F'$ iteratively on $X^i$. For $i = 1$, we can suppose that both $F$ and $F'$ send $T$ to $y_0$, then use $F_* = f = F'_*$ to homotope $F'$ to $F$ on $X^1$. The maps $F$ and $F'$ on a $i$-cell for $i \geqslant 2$ are homotopic because they glue to a map $S^i \to Y$ which is null-homotopic because $\pi_i(Y)$ is trivial. $\square$

Corollary 4.4. *Let $X$ and $Y$ be connected finite aspherical complexes. Every isomorphism $f\colon \pi_1(X) \to \pi_1(Y)$ is realized by a homotopic equivalence, unique up to homotopy.*

Corollary 4.5. *Two aspherical closed manifolds of distinct dimension have non-isomorphic fundamental groups.*

Proof. Two closed manifolds of different dimension cannot be homotopically equivalent because they have non-isomorphic homology groups. □

We cite for completeness this result, which we will never use.

Theorem 4.6 (Cartan-Hadamard). *A complete riemannian manifold $M$ with sectional curvature everywhere $\leqslant 0$ has a universal covering diffeomorphic to $\mathbb{R}^n$ and is hence aspherical.*

Sketch of the proof. Pick a point $x \in M$. Since $M$ is complete, the exponential map $\exp_x\colon T_x \to M$ is defined on $T_x$. The fact that the sectional curvatures are $\leqslant 0$ imply that $(d\exp_x)_y$ is invertible for any $y \in T_x$ and $\exp_x$ is a covering. □

CHAPTER 2

# Hyperbolic space

We introduce in this chapter the hyperbolic space $\mathbb{H}^n$.

## 1. The models of hyperbolic space

In every dimension $n \geqslant 2$ there exists a unique complete, simply connected riemannian manifold having constant sectional curvature 1, 0, or $-1$ up to isometries. These three manifolds are extremely important in riemannian geometry because they are the fundamental models to construct and study non-simply connected manifolds with constant curvature.

The three manifolds are respectively the *sphere* $S^n$, *euclidean space* $\mathbb{R}^n$, and *hyperbolic space* $\mathbb{H}^n$. As we will see, every complete manifold with constant curvature has one of these three spaces as its universal cover.

In contrast with $S^n$ and $\mathbb{R}^n$, hyperbolic space $\mathbb{H}^n$ can be constructed using various *models*, none of which is prevalent.

**1.1. Hyperboloid.** The sphere $S^n$ consists of all points with norm 1 in $\mathbb{R}^{n+1}$, considered with the euclidean scalar product. Analogously, we may define $\mathbb{H}^n$ as the set of all points of norm $-1$ in $\mathbb{R}^{n+1}$, considered with the usual *lorentzian* scalar product. This set forms a hyperboloid of two sheets, and we choose one.

Definition 1.1. Consider $\mathbb{R}^{n+1}$ equipped with the lorentzian scalar product of segnature $(n, 1)$:

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i - x_{n+1}y_{n+1}.$$

A vector $x \in \mathbb{R}^{n+1}$ is *time-like*, *light-like* or *space-like* if $\langle x, x\rangle$ is negative, null, or positive respectively. The *hyperboloid model* $I^n$ is defined as follows:

$$I^n = \big\{x \in \mathbb{R}^{n+1} \bigm| \langle x, x\rangle = -1,\ x_n > 0\big\}.$$

The set of points $x$ with $\langle x, x\rangle = -1$ is a *hyperboloid with two sheets*, and $I^n$ is the connected component (sheet) with $x_{n+1} > 0$. Let us prove a general fact. For us, a scalar product is a real non-degenerate symmetric bilinear form.

Proposition 1.2. *Let $\langle , \rangle$ be a scalar product on $\mathbb{R}^n$. The function $f \colon \mathbb{R}^n \to \mathbb{R}$ given by*

$$f(x) = \langle x, x\rangle$$

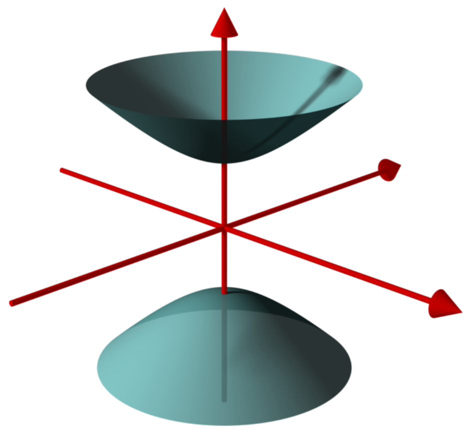


FIGURE 1. The hyperboloid with two sheets defined by the equation $\langle x, x\rangle = -1$. The model $I^n$ is the upper connected component.

*is everywhere smooth and has differential*

$$df_x(y) = 2\langle x, y\rangle.$$

PROOF. The following equality holds:

$$\langle x+y, x+y\rangle = \langle x, x\rangle + 2\langle x, y\rangle + \langle y, y\rangle.$$

The component $\langle x, y\rangle$ is linear in $y$ while $\langle y, y\rangle$ is quadratic. □

COROLLARY 1.3. *The hyperboloid $I^n$ is a riemannian manifold.*

PROOF. The hyperboloid is the set of points with $f(x) = \langle x, x\rangle = -1$; for all $x \in I^n$ the differential $y \mapsto 2\langle x, y\rangle$ is surjective and hence the hyperboloid is a differential submanifold of codimension 1.

The tangent space $T_xI^n$ at $x \in I^n$ is the hyperplane

$$T_x = \ker df_x = \big\{y \bigm| \langle x, y\rangle = 0\big\} = x^\perp$$

orthogonal to $x$ in the lorentzian scalar product. Since $x$ is time-like, the restriction of the lorentzian scalar product to $x^\perp$ is positive definite and hence defines a metric tensor on $I^n$. □

The hyperboloid $I^n$ is a model for hyperbolic space $\mathbb{H}^n$. We will soon prove that it is indeed simply connected, complete, and has constant curvature $-1$.

**1.2. Isometries of the hyperboloid.** The isometries of $I^n$ are easily classified using linear algebra.

Let $O(n, 1)$ be the group of linear isomorphisms $f$ of $\mathbb{R}^{n+1}$ that preserve the lorentzian scalar product, *i.e.* such that $\langle v, w\rangle = \langle f(v), f(w)\rangle$ for any $v, w \in \mathbb{R}^n$. An element in $O(n, 1)$ preserves the hyperboloid of two sheets, and the elements preserving the upper sheet $I^n$ form a subgroup of index two in $O(n, 1)$ that we indicate with $O_*(n, 1)$.

PROPOSITION 1.4. *The following equality holds:*

$$\mathrm{Isom}(I^n) = O_*(n,1).$$

PROOF. Pick $f \in O_*(n,1)$. If $x \in I^n$ then $f(x) \in I^n$ and $f$ sends $x^\perp$ to $f(x)^\perp$ isometrically, hence $f \in \mathrm{Isom}(I^n)$.

To prove $\mathrm{Isom}(I^n) \subseteq O_*(n,1)$ we show that for every pair $x, y \in I^n$ and every linear isometry $g\colon x^\perp \to y^\perp$ there is an element $f \in O_*(n,1)$ such that $f(x) = y$ and $f|_{x^\perp} = g$. Since isometries are determined by their first-order behavior at a point $x$, they are all contained in $O_*(n,1)$.

Simple linear algebra shows that $O_*(n,1)$ acts transitively on points of $I^n$ and hence we may suppose that $x = y = (0,\ldots,0,1)$. Now $x^\perp = y^\perp$ is the horizontal hyperplane and $g \in O(n)$. To define $f$ simply take

$$f = \begin{pmatrix} g & 0 \\ 0 & 1 \end{pmatrix}.$$

$\square$

Two analogous results hold for $S^n$ and $\mathbb{R}^n$:

PROPOSITION 1.5. *The following equalities hold:*

$$\mathrm{Isom}(S^n) = O(n+1),$$
$$\mathrm{Isom}(\mathbb{R}^n) = \big\{x \mapsto Ax + b \;\big|\; A \in O(n), b \in \mathbb{R}^n\big\}.$$

PROOF. The proof is analogous to the one above. $\square$

We have also proved the following fact. A *frame* at a point $p$ in a riemannian manifold $M$ is an orthonormal basis for $T_pM$.

COROLLARY 1.6. *Let $M = S^n$, $\mathbb{R}^n$, or $I^n$. Given two points $p, q \in M$ and two frames in $p$ and $q$, there is a unique isometry that carries the first frame to the second.*

The corollary says that $S^n$, $\mathbb{R}^n$ e $I^n$ have the "maximum possible number" of isometries.

**1.3. Subspaces.** Each $S^n$, $\mathbb{R}^n$, and $\mathbb{H}^n$ contains various *subspaces* of smaller dimension.

DEFINITION 1.7. A $k$-dimensional *subspace* of $\mathbb{R}^n$, $S^n$, $I^n$ is respectively:

- an affine $k$-dimensional space in $\mathbb{R}^n$,
- the intersection of a $(k+1)$-dimensional subspace of $\mathbb{R}^{n+1}$ with $S^n$,
- the intersection of a $(k+1)$-dimensional subspace of $\mathbb{R}^{n+1}$ with $I^n$, when it is non-empty.

Concerning non-emptyness, elementary linear algebra shows that the following conditions are equivalent for any $(k+1)$-dimensional subspace $W \subset \mathbb{R}^{n+1}$:

(1) $W \cap I^n \neq \emptyset$,
(2) $W$ contains at least a time-like vector,

(3) the segnature of $\langle , \rangle|_W$ is $(k, 1)$.

A $k$-subspace in $\mathbb{R}^n, S^n, \mathbb{H}^n$ is itself isometric to $\mathbb{R}^k, S^k, \mathbb{H}^k$. The non-empty intersection of two subspaces is always a subspace. An isometry of $\mathbb{R}^n, S^n, \mathbb{H}^n$ sends $k$-subspaces to $k$-subspaces.

A 1-subspace is a *line*. We now show that lines and geodesics are the same thing. We recall the hyperbolic trigonometric functions:

$$\sinh(t) = \frac{e^t - e^{-t}}{2}, \quad \cosh(t) = \frac{e^t + e^{-t}}{2}.$$

PROPOSITION 1.8. *A non-trivial complete geodesic in $S^n$, $\mathbb{R}^n$ and $\mathbb{H}^n$ is a line run at constant speed. Concretely, let $p \in M$ be a point and $v \in T_pM$ a unit vector. The geodesic $\gamma$ exiting from $p$ with velocity $v$ is:*

- $\gamma(t) = \cos(t) \cdot p + \sin(t) \cdot v$ *if* $M = S^n$,
- $\gamma(t) = p + tv$ *if* $M = \mathbb{R}^n$,
- $\gamma(t) = \cosh(t) \cdot p + \sinh(t) \cdot v$ *if* $M = I^n$.

PROOF. The proof for $\mathbb{R}^n$ is trivial. If $M = S^n$ or $I^n$ let $W \subset \mathbb{R}^{n+1}$ be the vector subspace generated by $p$ and $v$. Let $f \in O(n)$ or $f \in O_*(n, 1)$ be the isometry such that $f|_W = \mathrm{id}$ and $f|_{W^\perp} = -\mathrm{id}$. This induces an isometry of $S^n$ or $I^n$ that fixes $p$ and $v$, and hence fixes $\gamma$. Therefore $\gamma$ is contained in the line $W \cap S^n$ or $W \cap I^n$.

To prove that $\gamma(t)$ has the form described above we only need to check that it has constant unit speed. The velocity in $I^n$ is indeed

$$\gamma'(t) = \cosh'(t) \cdot p + \sinh'(t) \cdot v = \sinh(t) \cdot p + \cosh(t) \cdot v$$

which has squared norm $\sinh^2(t) - \cosh^2(t) = 1$. $\square$

COROLLARY 1.9. *The spaces $S^n$, $\mathbb{R}^n$ e $\mathbb{H}^n$ are complete.*

PROOF. The previous proposition shows that geodesics are defined on $\mathbb{R}$, hence the space is complete by Hopf-Rinow. $\square$

Finally, it is easy to show that two points in $\mathbb{H}^n$ are contained in a unique line.

REMARK 1.10. Euclid's V postulate holds only in $\mathbb{R}^2$: given a line $r$ and a point $P \notin r$, there is only one line passing through $P$ and disjoint from $r$ (in $\mathbb{R}^2$), there is noone (in $S^2$), or there are infinitely many (in $\mathbb{H}^2$).

**1.4. The Poincaré disc.** We introduce two models of $\mathbb{H}^n$ (the *disc* and *half-space*) that are easier to visualize especially in the dimensions $n = 2, 3$ we are interested in. The first model is the *Poincaré disc*

$$D^n = \{x \in \mathbb{R}^n \mid \|x\| < 1\}.$$

The metric tensor on $D^n$ is obviously not the euclidean one of $\mathbb{R}^n$, but it is the one induced by a particular diffeomorfism between $I^n$ and $D^n$ that we construct now. We identify $\mathbb{R}^n$ with the horizontal hyperplane $x_{n+1} = 0$ in $\mathbb{R}^{n+1}$ and note that the linear projection on $P = (0, \ldots, 0, -1)$ described

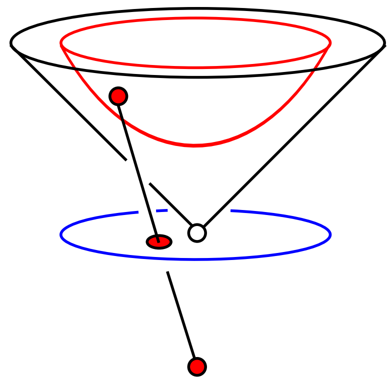


FIGURE 2. The projection on $P = (0, \ldots, 0, -1)$ induces a bijection between $I^n$ and $D^n$.

in Fig. 2 induces a bijection between $I^n$ and the horizontal disc $D^n \subset \mathbb{R}^n$, see Fig. 2. The projection $p$ may be written as:

$$p(x_1, \ldots, x_{n+1}) = \frac{(x_1, \ldots, x_n)}{x_{n+1} + 1}$$

and is indeed a diffeomorphism $p\colon I^n \to D^n$ that carries the metric tensor on $I^n$ onto a metric tensor $g$ on $D^n$.

PROPOSITION 1.11. *The metric tensor $g$ at $x \in D^n$ is:*

$$g_x = \left(\frac{2}{1 - \|x\|^2}\right)^2 \cdot g^E_x$$

*where $g^E$ is the euclidean metric tensor on $D^n \subset \mathbb{R}^n$.*

PROOF. The inverse map $p^{-1}$ is:

$$p^{-1}(x) = \left(\frac{2x_1}{1 - \|x\|^2}, \ldots, \frac{2x_n}{1 - \|x\|^2}, \frac{1 + \|x\|^2}{1 - \|x\|^2}\right).$$

Pick $x \in D^n$. A rotation around the $x_{n+1}$ axis is an isometry for both $I^n$ and $(D^n, g)$. Hence up to rotating we may take $x = (x_1, 0, \ldots, 0)$ and

$$p^{-1}(x) = \left(\frac{2x_1}{1 - x_1^2}, 0, \ldots, 0, \frac{1 + x_1^2}{1 - x_1^2}\right).$$

The differential at $x$ acts on the canonical basis $e_i$ of $\mathbb{R}^n$ as follows:

$$dp^{-1}_x\colon e_1 \longmapsto \frac{2}{(1 - x_1^2)^2}\left(1 + x_1^2, 0, \ldots, 0, 2x_1\right),$$

$$dp^{-1}_x\colon e_i \longmapsto \frac{2}{1 - x_1^2} e_i \quad \forall i = 2, \ldots, n.$$

The images are orthogonal vectors (with respect to the lorentzian scalar product) of norm $\frac{2}{1 - x_1^2}$, hence $g_x = \frac{4}{(1 - x_1^2)^2} g^E_x$. $\square$

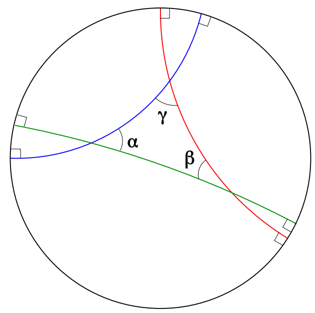



FIGURE 3. Three lines that determine a hyperbolic triangle in the Poincaré disc. The angles $\alpha, \beta$ e $\gamma$ coincide with the euclidean ones.

The Poincaré disc is a *conformal* model of $\mathbb{H}^n$: it is a model where the metric differs from the Euclidean metric only by multiplication by a positive scalar $(\frac{2}{1-\|x\|^2})^2$ that depends smoothly on $x$. We note that the scalar tends to infinity when $x$ tends to $\partial D^n$. On a conformal model lengths of vectors are different than the euclidean lengths, but the angles formed by two adjacent vectors coincide with the euclidean ones. Shortly: lengths are distorted but angles are preserved.

Let us see how to visualize $k$-subspaces in the disc model.

PROPOSITION 1.12. *The $k$-subspaces in $D^n$ are the intersections of $D^n$ with $k$-spheres and $k$-planes of $\mathbb{R}^n$ orthogonal to $\partial D^n$.*

PROOF. Since every $k$-subspace is an intersection of hyperplanes, we easily restrict to the case $k = n-1$. A hyperplane in $I^n$ is $I^n \cap v^\perp$ for some space-like vector $v$. If $v$ is horizontal (*i.e.* its last coordinate is zero) then $v^\perp$ is vertical and $p(I^n \cap v^\perp) = D^n \cap v^\perp$, a hyperplane orthogonal to $\partial D^n$.

If $v$ is not horizontal, up to rotating around $x_{n+1}$ we may suppose $v = (\alpha, 0, \ldots, 0, 1)$ with $\alpha > 1$. The hyperplane is

$$\left\{x_1^2 + \ldots + x_n^2 - x_{n+1}^2 = -1\right\} \cap \left\{x_{n+1} = \alpha x_1\right\}.$$

On the other hand the sphere in $\mathbb{R}^n$ of center $(\alpha, 0, \ldots, 0)$ and radius $\alpha^2 - 1$ is orthogonal to $\partial D^n$ and is described by the equation

$$\left\{(y_1 - \alpha)^2 + y_2^2 + \ldots + y_n^2 = \alpha^2 - 1\right\} = \left\{y_1^2 + \ldots + y_n^2 - 2\alpha y_1 = -1\right\}$$

which is equivalent to $||y||^2 = -1 + 2\alpha y_1$. If $y = p(x)$ the relations

$$y_1 = \frac{x_1}{x_{n+1}+1}, \quad \|y\|^2 = \frac{x_{n+1}-1}{x_{n+1}+1}$$

trasform the latter equation in $x_{n+1} = \alpha x_1$. $\square$

Three lines in $D^2$ are drawn in Fig. 3. Since $D^n$ is a conformal model, the angles $\alpha, \beta$, and $\gamma$ they make are precisely those one sees from the picture. In particular we verify easily that $\alpha + \beta + \gamma < \pi$.

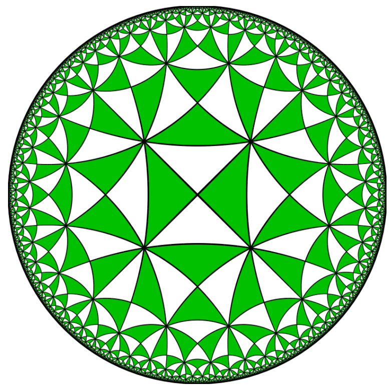


FIGURE 4. A *tessellation* of $S^2$, $\mathbb{R}^2$ o $\mathbb{H}^2$ is a subdivision of the plane into polygons. The tessellation of $\mathbb{H}^2$ shown here is obtained by drawing infinitely many lines in the plane. The triangles have inner angles $\frac{\pi}{2}, \frac{\pi}{10}, \frac{\pi}{10}$ and are all isometric.

**1.5. The half-space model.** We introduce another conformal model. The *half-space model* is the space

$$H^n = \big\{(x_1, \ldots, x_n) \in \mathbb{R}^n \ \big| \ x_n > 0\big\}.$$

The model $H^n$ is obtained from $D^n$ through a particular diffeomorphism, called *inversion*.

DEFINITION 1.13. Let $S = S(x_0, r)$ be the sphere in $\mathbb{R}^n$ centered in $x_0$ and with radius $r$. The *inversion* along $S$ is the map $\varphi\colon \mathbb{R}^n \setminus \{x_0\} \to \mathbb{R}^n \setminus \{x_0\}$ defined as follows:

$$\varphi(x) = x_0 + r^2 \frac{x - x_0}{\|x - x_0\|^2}.$$

The map may be extended on the whole sphere $S^n$, identified with $\mathbb{R}^n \cup \{\infty\}$ through the stereographic projection, by setting $\varphi(x_0) = \infty$ and $\varphi(\infty) = x_0$. A geometric description of inversion is given in Fig. 5.

We have already talked about conformal models. More generally, a diffeomorphism $f\colon M \to N$ between two oriented riemannian manifolds is *conformal* (risp. *anticonformal*) if for any $p \in M$ the differential $df_p$ is the product of a scalar $\lambda_p > 0$ and an isometry that preserves (risp. inverts) the orientation.

The scalar $\lambda_p$ depends on $p$. A conformal map preserves the angle between two vectors but modifies their lengths by multiplication by $\lambda_p$.

PROPOSITION 1.14. *The following holds:*

(1) *an inversion is a smooth and anticonformal map;*
(2) *an inversion sends $k$-spheres and $k$-planes into $k$-spheres and $k$-planes.*

PROOF. Up to translations we may suppose $x_0 = 0$. The inversion is $\varphi(x) = r^2 \frac{x}{\|x\|^2}$ and we now prove that $d\varphi_x$ is $\frac{r^2}{\|x\|^2}$ times a reflection

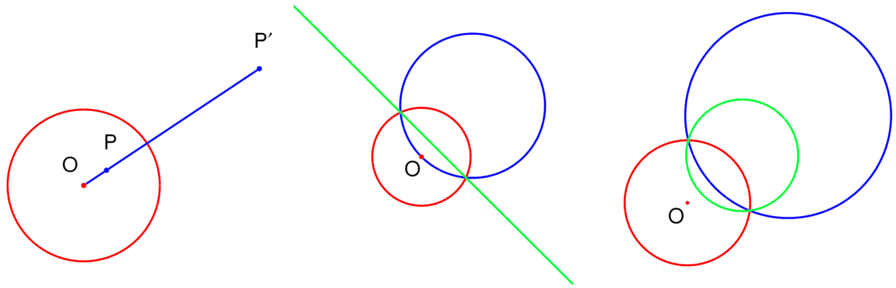



FIGURE 5. The inversion trrough a sphere of center $O$ and radius $r$ moves $P$ to $P'$ so that $OP \times OP' = r^2$ (left) and transforms a $k$-sphere $S$ into a $k$-plane if $O \in S$ (center) or a $k$-sphere if $O \notin S$ (right).

with respect to the hyperplane orthogonal to $x$. We may suppose $x = (x_1, 0, \ldots, 0)$ and calculate the partial derivatives:

$$\varphi(x_1, \ldots, x_n) = r^2 \frac{(x_1, \ldots, x_n)}{\|x\|^2},$$
$$\frac{\partial \varphi_i}{\partial x_j} = r^2 \frac{\delta_{ij}\|x\|^2 - 2x_i x_j}{\|x\|^4}.$$

By calculating the partial derivatives at $x = (x_1, 0, \ldots, 0)$ we get

$$\frac{\partial \varphi_1}{\partial x_1} = -\frac{r^2}{x_1^2}, \quad \frac{\partial \varphi_i}{\partial x_i} = \frac{r^2}{x_1^2}, \quad \frac{\partial \varphi_j}{\partial x_k} = 0$$

for all $i > 1$ and $j \neq k$. The fact that an inversion preserves sphere and planes may be easily reduced to the bidimensional case (with circles and lines), a classical fact of euclidean geometry. $\square$

The half-space model $H^n$ is obtained from the disc model $D^n$ by an inversion in $\mathbb{R}^n$ of center $(0, \ldots, 0, -1)$ and radius $\sqrt{2}$ as shown in Fig. 6. The boundary $\partial H^n$ is the horizontal hyperplane $\{x_n = 0\}$, to which we add an point $\infty$ at infinity, so to have a bijective correspondence between $\partial H^n$ and $\partial D^n$ through the inversion.

PROPOSITION 1.15. *The half-space $H^n$ is a conformal model for $\mathbb{H}^n$. The $k$-subspaces in $H^n$ are the $k$-planes e the $k$-spheres in $\mathbb{R}^n$ orthogonal to $\partial H^n$.*

PROOF. The inversion is anticonformal and hence preserve angles, in particular it transforms $k$-sfere and $k$-planes in $D^n$ orthogonal to $\partial D^n$ into $k$-spheres and $k$-planes in $H^n$ orthogonal to $\partial H^n$. $\square$

Some lines and planes in $H^3$ are drawn in Fig. 7. The metric tensor $g$ on $H^n$ has a particolarly simple form.

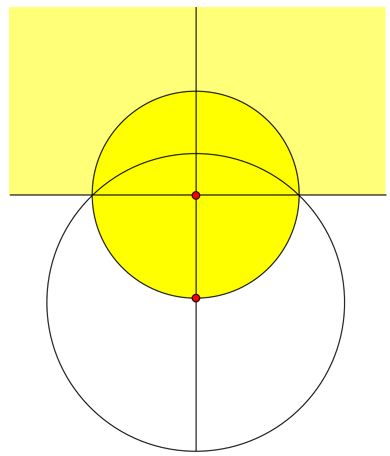


FIGURE 6. L'inversione lungo la sfera di centro $(0, \ldots, 0, -1)$ e raggio $\sqrt{2}$ trasforma il disco di Poincaré nel semispazio.

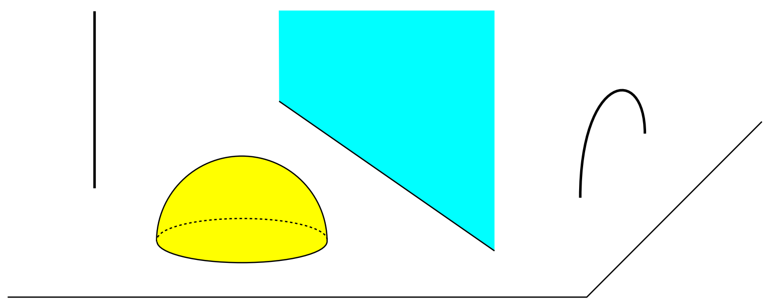


FIGURE 7. Rette e piani in $\mathbb{H}^3$ visualizzate con il modello del semispazio.

PROPOSITION 1.16. *The metric tensor on $H^n$ is:*

$$g_x = \frac{1}{x_n^2} \cdot g^E$$

*where $g^E$ is the euclidean metric tensor on $H^n \subset \mathbb{R}^n$.*

PROOF. The inversion $\varphi\colon D^n \to H^n$ is the function

$$\begin{aligned}\varphi(x_1, \ldots, x_n) &= (0, \ldots, 0, -1) + 2\frac{(x_1, \ldots, x_{n-1}, x_n + 1)}{\|(x_1, \ldots, x_{n-1}, x_n + 1)\|^2} \\ &= \frac{(2x_1, \ldots, 2x_{n-1}, 1 - \|x\|^2)}{\|x\|^2 + 2x_n + 1}.\end{aligned}$$

As seen in the proof of Proposition 1.14, the inversion $\varphi$ is anticonformal with scalar

$$\frac{2}{\|(x_1, \ldots, x_{n-1}, x_n + 1)\|^2} = \frac{2}{\|x\|^2 + 2x_n + 1}.$$

The map $\varphi$ hence trasforms the metric tensor $\left(\frac{2}{1-\|x\|^2}\right)^2 \cdot g^E$ in $x \in D^n$ into the metric tensor in $\varphi(x) \in H^n$ given by:

$$\left(\frac{2}{1-\|x\|^2}\right)^2 \cdot \left(\frac{\|x\|^2+2x_n+1}{2}\right)^2 \cdot g^E$$

which coincides with

$$\frac{1}{\varphi_n(x)^2} \cdot g^E.$$

□

In the half-space $H^n$ the lines are euclidean vertical half-lines or half-circles orthogonal to $\partial H^n$ as in Fig. 7. Verticali geodesics have a particularly simple form.

PROPOSITION 1.17. *A vertical geodesic in $H^n$ with unit speed is:*

$$\gamma(t) = (x_1, \ldots, x_{n-1}, e^t).$$

PROOF. We show that the speed of $\gamma$ is constantly one. A vector $v \in T_{(x_1,\ldots,x_n)}H^n$ has norm $\frac{\|v\|^E}{x_n}$ where $\|v\|^E$ indicates the euclidean norm. The velocity at time $t$ is $\gamma'(t) = (0, \ldots, 0, e^t)$ whose norm is

$$|\gamma'(t)| = \frac{e^t}{e^t} = 1.$$

□

We can easily deduce a parametrization for the geodesics in $D^n$ passing through the origin. Recall the hyperbolic tangent:

$$\tanh(t) = \frac{\sinh(t)}{\cosh(t)} = \frac{e^{2t}-1}{e^{2t}+1}.$$

PROPOSITION 1.18. *A geodesic in $D^n$ starting from the origin with velocity $x \in S^{n-1}$ is:*

$$\gamma(t) = \frac{e^t-1}{e^t+1} \cdot x = \left(\tanh \tfrac{t}{2}\right) \cdot x.$$

PROOF. We can suppose $x = (0, \ldots, 0, 1)$ and obtain this parametrization from that of the vertical line in $\mathbb{H}^n$ through inversion. □

We obtain in particular:

COROLLARY 1.19. *The exponential map* $\exp_0 \colon T_0D^n \to D^n$ *at the origin* $0 \in D^n$ *is the diffeomorphism:*

$$\exp_0(x) = \frac{e^{\|x\|}-1}{e^{\|x\|}+1} \cdot \frac{x}{\|x\|} = \left(\tanh \frac{\|x\|}{2}\right) \cdot \frac{x}{\|x\|}.$$

*The exponential maps are then all diffeomorphisms and* $\mathrm{inj}(\mathbb{H}^n) = \infty$.

In the half-space model it is easy to identify some isometries:

Proposition 1.20. *The horizontal translations $x \mapsto x + b$ with $b = (b_1, \ldots, b_{n-1}, 0)$ and the dilations $x \mapsto \lambda x$ with $\lambda > 0$ are isometries of $H^n$.*

Proof. Horizontal translations obiously preserve the tensor $g = \frac{1}{x_n^2} \cdot g^E$. A dilation $x \mapsto \lambda x$ sends a vector with eucldean norm 1 in $T_x H^n$ to a vector with euclidean norm $\lambda$ in $T_{\lambda x} H^n$. To get the hyperbolic norms we must divide them respectively by $x_n$ and $\lambda x_n$ and thus the dilation is norm-preserving and hence an isometry. □

Corollary 1.21. *In the conformal models (disc and half-space) every isometry of $\mathbb{H}^n$ sends $k$-spheres and euclidean $k$-planes to $k$-spheres and euclidean $k$-planes.*

Proof. With the half-space model, for every pair of points $p, q$ we may construct an isometry $\varphi$ with $\varphi(p) = q$ which is a composition of a dilation and a horizontal translation. Since dilations and translations satisfy the thesis, we may restrict to isometries that fix a point $P$. We now use the disc model (the two models are connected by an inversion, which also satisfies the thesis) and verify it for $P = 0$, whose stabilizer is just $O(n)$. □

Since $\operatorname{inj}(\mathbb{H}^n) = +\infty$, the ball $B(p, r) \subset \mathbb{H}^n$ centered at a point $p \in \mathbb{H}^n$ with radius $r$ is diffeomorphic to a euclidean ball. In the conformal models, it *is* actually a euclidean ball.

Corollary 1.22. *In the conformal models (disc and half-space) the balls are euclidean balls (with a different center!).*

Proof. In the disc model $B(0, r)$ is the euclidean ball of radius $\ln \frac{1+r}{1-r}$. The isometries of $\mathbb{H}^2$ act transitively on points and send $(n-1)$-spheres to $(n-1)$-spheres, whence the thesis. □

## 2. Compactification and isometries of hyperbolic space

**2.1. Points at infinity.** In this section we compactify the hyperbolic space $\mathbb{H}^n$ by adding its "points at infinity".

Let a *geodesic half-line* in $\mathbb{H}^n$ be a geodesic $\gamma \colon [0, +\infty) \to \mathbb{H}^n$ with constant unit speed.

Definition 2.1. The set $\partial\mathbb{H}^n$ of the *points at infinity* in $\mathbb{H}^n$ is the set of al geodesic half-lines, seen up to the following equivalence relation:

$$\gamma_1 \sim \gamma_2 \iff \sup\big\{\gamma_1(t), \gamma_2(t)\big\} < +\infty.$$

We can add to $\mathbb{H}^n$ its points at infinity and define

$$\overline{\mathbb{H}^n} = \mathbb{H}^n \cup \partial\mathbb{H}^n.$$

Proposition 2.2. *On the disc model there is a natural 1-1 correspondence between $\partial D^n$ and $\partial\mathbb{H}^n$ and hence between the closed disc $\overline{D^n}$ and $\overline{\mathbb{H}^n}$.*

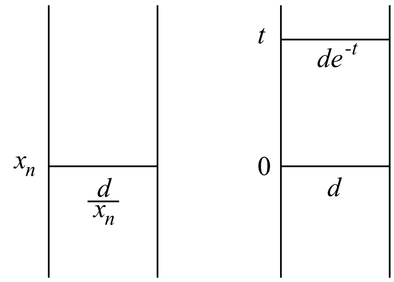



FIGURE 8. Two vertical lines in the half-space model $H^n$ at *euclidean* distance $d$. The *hyperbolic* length of the horizontal segment between them at height $x_n$ is $\frac{d}{x_n}$ and hence tends to zero as $x_n \to \infty$ (left). Using as a height parameter the more intrinsic hyperbolic arc-length, we see that the two vertical geodesics $\gamma_1$ e $\gamma_2$ approach at exponential rate, since $d(\gamma_1(t), \gamma_2(t)) \leqslant de^{-t}$ (right).

PROOF. A geodesic half-line $\gamma$ in $D^n$ is a circle or line arc orthogonal to $\partial D^n$ and hence the euclidean limit $\lim_{t\to\infty} \gamma(t)$ is a point in $\partial D^n$. We now prove that two half-lines tend to the same point if and only if they lie in the same equivalence class.

Suppose two half-geodesics $\gamma_1$, $\gamma_2$ tend to the same point in $\partial D^n$. We can use the half-space model and put this point at $\infty$, hence $\gamma_1$ and $\gamma_2$ are vertical and point upwards:

$$\gamma_1(t) = (x_1, \ldots, x_{n-1}, x_n e^t), \quad \gamma_2(t) = (y_1, \ldots, y_{n-1}, y_n e^t).$$

The geodesic

$$\gamma_3(t) = (y_1, \ldots, y_{n-1}, x_n e^t)$$

is equivalent to $\gamma_2$ since $d(\gamma_1(t), \gamma_3(t)) = |\ln \frac{y_n}{x_n}|$ for all $t$ and is also equivalent to $\gamma_1$ because $d(\gamma_1(t), \gamma_3(t)) \to 0$ as shown in Fig. 8.

Suppose $\gamma_1$ e $\gamma_2$ tend to distinct points in $\partial\mathbb{H}^n$. We can use the half-space model again and suppose that $\gamma_1$ is upwards vertical and $\gamma_2$ tends to some other point in $\{x_n = 0\}$. In that case we easily see that $d(\gamma_1(t), \gamma_2(t)) \to \infty$: for any $M > 0$ there is a $t_0 > 0$ such that $\gamma_1(t)$ and $\gamma_2(t)$ lie respectively in $\{x_{n+1} > M\}$ and $\left\{x_n < \frac{1}{M}\right\}$ for all $t > t_0$. Whatever curve connecting these two open sets has length at least $\ln M^2$, hence $d(\gamma_1(t), \gamma_2(t)) > \ln M^2$ for all $t > t_0$. $\square$

We can give $\overline{\mathbb{H}^n}$ the topology of $\overline{D^n}$: in that way we have *compactified* the hyperbolic space by adding its points at infinity. The interior of $\overline{\mathbb{H}^n}$ is $\mathbb{H}^n$, and the points at infinity form a sphere $\partial\mathbb{H}^n$.

Note that although $\mathbb{H}^n$ is a complete riemannian metric (and hence a metric space), its compactification $\overline{\mathbb{H}^n}$ is only a topological space: a point in $\partial\mathbb{H}^n$ has infinite distance from any other point in $\overline{\mathbb{H}^n}$.

The topology on $\overline{\mathbb{H}^n}$ may be defined intrinsically without using a particular model $D^n$: for any $p \in \partial\mathbb{H}^n$ we define a system of open neighborhoods

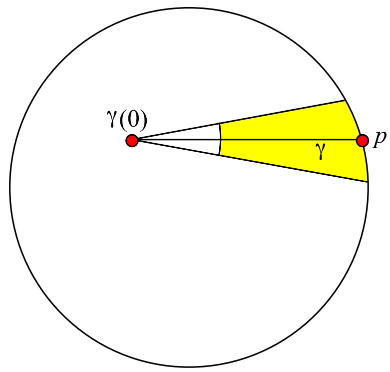



FIGURE 9. An open neighborhood $U(\gamma, V, r)$ of $p \in \partial\mathbb{H}^n$ (in yellow).

of $p$ in $\overline{\mathbb{H}^n}$ as follows. Let $\gamma$ be a geodesic that represents $p$ and $V$ be an open neighborhood of the vector $\gamma'(0)$ in the unitary sphere in $T_{\gamma(0)}$. Pick $r > 0$. We define the following subset of $\overline{\mathbb{H}^n}$:

$$U(\gamma, V, r) = \big\{\alpha(t) \bigm| \alpha(0) = \gamma(0),\ \alpha'(0) \in V,\ t > r\big\}$$
$$\bigcup \big\{[\alpha] \bigm| \alpha(0) = \gamma(0),\ \alpha'(0) \in V\big\}$$

where $\alpha$ indicates a half-line in $\mathbb{H}^n$ and $[\alpha] \in \partial\mathbb{H}^n$ its class, see Fig. 9. We define an open neighborhoods system $\{U(\gamma, V, r)\}$ for $p$ by letting $\gamma$, $V$, and $r$ vary. The resulting topology on $\overline{\mathbb{H}^n}$ coincides with that induced by $\overline{D^n}$.

PROPOSITION 2.3. *Two distinct points in $\partial\mathbb{H}^2$ are the endpoints of a unique line.*

PROOF. Take $H^n$ with one point at $\infty$ an the other lying in $\{x_{n+1} = 0\}$. There is only one euclidean vertical line connecting them. $\square$

**2.2. Elliptic, parabolic, and hyperbolic isometries.** Every isometry of $\mathbb{H}^n$ extends to the boundary.

PROPOSITION 2.4. *Every isometry $\varphi\colon \mathbb{H}^n \to \mathbb{H}^n$ estends to a unique homeomorphism $\varphi : \overline{\mathbb{H}^n} \to \overline{\mathbb{H}^n}$. An isometry $\varphi$ is determined by its trace $\varphi|_{\partial\mathbb{H}^n}$ at the boundary.*

PROOF. The extension of $\varphi$ to $\partial\mathbb{H}^n$ is defined in a natural way: a boundary point is a class $[\gamma]$ of geodesic half-lines and we set $\varphi([\gamma]) = [\varphi(\gamma)]$. Since the topology on $\overline{\mathbb{H}^n}$ may be defined intrinsically, the extension is a homeomorphism.

To prove the second assertion we show that an isometry $\varphi$ that fixes the points at infinity is the identity. The isometry $\varphi$ fixes every line as a set (because it fixes its endpoints), and since every point is the intersection of two lines it fixes also every point. $\square$

PROPOSITION 2.5. *Let $\varphi$ be an isometry of $\mathbb{H}^n$. Precisely one of the following holds:*

(1) *$\varphi$ has at least one fixed point in $\mathbb{H}^n$,*
(2) *$\varphi$ has no fixed points in $\mathbb{H}^n$ and has exactly one in $\partial\mathbb{H}^n$,*
(3) *$\varphi$ has no fixed points in $\mathbb{H}^n$ and has exactly two in $\partial\mathbb{H}^n$.*

PROOF. The extension $\varphi : \overline{\mathbb{H}^n} \to \overline{\mathbb{H}^n}$ is continuous and has a fixed point by Brouwer theorem. We only need to prove that if $\varphi$ has three fixed points $P_1, P_2, P_3$ at the boundary then it has some fixed point in the interior. The isometry $\varphi$ fixes the line $r$ with endpoints $P_1$ and $P_2$. There is only line $s$ with endpoint $P_3$ and orthogonal to $r$: the isometry $\varphi$ must also fix $s$ and hence fixes the point $r \cap s$. □

Figura?

Isometries of type (1), (2), and (3) are called respectively *elliptic*, *parabolic*, and *hyperbolic*. A hyperbolic isometry fixes two points $p, q \in \partial\mathbb{H}^n$ and hence preserves the unique line $l$ with endpoints $p$ and $q$. The line $l$ is the *axis* of the hyperbolic isometry, which acts on $l$ as a translation.

**2.3. Incident, parallel, and ultraparallel subspaces.** In the compactification, every $k$-subspace $S \subset \mathbb{H}^n$ has a topological closure $\overline{S} \subset \overline{\mathbb{H}^n}$. In the two conformal models, the boundary $\partial S = \overline{S} \cap \partial\mathbb{H}^n$ is a $(k-1)$-sphere (or a $(k-1)$-plane plus $\infty$ in $H^n$).

The usual *distance* $d(A, B)$ between two subsets $A, B$ in a metric space is defined as

$$d(A, B) = \inf_{x\in A, y\in B} \big\{d(x, y)\big\}.$$

PROPOSITION 2.6. *Let $S$ and $S'$ be subspaces of $\mathbb{H}^n$ arbitrary dimension. Precisely one of the following holds:*

(1) *$S \cap S' \neq \emptyset$,*
(2) *$S \cap S' = \emptyset$ and $\overline{S} \cap \overline{S'}$ is a point in $\partial\mathbb{H}^n$; moreover $d(S, S') = 0$ and there is no geodesic orthogonal to both $S$ and $S'$,*
(3) *$\overline{S} \cap \overline{S'} = \emptyset$; moreover $d = d(S, S') > 0$ and there is a unique geodesic $\gamma$ orthogonal to both $S$ and $S'$: the segment of $\gamma$ between $S$ and $S'$ is the only arc connecting them having length $d$.*

PROOF. If $\partial S \cap \partial S'$ contains two points then it contains the line connecting them and hence $S \cap S' \neq \emptyset$.

In (2) we use the half-space model and send $\overline{S} \cap \overline{S'}$ at infinity. Then $S$ and $S'$ are euclidean vertical subspaces and Fig. 8 shows that $d(S, S') = 0$. Lines are vertical or half-circles and cannot be orthogonal to both $S$ and $S'$.

In (3), let $x_i \in S$ e $x'_i \in S'$ be such that $d(x_i, x'_i) \to d$. Since $\overline{\mathbb{H}^n}$ is compact, on a subsequence $x_i \to x \in \overline{S}$ and $x'_i \to x' \in \overline{S'}$. By hypothesis $x \neq x'$ and hence $d > 0$ and $x, x' \in \mathbb{H}^n$ since $0 < d < \infty$.

Let $\gamma$ be the line passing through $x$ and $x'$. The segment between $x$ and $x'$ has length $d(x, x') = d$. The line is orthogonal to $S$ and $S'$: if it had with $S'$ an angle smaller than $\frac{\pi}{2}$ we could find another point $x'' \in S'$ near $x'$ with $d(x, x'') < d$. We can draw $S, S', \gamma$ as in Fig. 10 by placing the origin between $x$ and $x'$: no other line can be orthogonal to both $S$ and $S'$. □

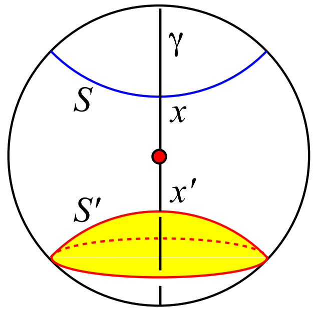



FIGURE 10. Two disjoint subspaces $S, S'$ and a line $\gamma$ orthogonal to both.

Two subspaces of type (1), (2) or (3) are called respectively *incident*, *asymptotically parallel*, and *ultra-parallel*.

**2.4. Möbius transformations.** If $M$ is an orientable riemannian manifold. è una varietà riemanniana orientabile, we denote be $\mathrm{Isom}^+(M)$ the group of all orientation-preserving isometries of $M$. We describe $\mathrm{Isom}^+(\mathbb{H}^2)$ and $\mathrm{Isom}^+(\mathbb{H}^3)$ conveniently as some groups of $2 \times 2$ matrices.

Recall that the *Riemann sphere* $S = \mathbb{C} \cup \{\infty\}$ is a fundamental object in complex analysis and projective geometry. In complex analysis an *automorphism* of the Riemann sphere is a *biolomorphism* of $S$, and one proves that the automorphisms are precisely the *Möbius transformations*

$$z \mapsto \frac{az+b}{cz+d}$$

where $a, b, c, d$ are complex numbers with $ad-bc \neq 0$. A Möbius transformation is hence determined by an invertible matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ and two transformations compose as matrices do. Moreover two matrices $A$ and $B$ determine the same transformation if and only if $B = \lambda A$ for some $\lambda \in \mathbb{C}^*$. The group of all Möbius transformations is hence isomorphic to

$$\mathbb{P}\mathrm{GL}_2(\mathbb{C}) = \mathbb{P}\mathrm{SL}_2(\mathbb{C}) = \mathrm{GL}_2(\mathbb{C})/_{\{\lambda I\}} = \mathrm{SL}_2(\mathbb{C})/_{\pm I}.$$

The symbol $\mathbb{P}$ before a set of matrices (or other objects) indicate that we quotient the set by the relation $A \sim \lambda A$ for any $\lambda \neq 0$.

In projective geometry the Riemann sphere is the complex projective line $\mathbb{CP}^1 = \mathbb{C} \cup \{\infty\}$ and its automorphisms are the *projective transformations*

$$[z, w] \mapsto [az + bw, cz + dw]$$

with (as above) $ad - bc \neq 0$. Also in this context the automorphisms group is $\mathbb{P}\mathrm{SL}_2(\mathbb{C})$. We recall that a projective transformation of $\mathbb{CP}^n$ is determined by its behaviour on $n + 2$ points in general position: when $n = 1$ we get

PROPOSITION 2.7. *Given two triples $\{P_1, P_2, P_3\}$ and $\{Q_1, Q_2, Q_3\}$ of distinct points in $\mathbb{CP}^1$ there is a unique Möbius transformation $\varphi$ such that $\varphi(P_i) = Q_i$ for all $i$.*

The Möbius transformations with $a, b, c, d \in \mathbb{R}$ and $ad - bc > 0$ form the subgroup of $\mathbb{PSL}_2(\mathbb{C})$ denoted as

$$\mathbb{PSL}_2(\mathbb{R}) = \mathrm{SL}_2(\mathbb{R})/_{\pm I}.$$

These are the Möbius transformations that preserve the extended real line $\mathbb{R} \cup \{\infty\}$ and do not interchange the two half-planes in $(\mathbb{C} \cup \infty) \setminus (\mathbb{R} \cup \infty)$.

Exercise 2.8. Given two triples $\{P_1, P_2, P_3\}$ e $\{Q_1, Q_2, Q_3\}$ of distinct points in $\mathbb{R} \cup \{\infty\}$ with the same cyclic orientation there is a unique Möbius transformation $\varphi \in \mathbb{PSL}_2(\mathbb{R})$ such that $\varphi(P_i) = Q_i$ for all $i$.

**2.5. Isometries of $H^2$.** We identify $\mathbb{R}^2$ with $\mathbb{C}$ and consider the conformal models

$$D^2 = \{z \in \mathbb{C} \mid |z| < 1\}, \quad H^2 = \{z \in \mathbb{C} \mid \Im z > 0\}.$$

The inversion relating them is the Möbius anti-transformation

$$z \mapsto \frac{\bar{z} + i}{i\bar{z} + 1}. \tag{1}$$

Proposition 2.9. *We have* $\mathrm{Isom}^+(H^2) = \mathbb{PSL}_2(\mathbb{R})$.

Proof. The group $\mathbb{PSL}_2(\mathbb{R})$ is generated by:

(1) translations $z \mapsto z + b$ with $b \in \mathbb{R}$, corresponding to matrices $\left(\begin{smallmatrix} 1 & b \\ 0 & 1 \end{smallmatrix}\right)$,
(2) dilations $z \mapsto \lambda z$ with $\lambda \in \mathbb{R}^*$, corresponding to matrices $\begin{pmatrix} \sqrt{\lambda} & 0 \\ 0 & \frac{1}{\sqrt{\lambda}} \end{pmatrix}$,
(3) inversion $z \mapsto -\frac{1}{z}$, corresponding to $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$.

To prove this, check that with these transformations we can send $0, 1, \infty$ to any co-oriented triple of points in $\mathbb{R} \cup \{\infty\}$ and use Exercise 2.8. Each such transformation is an isometry: use Proposition 1.20 and check that the inversion transforms via the Möbius anti-transformation (1) into a reflection along the origin in $D^2$. Hence $\mathbb{PSL}_2(\mathbb{R}) \subset \mathrm{Isom}^+(H^2)$.

Conversely, take $\varphi \in \mathrm{Isom}^+(H^2)$. Since $\varphi$ is orientatio-preserving, it sends $0, 1, \infty$ into a co-oriented triple of points in $\mathbb{R} \cup \{\infty\}$. Let $\psi \in \mathbb{PSL}_2(\mathbb{R})$ act on $0, 1, \infty$ like $\varphi$. Then $f = \psi^{-1} \circ \varphi$ is an isometry that fixes $0, 1, \infty$ and we easily conclude that $f = \mathrm{id}$ and hence $\varphi = \psi$. To prove that, note that $f$ has a fixed point $x \in \mathbb{H}^2$ by Proposition 2.5. The isometry $f$ fixes $x$ and the three half-lines starting from $x$ and pointing towards $0, 1, \infty$: therefore $df_x = \mathrm{id}$ and hence $f = \mathrm{id}$. $\square$

The isometry type is determined by the trace of $A \in \mathbb{PSL}_2(\mathbb{R})$, which is well-defined up to sign:

Proposition 2.10. *A non-trivial isometry $A \in \mathbb{PSL}_2(\mathbb{R})$ is elliptic, parabolic, hyperbolic if and only if respectively* $|\mathrm{tr} A| < 2$, $|\mathrm{tr} A| = 2$, $|\mathrm{tr} A| > 2$.

Proof. Take $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ with $\det A = ad - bc = 1$. The Möbius transformation $z \mapsto \frac{az+b}{cz+d}$ has a fixed point $z \in \mathbb{C}$ if and only if

$$\frac{az + b}{cz + d} = z \Longleftrightarrow cz^2 + (d - a)z - b = 0$$

We find

$$\Delta = (d-a)^2 + 4bc = (d+a)^2 - 4 = \mathrm{tr}^2 A - 4.$$

There is a fixed point in $\mathbb{H}^2$ if and only if $\Delta < 0$; if $\Delta > 0$ we find two fixed points in $\mathbb{R} \cup \{\infty\}$ and if $\Delta = 0$ only one. □

**2.6. Isometries of $H^3$.** We identify

$$\mathbb{R}^3 = \mathbb{C} \times \mathbb{R} = \{(z,t) \mid z \in \mathbb{C}, t \in \mathbb{R}\}$$

hence $H^3 = \{(z,t) \mid \Im z > 0\}$. We also identify $\mathbb{C} \times \{0\}$ with $\mathbb{C}$. Every isometry of $H^3$ extends to the boundary

$$\partial H^3 = \mathbb{C} \cup \{\infty\}.$$

PROPOSITION 2.11. *The boundary trace induces an identification*

$$\mathrm{Isom}^+(H^3) = \mathbb{P}\mathrm{SL}_2(\mathbb{C}).$$

PROOF. The proof is similar to that of Proposition 2.9. We prove analogously that $\mathbb{P}\mathrm{SL}_2(\mathbb{C})$ is generated by:

(1) translations $z \mapsto z + b$ con $b \in \mathbb{C}$, corresponding to matrices $\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$,
(2) dilations $z \mapsto \lambda z$ with $\lambda \in \mathbb{C}^*$, corresponding to matrices $\begin{pmatrix} \sqrt{\lambda} & 0 \\ 0 & \frac{1}{\sqrt{\lambda}} \end{pmatrix}$,
(3) inversion $z \mapsto -\frac{1}{z}$ corresponding to $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

Each such transformation is the trace of an isometry of $\mathbb{H}^3$:

(1) the horizontal translations $(z,t) \mapsto (z+b,t)$,
(2) if $\lambda = \rho e^{i\theta}$, the composition of a dilation $(z,t) \mapsto \rho(z,t)$ and a rotation $(z,t) \mapsto (e^{i\theta}z, t)$,
(3) inversion with respect to the sphere $|z|^2 + t^2 = 1$.

Mettere prima che le inversioni sono isometrie?

Conversely, we prove as above that the trace of an isometry $\varphi$ lies in $\mathbb{P}\mathrm{SL}_2(\mathbb{C})$, because an isometry that fixes 3 distinct points $P, Q, R \in \partial\mathbb{H}^3$ fixes pointwise the plane $\pi$ containing them and is hence either the identity or a reflection along $\pi$, but only the first one is orientation-reversing. □

Also in $H^3$ the isometry type is determined by the trace of $A \in \mathbb{P}\mathrm{SL}_2(\mathbb{C})$, well-defined up to sign:

PROPOSITION 2.12. *A non-trivial isometry $A \in \mathbb{P}\mathrm{SL}_2(\mathbb{C})$ is elliptic, parabolic, hyperbolic if and only if respectively* $\mathrm{tr}A \in (-2,2)$, $\mathrm{tr}A = \pm 2$, $\mathrm{tr}A \in \mathbb{C} \setminus (-2,2)$.

PROOF. A non-trivial matrix $A \in \mathrm{SL}_2(\mathbb{C})$ is conjugate to one of:

$$\pm\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} \lambda & 0 \\ 0 & \frac{1}{\lambda} \end{pmatrix}$$

for some $\lambda \in \mathbb{C}^*$ and represents an isometry:

$$(z,t) \longmapsto (z+1,t), \quad (z,t) \longmapsto (\lambda^2 z, |\lambda|^2 t).$$

In the first case $\mathrm{tr}A = \pm 2$ and $A$ is parabolic with fixed point $\infty$, in the second case $A$ has a fixed point in $H^3$ if and only if $|\lambda| = 1$, *i.e.* $\mathrm{tr}A =$

$\lambda + \lambda^{-1} \in (-2, 2)$, the fixed point being $(0, 1)$. If $|\lambda| \neq 1$ there are two fixed points 0 and $\infty$ at infinity and hence $A$ is hyperbolic. $\square$

Finally, we note that the orientation-reversing isometries of $H^2$ and $H^3$ may be described analogously as Möbius anti-transformations

$$z \mapsto \frac{a\bar{z} + b}{c\bar{z} + d}.$$

For $H^2$ we only consider anti-trasformations with real coefficients and negative determinant $ad - bc < 0$. We may identify the groups

$$\mathrm{Isom}(H^2) = \mathbb{P}\mathrm{GL}_2(\mathbb{R})$$

using the convention that matrices with negative determinant act like Möbius anti-transformations.

**2.7. Horospheres.** Parabolic transformations are related to some objects in $\mathbb{H}^n$ called *horospheres*.

DEFINITION 2.13. Let $p$ be a point in $\partial\mathbb{H}^n$. A *horosphere* centered in $p$ is a connected hypersurface orthogonal to all the lines exiting from $p$.

Horospheres may be easily visualized in the half-space model $H^n$ by sending $p$ at infinity. The lines exiting from $p$ are the euclidean vertical lines and the horospheres centered at $p$ are precisely the horizontal hyperplanes $\{x_n = k\}$ with $k > 0$.

REMARK 2.14. Since the metric tensor $g = \frac{1}{x_n^2} g^E$ is constant on each hyperplane $\{x_n = k\}$, each horosphere is *isometric* to the euclidean $\mathbb{R}^n$. This quite surprising fact is peculiar to hyperbolic geometry: hyperbolic space $\mathbb{H}^n$ contains hypersurfaces isometric to $\mathbb{H}^{n-1}$ (the hyperplanes), $\mathbb{R}^{n-1}$ (the horospheres), and $S^{n-1}$ (the spheres).

The horospheres centered at $p \neq \infty$ in $\partial H^n$ or in any point $p \in \partial D^n$ are precisely the euclidean spheres tangent in $p$ to the sphere at infinity. The horospheres in $\mathbb{H}^2$ are circles and are called *horocycles*, see Fig. 11.

Let us go back to the isometries of $\mathbb{H}^n$. In the half-space model $H^n$ we denote a point as a pair $(x, t)$ with $x = (x_1, \ldots, x_{n-1})$ and $t = x_n$.

PROPOSITION 2.15. *Let $\varphi$ be an isometry of $\mathbb{H}^n$:*

(1) *if $\varphi$ is elliptic with fixed point $0 \in D^n$ then*

$$\varphi(x) = Ax$$

*for some matrix $A \in O(n)$;*

(2) *if $\varphi$ is parabolic with fixed point $\infty$ in $H^n$ then*

$$\varphi(x, t) = (Ax + b, t)$$

*for some matrix $A \in O(n)$ and some vector $b$;*

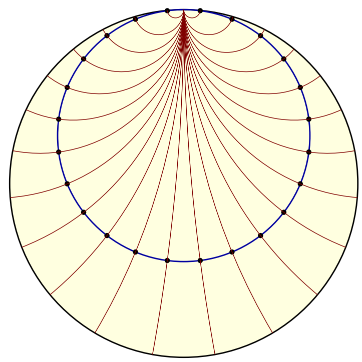


FIGURE 11. A horocycle in $\mathbb{H}^2$ centered in $p \in \partial\mathbb{H}^2$ is a circle tangent to $p$. It is orthogonal to all the lines exiting from $p$.

(3) *if* $\varphi$ *is hyperbolic with fixed points* $0$ *and* $\infty$ *then*

$$\varphi(x,t) = \lambda(Ax,t)$$

*for some matrix* $A \in O(n)$ *and some positive scalar* $\lambda \neq 1$.

PROOF. Point (1) is obvious. In (2) the isometry $\varphi$ fixes $\infty$ and hence permutes the horospheres centered at $\infty$: we first prove that this permutation is trivial. The map $\varphi$ sends a horosphere $O_0$ at height $t = t_0$ to a horosphere $O_1$ at some height $t = t_1$. If $t_1 \neq t_0$, up to change $\varphi$ with its inverse we may suppose that $t_1 < t_0$. We know that the map $\psi\colon O_1 \to O_0$ sending $(x,t_1)$ to $(x,t_0)$ is a contraction: hence $\varphi \circ \psi\colon O_1 \to O_1$ is a contraction and thus has a fixed point $(x,t_1)$. Therefore $\varphi(x,t_0) = (x,t_1)$. Since $\varphi(\infty) = \infty$, the vertical geodesic passing through $(x,t_0)$ and $(x,t_1)$ is preserved by $\varphi$, and hence we have found another fixed point $(x,0) \in \partial\mathbb{H}^n$: a contradiction.

We now know that $\varphi$ preserves each horosphere $O$ centered at $\infty$. On one geodesic it acts like an isometry $x \mapsto Ax + b$ of euclidean space. Since $\varphi$ sends vertical geodesics to vertical geodesics, it acts with the same formula on each horosphere and we are done.

Concerning (3), the axis $l$ of $\varphi$ is the vertical line with endpoints $0$ and $\infty$, and $\varphi$ acts on $l$ by translation; hence it sends $(0,1)$ to some $(0,\lambda)$. The differential $d\varphi$ at $(0,1)$ is necessarily $\left(\begin{smallmatrix} A & 0 \\ 0 & \lambda \end{smallmatrix}\right)$ for some $A \in O(n)$ and hence $\varphi$ is globally as stated. The case $\lambda = 1$ is excluded because $(0,1)$ would be a fixed point in $\mathbb{H}^n$. $\square$

The *minimum displacement* $d(\varphi)$ of an isometry $\varphi$ of $\mathbb{H}^n$ is

$$d(\varphi) = \inf_{x \in \mathbb{H}^n} d\big(x, \varphi(x)\big).$$

A point $x$ realizes the minimum displacement if $d(x,\varphi(x)) = d(\varphi)$.

COROLLARY 2.16. *The following holds:*

(1) *an elliptic transformation has* $d = 0$ *realized on its fixed points,*
(2) *a parabolic transformation with fixed point* $p \in \partial\mathbb{H}^n$ *has* $d = 0$ *realized nowhere and fixes every horosphere centered in* $p$*;*
(3) *a hyperbolic transformation with fixed points* $p, q \in \partial\mathbb{H}^n$ *has* $d > 0$ *realized on its axis.*

PROOF. Point (1) is obvious. Point (2) was already noticed while proving Proposition 2.15. Concerning (3), let $l$ be the axis of the hyperbolic transformation $\varphi$. The hyperplane orthogonal to $l$ in a point $x \in l$ is sent to the hyperplane orthogonal to $l$ in $\varphi(x)$. The two hyerplanes are ultraparallel and by Proposition 2.6 their minimum distance is at the points $x$ and $\varphi(x)$. Hence the points on $l$ realize the minimum displacement for $\varphi$. □

COROLLARY 2.17. *The iterate* $\varphi^k$ *of a elliptic, parabolic, hyperbolic transformation is again elliptic, parabolic, hyperbolic.*

PROOF. Vero per paraboliche? □

## 3. Geometry of hyperbolic space

We investigate the geometry of $\mathbb{H}^n$.

**3.1. Area and curvature.** We can verify that $\mathbb{H}^n$ has constant sectional curvature $-1$. It should be no surprise that $\mathbb{H}^n$ has constant curvature, since it has many symmetries (*i.e.* isometries). To calculate its sectional curvature we calculate the area of a disc.

PROPOSITION 3.1. *The disc of radius* $r$ *in* $\mathbb{H}^2$ *has area*

$$A(r) = \pi\left(e^{\frac{r}{2}} - e^{-\frac{r}{2}}\right)^2 = 4\pi\sinh^2\tfrac{r}{2} = 2\pi(\cosh r - 1).$$

PROOF. In general, let $U \subset \mathbb{R}^n$ be an open set with a metric tensor $g$, expressed as a square matrix $g_x$ depending smoothly on $x \in U$. The induced volume form on $U$ is

$$\sqrt{\det g}\cdot dx_1\cdots dx_n.$$

Let now $D(r)$ be a disc in $\mathbb{H}^2$ of radius $r$. If we center it in 0 in the disc model, its euclidean radius is $\tanh\frac{r}{2}$ by Corollary 1.19 and we get

$$\begin{aligned} A(r) &= \int_{D(r)}\sqrt{\det g}\cdot dxdy = \int_{D(r)}\left(\frac{2}{1-x^2-y^2}\right)^2 dxdy \\ &= \int_0^{2\pi}\int_0^{\tanh\frac{r}{2}}\left(\frac{2}{1-\rho^2}\right)^2\rho\cdot d\rho d\theta = 2\pi\left[\frac{2}{1-\rho^2}\right]_0^{\tanh\frac{r}{2}} \\ &= 4\pi\left(\frac{1}{1-\tanh^2\frac{r}{2}} - 1\right) = 4\pi\sinh^2\tfrac{r}{2}. \end{aligned}$$

□

COROLLARY 3.2. *Hyperbolic space* $\mathbb{H}^n$ *has sectional curvature* $-1$.

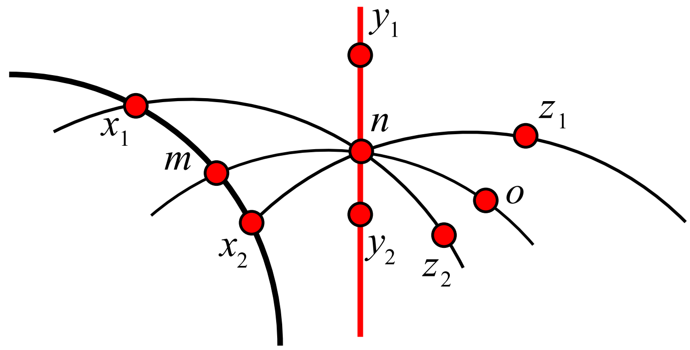



FIGURE 12. Distance between points in disjoint lines is a strictly convex function in hyperbolic space.

PROOF. Pick $p \in \mathbb{H}^n$ and $W \subset T_p$ a 2-dimensional subspace. The image $\exp_p(W)$ is the hyperbolic plane tangent to $W$ in $p$. On a hyperbolic plane

$$A(r) = 2\pi(\cosh r - 1) = 2\pi\left(\frac{r^2}{2!} + \frac{r^4}{4!} + o(r^4)\right) = \pi r^2 + \frac{\pi r^4}{12} + o(r^4)$$

and hence $K = -1$ following Definition 2.14. $\square$

**3.2. Convexity of the distance function.** We recall that a function $f\colon \mathbb{R}^n \to \mathbb{R}$ is *strictly convex* if

$$f(tv + (1-t)w) < tf(v) + (1-t)f(w)$$

for any pair $v, w \in \mathbb{R}^k$ of distinct points and any $t \in (0, 1)$. The following is immediate.

EXERCISE 3.3. A positive strictly convex function is continuous and admits a minimum if and only if it is proper.

As opposite to euclidean space, in $\mathbb{H}^n$ the distance function is strictly convex on disjoint lines. The product of two line $l \times l'$ is identified to $\mathbb{R} \times \mathbb{R}$ via an isometry which is unique up to translations.

PROPOSITION 3.4. *Let $l, l' \subset \mathbb{H}^2$ be two disjoint lines. The map*

$$\begin{aligned} l \times l' &\longrightarrow \mathbb{R}_{\geqslant 0} \\ (x, y) &\longmapsto d(x, y) \end{aligned}$$

*is strictly convex; it is proper if and only if the lines are ultraparallel.*

PROOF. The function $d$ is clearly continuous, hence to prove its convexity it suffices to show that

$$d\left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2}\right) < d(x_1, y_1) + d(x_2, y_2)$$

for any pair of distinct points $(x_1, y_1), (x_2, y_2) \in l \times l'$. Suppose $x_1 \neq x_2$ and denote by $m$ and $n$ the midpoints $\frac{x_1+x_2}{2}$ and $\frac{y_1+y_2}{2}$ as in Fig. 12.

Let $\sigma_p$ be the reflection at the point $p \in \mathbb{H}^2$. The map $\tau = \sigma_n \circ \sigma_m$ translates the line $r$ containing the segment $mn$ by the quantity $2d(m,n)$: hence it is a hyperbolic transformation with axis $r$. We draw the points $o = \tau(m)$ and $z_i = \tau(x_i)$ in the figure and note that $z_1 = \sigma_n(x_2)$, hence $d(x_2, y_2) = d(z_1, y_1)$. The triangular inequality implies that

$$d(x_1, z_1) \leqslant d(x_1, y_1) + d(y_1, z_1) = d(x_1, y_1) + d(x_2, y_2).$$

A hyperbolic transformation has minimum displacement on its axis $r$ and $x_1 \neq m$ is not in $r$, hence

$$2d(m,n) = d(m,o) = d(m, \tau(m)) < d(x_1, \tau(x_1)) = d(x_1, z_1).$$

Finally we get $2d(m,n) < d(x_1, y_1) + d(x_2, y_2)$. The function $d$ is proper, that is it has minimum, if and only if the two lines are ultraparallel by Proposition 2.6. □

**3.3. Convex combinations.** Let $p_1, \ldots, p_k \in$ be $k$ points in $\mathbb{H}^n, \mathbb{R}^n$ or $S^n$ and $t_1, \ldots, t_k$ be non-negative numbers with $t_1 + \ldots + t_k = 1$. The *convex combination*

$$p = t_1 p_1 + \ldots + t_k p_k$$

is another point in space defined as follows:

$$\text{in } \mathbb{R}^n: \quad p = t_1 p_1 + \ldots + t_k p_k$$
$$\text{in } I^n, S^n: \quad p = \frac{t_1 p_1 + \ldots + t_k p_k}{\|t_1 p_1 + \ldots + t_k p_k\|}$$

where $|v| = \sqrt{-\langle v, v\rangle}$ in the $I^n$ case. Using convex combination we may define the *barycenter* of the points as $\frac{1}{2}p_1 + \ldots \frac{1}{2}p_k$.

**3.4. Unimodularity.** As a group of matrices $O_*(n,1)$ the isometry group $\mathrm{Isom}(\mathbb{H}^n)$ is naturally a Lie group. Recall that a Lie group is unimodular if it admits a Haar measure that is both left- and right-invariant.

COROLLARY 3.5. *The isometry group* $\mathrm{Isom}(\mathbb{H}^n)$ *is unimodular.*

PROOF. The group $\mathrm{Isom}(\mathbb{H}^n)$ contains a simple subgroup of index two, hence the modular function has finite - thus trivial - image. □

REMARK 3.6. A Haar measure for $\mathrm{Isom}(\mathbb{H}^n)$ may be constructed as follows. Let $x \in \mathbb{H}^n$ be a fixed point. Define the measure of a Borel set $S \subset \mathrm{Isom}(\mathbb{H}^n)$ as the measure of $S(x) = \cup_{f \in S} f(x) \subset \mathbb{H}^n$. This measure is left-invariant, and is hence also right-invariant by Corollary 3.5: in particular it does not depend on the choice of $x$.

CHAPTER 3

# Hyperbolic manifolds

## 1. Discrete groups of isometries

**1.1. Hyperbolic, flat, and elliptic manifolds.** We introduce three important classes of riemannian manifolds.

Definition 1.1. A *hyperbolic* (resp. *flat* or *elliptic*) *manifold* is a riemannian $n$-manifold that may covered by open sets isometric to open sets of $\mathbb{H}^n$ (resp. $\mathbb{R}^n$ o $S^n$).

A hyperbolic (resp. flat or elliptic) manifold has constant sectional curvature $-1$ (resp. 0 or $+1$). We show that the model $\mathbb{H}^n$ is indeed unique.

A *local isometry* $f: M \to N$ between riemannian manifold is a map such that every $x \in M$ has an open neighborhood $U$ such that $f|_U$ is an isometry onto its image.

Theorem 1.2. *A complete simply connected hyperbolic $n$-manifold $M$ is isometric to $\mathbb{H}^n$.*

Proof. Pick a point $x$ and choose an isometry $\varphi\colon U_x \to V$ between an open ball $U_x$ containing $x$ and an open ball $V \subset \mathbb{H}^n$. We show that $\varphi$ extends (uniquely) to an isometry $\varphi\colon M \to \mathbb{H}^n$.

For every $y \in M$, choose an arc $\alpha\colon [0,1] \to M$ from $x$ to $y$. By compactness there is a partition $0 = t_0 < t_1 < \ldots < t_k = 1$ and for each $i = 0, \ldots, k-1$ an isometry $\varphi_i\colon U_i \to V_i$ from an open set $U_i$ in $M$ containing $\alpha([t_i, t_{i+1}])$ to an open set $V_i \subset \mathbb{H}^n$.

Inductively on $i$, we now modify $\varphi_i$ so that $\varphi_{i-1}$ and $\varphi_i$ coincide on the component $C$ of $U_{i-1} \cap U_i$ containing $\alpha(t_i)$. To do so, note that

$$\varphi_{i-1} \circ \varphi_i^{-1}\colon \varphi_i(C) \longrightarrow \varphi_{i-1}(C)$$

is an isometry of open connected sets in $\mathbb{H}^n$ and hence extends to an isometry of $\mathbb{H}^n$. Then it makes sense to substitute $V_i$ with $\varphi_{i-1} \circ \varphi_i^{-1}(V_i)$, so that the new maps $\varphi_{i-1}$ and $\varphi_i$ coincide on $C$. The curve $\alpha$ now projects to a curve $\breve{\alpha}\colon [0,1] \to \mathbb{H}^n$. We define $\varphi(y) = \breve{\alpha}(1) = \varphi_{k-1}(\alpha(1))$.

To prove that $\varphi(y)$ is well-defined, we consider another path $\beta$ connecting $x$ to $y$. Since $M$ is simply-connected, there is a homotopy connecting $\alpha$ and $\beta$. The image of the homotopy is compact and is hence covered by finitely many open balls $U_i$ isometric to open balls $V_i \subset \mathbb{H}^n$ via some maps $\varphi_i$. By the Lebesgue number theorem, there is a $N > 0$ such that in the grid in $[0,1] \times [0,1]$ of $\frac{1}{N} \times \frac{1}{N}$ squares, the image of every square is entirely contained in at least one $U_i$. We can now modify as above the isometries $\varphi_i$ inductively on the grid, starting from the top-left square, so that they all glue up and the homotopy can be pushed to a homotopy between $\breve{\alpha}$ and $\breve{\beta}$, showing in particular that $\breve{\alpha}(1) = \breve{\beta}(1)$ and hence $\varphi(y)$ is well-defined.

The resulting map $\varphi\colon M \to \mathbb{H}^n$ is a local isometry, hence in particular a local homeomorphism. The fact that $\varphi$ is a local isometry and $M$ is complete together imply easily that every path in $\mathbb{H}^n$ can be lifted to a path in $M$. This path-lifting property implies that $\varphi$ is a covering, and since $\mathbb{H}^n$ is simply connected it is a homeomorphism. Hence $\varphi$ is an isometry. $\square$

The isometry $\varphi\colon M \to \mathbb{H}^n$ constructed in the proof is called a *developing map*. The same proof shows that a complete simply connected flat (or elliptic) $n$-manifold is isometric to $\mathbb{R}^n$ (or $S^n$).

**1.2. Discrete groups of isometries.** Let $M$ be a riemannian manifold. We give $\mathrm{Isom}(M)$ the compact-open topology: a pre-basis consists of the sets of all isometries such that $\varphi(K) \subset U$, where $K$ and $U$ vary among all (respectively) compact and open sets in $M$. With this topology $\mathrm{Isom}(M)$ is a topological group.

Proposition 1.3. *The following map is continuous and proper:*

$$\begin{aligned} F\colon \mathrm{Isom}(M) \times M &\to M \times M \\ (\varphi, p) &\mapsto (\varphi(p), p) \end{aligned}$$

Proof. Pick two open balls $B, B' \subset M$. We prove that the counterimage $F^{-1}(B' \times B)$ is relatively compact: this implies that $F$ is proper.

The counterimage consists of all pairs $(\varphi, p)$ with $p \in B$ and $\varphi(p) \in B'$. Since an isometry is determined by its first-order action on a point, the pair $(\varphi, p)$ is determined by the triple $(p, \varphi(p), d\varphi_p)$. The points $(p, \varphi(p))$ vary in the relatively compact set $B \times B'$ and $d\varphi_p$ then vary in a compact set homemorphic to $O(n)$. Therefore $F^{-1}(B' \times B)$ is contained in a relatively compact space, hence its is relatively compact. $\square$

Corollary 1.4. *If $M$ is compact then* $\mathrm{Isom}(M)$ *is compact*

A *discrete* group $\Gamma < \mathrm{Isom}(M)$ is a group which is discrete subset. Let $G$ be a group acting on a set $X$. Recall that

- the *stabilizer* of a point $x \in X$ is the subgroup $\{g \mid g(x) = x\} < G$,
- the *orbit* of $x$ is $\{g(x) \mid g \in G\} \subset X$,
- the *quotient* $X/_G$ is defined by quotienting every orbit to a point,
- the action is *free* if for any non-trivial $g \in G$ and any $x \in X$ we have $g(p) \neq p$.

Proposition 1.5. *Let $M$ be a connected riemannian manifold and $\Gamma <$* $\mathrm{Isom}(M)$ *a discrete group acting freely on $M$. There is a riemannian structure on $M/_\Gamma$ such that $\pi\colon M \to M/_\Gamma$ is a local isometry.*

PROOF. Since $\Gamma$ is discrete, Proposition 1.3 easily implies that every orbit in $M$ is discrete. Pick a point $p \in M/_\Gamma$: its counterimage $O = \pi^{-1}(p) \subset M$ is an orbit and is hence discrete. For every $q \in O$ we define

$$D(q) = \left\{ x \in M \mid d(x,q) < d(x,q')\ \forall q' \in O, q' \neq q \right\}.$$

The set $D(q)$ is open and contains $q$. We have $D(q) \cap DD(q') = \emptyset$ for all $q \neq q'$ and $g(D(q)) = D(g(q))$ for all $g \in \Gamma$. Therefore $\pi$ restricts on each $D(q)$ to an homeomorphism onto an open set $D(p)$ containing $p$. Since $\pi^{-1}(D(p)) = \sqcup_q D(q)$, the open set $D(p)$ is well-covered and hence $\pi$ is a covering map.

We can give $D(p)$ the metric tensor of $D(q)$: this definition does not depend on $q \in O$ since $\Gamma$ consists of isometries. If $p \neq p'$ then the metric tensors coincide on $D(p) \cap D(p')$. We get a riemannian structure on $M/_\Gamma$ and $\pi$ is a local isometry by construction. $\square$

The open set $D(p) \subset M/_\Gamma$ containing $p$ is called a *Dirichlet domain* and we will use it again soon.

COROLLARY 1.6. *Let $\Gamma$ be a discrete group of isometries acting freely on $\mathbb{H}^n$. Then $\mathbb{H}^n/_\Gamma$ is a complete hyperbolic $n$-manifold.*

An analagous corollary holds also for flat and elliptic manifolds. Concerning completeness, we use the following.

EXERCISE 1.7. Let $p\colon M \to N$ be a covering and local isometry between riemannian manifolds. Then $M$ is complete if and only if $N$ is.

A converse of Corollary 1.6 holds.

PROPOSITION 1.8. *Every complete hyperbolic $n$-manifold $M$ is isometric to $\mathbb{H}^n/_\Gamma$ for some discrete group $\Gamma < \mathrm{Isom}(\mathbb{H}^n)$ acting freely.*

PROOF. The universal cover of $M$ is complete, hyperbolic, and simply connected: hence it is isometric to $\mathbb{H}^n$ by Theorem 1.2. The deck transformations of the covering $\mathbb{H}^n \to M$ are necessarily isometries: they form a subgroup $\Gamma < \mathrm{Isom}(\mathbb{H}^n)$ which acts freely.

The fiber of a point in $M$ is a discrete orbit $O \subset \mathbb{H}^n$, and this implies that $\Gamma$ is discrete: fix $x \in O$ and note that the continuous map $\mathrm{Isom}(M) \to \mathbb{H}^n$ that sends $g$ to $g(x)$ sends $\Gamma$ injectively to $O$. $\square$

REMARK 1.9. As usual with universal coverings, the deck transformation group $\Gamma$ is isomorphic to the fundamental group $\pi_1(M)$.

The same proof shows that every complete flat or spherical $n$-manifold is isometric to $\mathbb{R}^n/_\Gamma$ or $S^n/_\Gamma$ for some discrete group $\Gamma$ of isometries acting freely on $\mathbb{R}^n$ or $S^n$.

COROLLARY 1.10. *There is a natural 1-1 correspondence*

$$\left\{ \begin{array}{c} \text{complete hyperbolic} \\ \text{manifolds up to isometry} \end{array} \right\} \longleftrightarrow \left\{ \begin{array}{c} \text{discrete subgroups of } \mathrm{Isom}(\mathbb{H}^n) \\ \text{without non} - \text{trivial elliptics} \\ \text{up to conjugation} \end{array} \right\}.$$

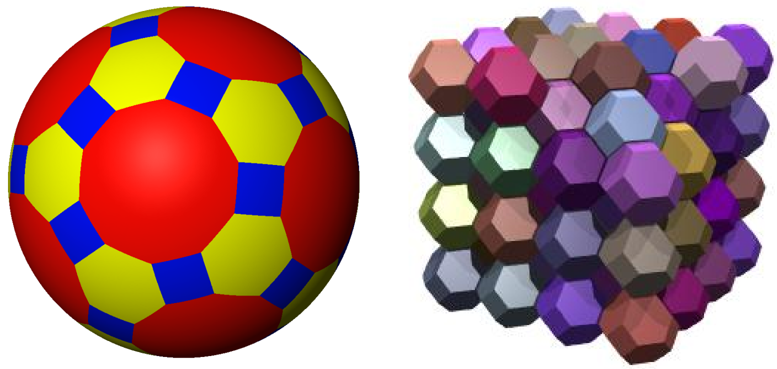


FIGURE 1. A tessellation of $S^2$ into squares, hexagons, and decagons, and a tassellation of $\mathbb{R}^3$ into truncated octahedra.

PROOF. A group $\Gamma$ acts freely on $\mathbb{H}^n$ if and only if it does not contain non-trivial elliptics. The proof of Proposition 1.8 shows how to pass from $M$ to $\Gamma$: the only choice we make is an isometry between the universal covering of $M$ and $\mathbb{H}^n$. Different choices produce conjugate groups. $\square$

**1.3. Polyhedra and tessellations.** Another method to construct hyperbolic (flat, elliptic) manifolds consists in assembling polyhedra. A half-space in $\mathbb{H}^n$, $\mathbb{R}^n$, or $S^n$ is the closure of one of the two portions of space delimited by a hyperplane. A set of half-spaces is *locally finite* if a compact set intersects only finitely many of their boundary hyperplanes.

DEFINITION 1.11. A $n$-dimensional *polyhedron* in $\mathbb{H}^n$ (or $\mathbb{R}^n$, $S^n$) is the intersection $P = \cap_i H_i$ of a locally finite set of half-spaces. We also assume that $P$ has non-empty interior.

A subspace $S \subset \mathbb{H}^n$ is *convex* if $x, y \in S$ implies that the segment connecting $x, y$ is also contained in $S$. A polyhedron is clearly convex because it is the intersection of convex sets.

If non-empty, the intersection $\partial H_i \cap P$ is always a $k$-dimensional polyhedron for some $k \leqslant n$, called *face*. A face of dimension $n-1$ is a *facet*.

DEFINITION 1.12. A *tessellation* of $\mathbb{H}^n$ (or $\mathbb{R}^n$, $S^n$) is a set of polyhedra that cover the space and may intersect in pairs only in common faces.

Some examples are shown in Fig. 1.

**1.4. Fundamental and Dirichlet domains.** Let now $\Gamma$ be a discrete group acting freely on $\mathbb{H}^n$.

DEFINITION 1.13. A *fundamental domain* for $\Gamma$ is an open connected set $D \subset \mathbb{H}^n$ such that $D$ intersects every orbit in *at most* one point and $\overline{D}$ intersects every orbit in *at least* one point.

In other words, the open sets $\{g(D)|g \in \Gamma\}$ are disjoint, while their closures $\{\overline{g(D)}|g \in \Gamma\}$ cover $\mathbb{H}^n$. In presence of a fundamental domain $D$ we may see $M$ as $\overline{D}$ with the points in $\partial D$ lying in the same orbit identified.

We have already encountered a fundamental domain in the proof of Proposition 1.5, the *Dirichlet domain*. For a point $q \in \mathbb{H}^n$ we set

$$D(q) = \big\{x \in M \ \big|\ d(x,q) < d(x,g(q)) \ \forall g \in \Gamma, g \neq \mathrm{id}\big\}.$$

PROPOSITION 1.14. *Pick $q \in \mathbb{H}^n$. The Dirichlet domain $D(q)$ is a fundamental domain for $\Gamma$. Its closure $\overline{D(q)}$ is a polyhedron and*

$$\big\{\overline{D(g(q))} \ \big|\ g \in \Gamma\big\}$$

*is a tessellation of $\mathbb{H}^n$.*

PROOF. We have $g(D(q)) = D(g(q))$ and the open sets $\{D(g(q)) \mid g \in \Gamma\}$ are disjoint. Every point $x \in \mathbb{H}^n$ has at least one nearest point in the orbit of $p$: hence $x \in \overline{D(g(p))}$ for some $g \in \Gamma$. Therefore the elements $\big\{\overline{g(D(q))} \mid g \in \Gamma\big\}$ cover $\mathbb{H}^n$. Hence $D(q)$ is a fundamental domain.

The domain $D(q)$ is the intersection of the open half-paces

$$H_g = \big\{x \in \mathbb{H}^n \ \big|\ d(x,q) < d(x,g(q))\big\}.$$

Its closure $\overline{D(q)}$ is the intersection of the same closed half-spaces. Since the orbit of $q$ is discrete, the intersection of these half-spaces is locally finite and hence $\overline{D(q)}$ is a polyhedron.

Two adjacent polyhedra $\overline{D(g(q))}$ and $\overline{D(g'(q))}$ intersect in the common face obtained by intersecting each polyhedron with the hyperplane

$$\big\{x \in \mathbb{H}^n \mid d(x,g(q)) = d(x,g'(q))\big\}.$$

Hence they form a tessellation of $\mathbb{H}^n$. $\square$

Everything we said also holds for $\mathbb{R}^n$ and $S^n$.

**1.5. Flat manifolds.** We show some examples, starting with some simple (and well-known) flat manifolds. The group of translations in $\mathbb{R}^n$ may be identified with $\mathbb{R}^n$ itself. Let $\Gamma = \mathbb{Z}^n$ be the integer translations. The quotient $\mathbb{R}^n/_\Gamma$ is naturally diffeomorphic to the $n$-dimensional torus:

$$\mathbb{R}^n/_{\mathbb{Z}^n} = (\mathbb{R}/_{\mathbb{Z}})^n \cong \underbrace{S^1 \times \ldots \times S^1}_{n}$$

which is hence a flat manifold.

EXERCISE 1.15. For any $q \in \mathbb{R}^n$ the Dirichlet domain $D(q)$ is a $n$-dimensional unit cube centered at $q$.

The flat $n$-torus may be seen as the unit $n$-cube with its opposite facets identified by a translation. The two-dimensional case is shown in Fig. 2-(left): by identifying the opposite sides of a square we get a torus.

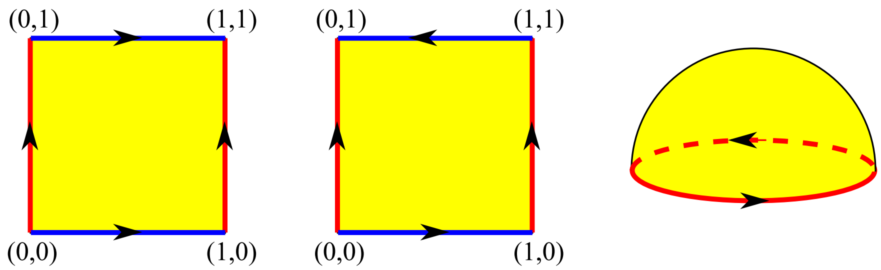



FIGURE 2. A fundamental domain in $\mathbb{R}^2$ for the torus (left) and the Klein bottle (center): opposite sides should be identified as indicated by the arrows. A fundamental domain in $S^2$ for $\mathbb{RP}^2$ (right).

Among flat surfaces we also find the *Klein bottle* by taking $\Gamma$ as the group generated by the following isometries:

$$\tau\colon (x,y)\mapsto (x+1,y), \qquad \eta\colon (x,y)\mapsto (1-x,y+1).$$

A fundamental domain for the Klein bottle is shown in Fig. 2-(center).

REMARK 1.16. The subgroups $\langle\tau\rangle$ and $\langle\eta\rangle$ generated respectively by $\tau$ and $\eta$ are both isomorphic to $\mathbb{Z}$. Note however that $\mathbb{R}^2/_{\langle\tau\rangle}$ is an infinite cilinder and $\mathbb{R}^2/_{\langle\eta\rangle}$ is an infinite Möbius strip. Being subgroups of $\Gamma$, both spaces cover the Klein bottle.

The subgroup of $\Gamma$ generated by the traslations $\tau$ and $\eta^2$ is isomorphic to $\mathbb{Z}^2$ and has index 2 in $\Gamma$. The Klein bottle has indeed a double covering isometric to a flat torus. Its fundamental domain is a rectangle with vertices $(0,0)$, $(1,0)$, $(0,2)$, $(1,2)$.

The $n$-torus possesses a continuous family of non-isometric flat metrics. A *lattice* $\Gamma<\mathbb{R}^n$ is a discrete subgroup isomorphic to $\mathbb{Z}^n$ which spans $\mathbb{R}^n$ as a vector space. We see $\Gamma$ as a group of translations.

EXERCISE 1.17. The flat manifold $\mathbb{R}^n/_\Gamma$ is diffeomorphic to the $n$-torus. A fundamental domain is the parallelotope spanned by $n$ generators of $\Gamma$.

REMARK 1.18. A Dirichlet domain is not necessarily a parallelotope. For instance, consider the *hexagonal torus* $\mathbb{C}/_\Gamma$ with $\Gamma$ generated by 1 and $e^{\frac{\pi i}{3}}$. The Dirichlet domain of 0 is a regular hexagon.

Mettere figura toro esagonale?

**1.6. Elliptic manifolds.** Every elliptic manifold is covered by $S^n$ and is hence compact and with finite fundamental group (because the covering has finite degree).

An important example is real projective space $\mathbb{RP}^n = S^n/_\Gamma$ where is the cyclic group of order two generated by the antipodal map $\iota\colon v\to -v$.

EXERCISE 1.19. For any $q\in S^n$, the Dirichlet domain is the emisphere centered as $q$.

The two-dimensional case is shown in Fig. 2-(right). In dimension $n = 3$ there are more elliptic manifolds. Let $p$ and $q$ be coprime integers and set $\omega = e^{\frac{2\pi i}{p}}$. We identify $\mathbb{R}^4$ with $\mathbb{C}^2$ and see $S^3$ as

$$S^3 = \big\{(z,w) \in \mathbb{C}^2 \ \big|\ |z|^2 + |w|^2 = 1\big\}.$$

The map

$$f(z,w) = (\omega z, \omega^q w)$$

is an isometry of $\mathbb{R}^4$ because it consists of two simultaneous rotations on the coordinate planes $w = 0$ e $z = 0$. The map $f$ hence induces an isometry of $S^3$. It has order $p$ and noone of its iterates $f, f^2, \ldots, f^{p-1}$ has a fixed point. Therefore the group $\Gamma = \langle f \rangle$ generated by $f$ acts freely on $S^3$, and is discrete because it is finite.

The fundamental group of $S^3/_\Gamma$ is isomorphic to $\Gamma \cong \mathbb{Z}_p$. This elliptic manifold is called a *lens space* and indicated with the symbol $L(p,q)$.

**1.7. Selberg lemma.** The reader might now expect to find some examples of compact hyperbolic manifolds, constructed as above from explicit discrete subgroups $\Gamma$ of $\mathrm{Isom}(\mathbb{H}^n)$ acting freely on $\mathbb{H}^n$. It turns out however that exhibiting such groups is quite hard, and one usually constructs hyperbolic manifolds by other means. The method we present here is nonconstructive and based on the following algebraic result, while more geometric methods will follow in the next chapters.

LEMMA 1.20 (Selberg lemma). *Let $G$ be a finitely generated subgroup of $\mathrm{GL}(n,\mathbb{C})$. There is a finite-index normal subgroup $H \triangleleft G$ without non-trivial finite-order elements.*

COROLLARY 1.21. *Let $\Gamma$ be a finitely generated discrete subgroup of isometries of $\mathbb{H}^n$. There is a finite-index normal subgroup $\Gamma' \triangleleft \Gamma$ that acts freely on $\mathbb{H}^n$.*

PROOF. The group $\mathrm{Isom}(\mathbb{H}^n)$ is isomorphic to $O(n,1)_* < \mathrm{GL}(n+1,\mathbb{C})$, hence Selberg lemma applies to $\Gamma$. We get a finite-index normal subgroup $\Gamma' \triangleleft \Gamma$ without finite-order elements.

The group $\Gamma'$ acts freely unless it contains a non-trivial elliptic element $\varphi$ which fixes some point $x \in \mathbb{H}^n$. The element $\varphi$ would have finite order because $\Gamma$ is discrete and the stabilizer $O(n)$ of a point is compact: a contradiction. $\square$

By Selberg lemma we can forget about the "acting freely" hypothesis and concentrate on the construction of discrete subgroups of $\mathrm{Isom}(\mathbb{H}^n)$. The isometry group of a tessellation of $\mathbb{H}^n$ is the group formed by all the isometries that fix the tessellation as a set of polyhedra. Here is a source of discrete groups:

PROPOSITION 1.22. *If a tessellation consists of compact polyhedra in $\mathbb{H}^n$, its isometry group $\Gamma$ is discrete.*

FIGURE 3. Le tassellazioni $(2,3,3)$, $(2,3,4)$ e $(2,3,5)$ nella sfera.

PROOF. A compact polyhedron is the convex hull of finitely many points, its vertices. An isometry that fixes the vertices fixes the polyhedron pointwise. Therefore the isometry group of every polyhedron is finite. A bounded set $B \subset \mathbb{H}^n$ contains only finitely many polyhedra of the tessellation, hence there are only finitely many $\varphi \in \Gamma$ such that $\varphi(B) \cap B \neq \emptyset$. $\square$

**1.8. Triangular groups.** We construct here some discrete subgroups of isometries. We start with the following.

EXERCISE 1.23. Given three real numbers $0 < \alpha, \beta, \gamma < \pi$ there is a triangle $\Delta$ with inner angles $\alpha, \beta, \gamma$ inside $\mathbb{H}^2, \mathbb{R}^2$, or $S^2$ depending on whether the sum $\alpha + \beta + \gamma$ is smaller, equal, or bigger than $\pi$.

Let $a, b, c \geqslant 2$ be three natural numbers and $\Delta$ be a triangle with inner angles $\frac{\pi}{a}, \frac{\pi}{b}, \frac{\pi}{c}$. The triangle $\Delta$ lies in $\mathbb{H}^2$, $\mathbb{R}^2$, or $S^2$ depending on whether $\frac{1}{a} + \frac{1}{b} + \frac{1}{c}$ is smaller, equal, or bigger than 1. In any case, by mirroring iteratively $\Delta$ along its edges we construct a tessellation of the space.

EXAMPLE 1.24. The triples realizable in $S^2$ are $(2,2,c)$, $(2,3,3)$, $(2,3,4)$, and $(2,3,5)$: the last three tessellations are shown in Fig. 3 and are linked to platonic solids. Their isometry groups are respectively the isometry group of the tetrahedron, of the cube (or octahedron), and of the icosahedron (or dodecahedron). These groups act transitively on the triangles of the tessellations, and have order 24, 48, and 120 respectively. They are actually isomorphic to $S_4$, $S_4 \times \mathbb{Z}_2$, and $A_5 \times \mathbb{Z}_2$.

EXAMPLE 1.25. The triples realizable in $\mathbb{R}^2$ are $(2,3,6)$, $(2,4,4)$, and $(3,3,3)$: the tessellations are shown in Fig. 4.

EXAMPLE 1.26. There are infinitely many triples realizable in $\mathbb{H}^2$, some are shown in Fig. 5.

The isometry group of this tessellation is called a *triangular group*.

EXERCISE 1.27. The triangular group acts transitively on the triangles of the tessellation and is generated by the reflections $x, y, z$ along the three sides of $\Delta$. A presentation for the group is

$$\langle x, y, z \mid x^2, y^2, z^2, (xy)^c, (yz)^a, (zx)^b \rangle.$$

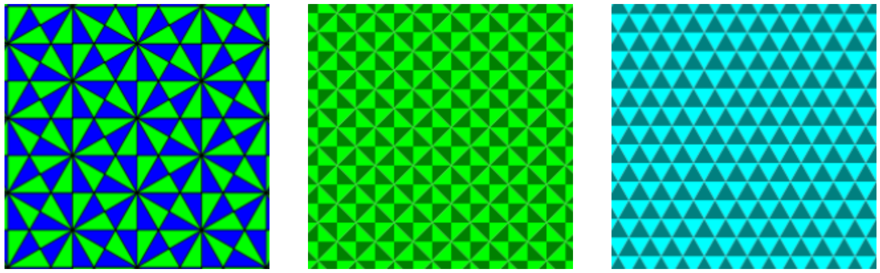


FIGURE 4. Le tassellazioni $(2,3,6)$, $(2,4,4)$ e $(3,3,3)$ nel piano.

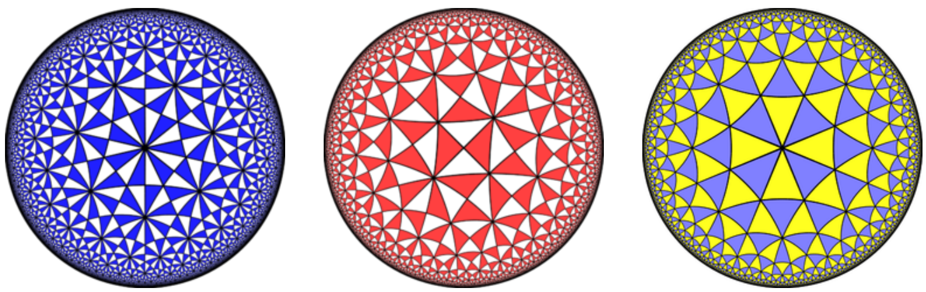


FIGURE 5. Le tassellazioni $(2,3,7)$, $(2,4,5)$ e $(3,3,4)$ nel piano iperbolico.

We have constructed some discrete groups $\Gamma$ of $\mathbb{H}^2$. Each such contains infinitely many elliptic elements, such as reflections along lines and finite-order rotations around vertices of the triangles. However, Selberg lemma guarantees the existence of a finite-index subgroup $\Gamma' < \Gamma$ that avoids all the non-trivial elliptics elements and hence acts freely.

Suppose for simplicity that $a, b, c$ are different, hence $\Delta$ has no symmetries: the quotient $\mathbb{H}^n/_{\Gamma}$ is isometric to $\Delta$ and the surface $\mathbb{H}^n/_{\Gamma'}$ is compact in virtue of the following:

EXERCISE 1.28. If $\Gamma' < \Gamma$ has index $h$, the surface $\mathbb{H}^2/_{\Gamma'}$ tessellates into $h$ copies of $\Delta$.

**1.9. Ideal polyhedra.** We can try to generalize the triangular groups in two natural ways: by taking triangles with vertices "at infinity", or polyhedra of higher dimension.

A polyhedron $P \subset \mathbb{H}^n$ has its closure $\overline{P} \subset \overline{\mathbb{H}^n}$ and its points at infinity $\overline{P} \setminus P$. An isolated point in $\overline{P} \setminus P$ is called a *vertex at infinity*, while the ordinary vertices of $P$ in $\mathbb{H}^n$ are the *finite vertices*. An *ideal polyhedron* is a polyhedron without finite vertices, whose points at infinity form a discrete (hence finite) set. Fare figura

For instance, an *ideal polygon* is shown in Fig. A peculiar aspect of hyperbolic geometry is that ideal polyhedra are non-compact but have finite volume, as we now see. Given a horosphere $O$ centered at $p$ and a domain

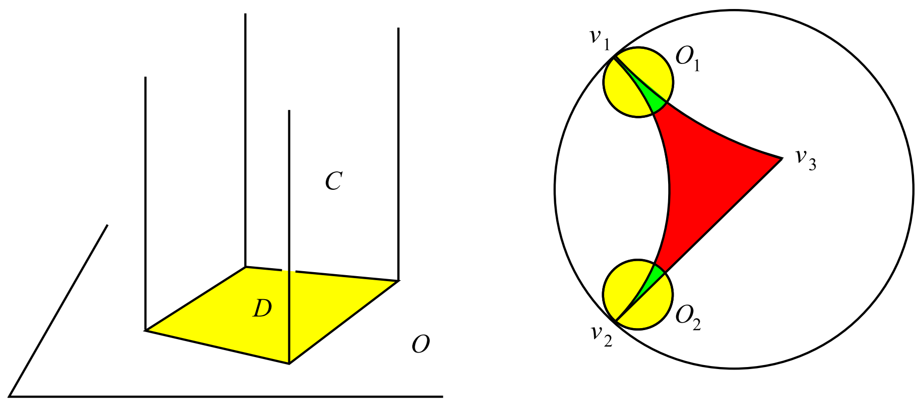



FIGURE 6. The cone $C$ over a domain $D \subset O$ has volume proportional to the area of $D$ (left). If the domain is compact, the cone has finite volume: therefore an ideal polyhedron has finite volume (right).

$D \subset O$, the *cone* $C$ of $D$ over $p$ is the union of all half-lines exiting from $D$ towards $p$, see Fig. 6.

PROPOSITION 1.29. *Let $O$ be a horosphere centered at $p \in \partial\mathbb{H}^n$, $D \subset O$ any domain and $C$ the cone over $D$. The following equality holds:*

$$\mathrm{Vol}(C) = \frac{\mathrm{Vol}_O(D)}{n-1}$$

*where* $\mathrm{Vol}_O$ *indicates the volume in the flat $(n-1)$-manifold $O$.*

PROOF. Consider the half-space model with $p = \infty$ and $O$ at some height $x_n = h$ as in Fig. 6-(left). We obtain

$$\mathrm{Vol}(C) = \int_D dx \int_h^\infty \frac{1}{t^n} dt = \frac{1}{n-1}\int_D \frac{dx}{h^{n-1}} = \frac{1}{n-1} \cdot \mathrm{Vol}_O(D).$$

□

COROLLARY 1.30. *A polyhedron $P \subset \mathbb{H}^n$ with $\partial P \subset \partial\mathbb{H}^n$ consisting of finitely many points has finite volume.*

PROOF. For every $p \in \partial P$, a small horoball centered at $p$ intersects $P$ into a cone which has finite volume. The polyhedron $P$ decomposes into finitely many cones and a bounded region, see Fig. 6-(right). □

The area of a polygon with both finite and infinite vertices can be calculated using an extremely simple formula. The sum of the inner angles of a hyperbolic polygon is strictly smaller than that of a euclidean polygon with the same number of sides, and the difference is precisely its area. The interior angle of a vertex at infinity is zero.

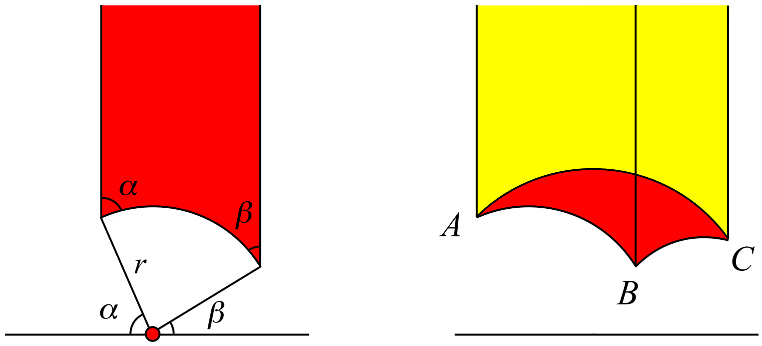



FIGURE 7. A triangle with at least an ideal vertex (left). The area of a triangle with finite vertices can be derived as the area difference of triangles with one ideal vertex (right).

PROPOSITION 1.31. *A polygon $P$ with inner angles $\alpha_1, \ldots, \alpha_n$ has area*

$$\text{Area}(P) = (n-2)\pi - \sum_{i=1}^{n} \alpha_i.$$

PROOF. Every polygon decomposes into triangles, and it suffices to prove the formula on these. Consider first a triangle $T$ with at least one vertex at infinity. We use the half-plane model and send this vertex to $\infty$ as in Fig. 7-(left). We suppose that the red dot is the origin of $\mathbb{R}^2$, hence

$$T = \big\{(r\cos\theta, y) \mid \beta \leqslant \theta \leqslant \pi - \alpha,\ \ y \geqslant r\sin\theta\big\}$$

and we get

$$\begin{aligned}
\text{Area}(T) &= \int_T \frac{1}{y^2} dxdy = \int_{\pi-\alpha}^{\beta} \int_{r\sin\theta}^{\infty} \frac{-r\sin\theta}{y^2} dyd\theta \\
&= \int_{\pi-\alpha}^{\beta} -r\sin\theta \left[-\frac{1}{y}\right]_{r\sin\theta}^{\infty} d\theta = \int_{\beta}^{\pi-\alpha} \frac{r\sin\theta}{r\sin\theta} d\theta \\
&= \int_{\beta}^{\pi-\alpha} 1 = \pi - \alpha - \beta.
\end{aligned}$$

The area of a triangle with finite vertices $ABC$ is deduced as in Fig. 7-(right) using the formula

$$\text{Area}(ABC) = \text{Area}(AB\infty) + \text{Area}(BC\infty) - \text{Area}(AC\infty).$$

□

We may construct various tessellations using ideal polygons, a natural one being the *Farey tessellation* which is defined as follows. Consider $\mathbb{H}^2$ with the half-plane model. For any pair of rational numbers $\frac{p}{q}, \frac{r}{s} \in \mathbb{Q} \cup \{\infty\} \subset \partial H^2$ such that $ps - qr = \pm 1$ we draw the geodesic in $H^2$ connecting them. We obtain a tessellation of $\mathbb{H}^2$ in ideal triangles, shown in Fig. 8. Recall that $\text{Isom}(H^2) = \mathbb{P}\text{GL}_2(\mathbb{R})$.

Fare o trovare figura con $H^2$.

FIGURE 8. The Farey tessellation of $\mathbb{H}^2$ ideal triangles.

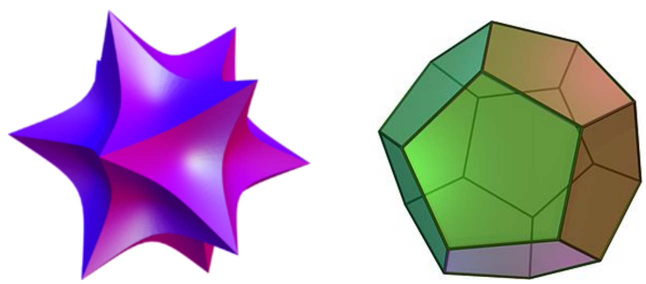


FIGURE 9. A hyperbolic regular icosahedron and a euclidean regular dodecahedron.

EXERCISE 1.32. The Farey tessellation has isometry group $\mathbb{P}\mathrm{GL}_2(\mathbb{Z})$.

**1.10. Platonic solids.** Each platonic solid $P$ has a nice continuous family of representations in the three geometries $\mathbb{H}^3, \mathbb{R}^3$, and $S^3$, which at few discrete points generate a tessellation of the space.

To construct this family pick any point $x$ in $\mathbb{H}^3$, $\mathbb{R}^3$, or $S^3$ and represent $P$ centered at $x$ with varying size. To do this, represent $P$ in the tangent space $T_x$ centered in the origin and with some radius $t > 0$. Consider the image of its vertices by the exponential map in $\mathbb{H}^3$, $\mathbb{R}^3$, or $S^3$ and take their convex hull. We indicate by $P(t)$ the resulting platonic solid, with this convention: if $t < 0$ then $P(t)$ is the solid obtained in $\mathbb{H}^3$ with parameter $-t$, if $t = 0$ then $P(0)$ is the usual euclidean solid (unique up to dilations), if $t > 0$ then $P(t)$ is the solid obtained in $S^3$.

The platonic solid is defined for all $t \in [-\infty, \frac{\pi}{2}]$: when $t = -\infty$ we get an ideal platonic solid with all vertices at infinity, while as $t = \frac{\pi}{2}$ the platonic solid degenerates to a half-sphere in $S^n$. The dihedral angle $\theta(t)$ varies continuously with $t$, since when $t \to 0$ the polyhedron shrinks and every geometry tends to the euclidean one when we shrink objects. It is a strictly monotone increasing function.

| polyhedron | $\theta = \frac{\pi}{3}$ | $\theta = \frac{2\pi}{5}$ | $\theta = \frac{\pi}{2}$ | $\theta = \frac{2\pi}{3}$ |
|---|---|---|---|---|
| tetrahedron | ideal $\mathbb{H}^3$ | $S^3$ | $S^3$ | $S^3$ |
| cube | ideal $\mathbb{H}^3$ | $\mathbb{H}^3$ | $\mathbb{R}^3$ | $S^3$ |
| octahedron | | | ideal $\mathbb{H}^3$ | $S^3$ |
| icosahedron | | | | $\mathbb{H}^3$ |
| dodecahedron | ideal $\mathbb{H}^3$ | $\mathbb{H}^3$ | $\mathbb{H}^3$ | $S^3$ |

TABLE 1. The platonic solids with dihedral angle $\theta$ that divide $2\pi$.

The *vertex valence* of $P$ is the number of edges incident to each vertex of $P$.

PROPOSITION 1.33. *Let $n \in \{3, 4, 5\}$ be the vertex valence of $P$. It holds*

$$\theta\left(\left[-\infty, \frac{\pi}{2}\right]\right) = \left[\frac{n-2}{n}\pi, \pi\right].$$

PROOF. Since $\theta$ is continuous and monotone increasing, it suffices to show that $\theta(-\infty) = \frac{n-2}{n}$ and $\theta(\frac{\pi}{2}) = \pi$.

By intersecting the ideal polyhedron $P(-\infty)$ with a small horosphere $O$ centered at an ideal vertex $v$ we get a regular $n$-gon $P$ in the euclidean plane $O$, with interior angles $\frac{n-2}{n}\pi$. The dihedral angle at an edge $e$ is measured by intersecting the polyhedron with a hypersurface orthogonal to $e$: since $O$ is orthogonal we get $\theta(-\infty) = \frac{n-2}{n}$.

The polyhedron $P(\frac{\pi}{2})$ is a semisphere and hence $\theta(\frac{\pi}{2}) = \pi$. □

When $\theta(t)$ divides $2\pi$, then by repeatedly mirroring $P(t)$ along its faces we get a tessellation of the space: these tessellations are listed in Table 1 and some pictures are shown in Fig. 10 and 11.

Every such tessellation of $\mathbb{H}^3$ has a discrete isometry group $\Gamma$ and by Selberg's lemma there is a finite-index subgroup $\Gamma' < \Gamma$ acting on $\mathbb{H}^3$ without fixed points. We mention few examples.

EXAMPLE 1.34. The *Seifert-Weber space* is a closed hyperbolic 3-manifold $M = \mathbb{H}^3/_{\Gamma'}$ related to the tessellation into hyperbolic dodecahedra with dihedral angle $\frac{2\pi}{5}$. The manifold $M$ may be obtained from a single such dodecahedron by identifying the opposite faces after making a $\frac{3\pi}{5}$-turn.

EXAMPLE 1.35. The *Poincaré homology sphere* is a closed elliptic 3-manifold $M = S^3/_{\Gamma'}$ related to the tessellation into spherical dodecahedra with dihedral angle $\frac{2\pi}{3}$. The manifold $M$ may be obtained from a single such dodecahedron by identifying the opposite faces after making a $\frac{\pi}{5}$-turn.

EXAMPLE 1.36. The *figure-eight knot complement* is a hyperbolic 3-manifold $M = S^3/_{\Gamma'}$ related to the tessellation into ideal regular tetrahedra with dihedral angle $\frac{\pi}{3}$. It is diffeomorphic to the complement in $S^3$ of the

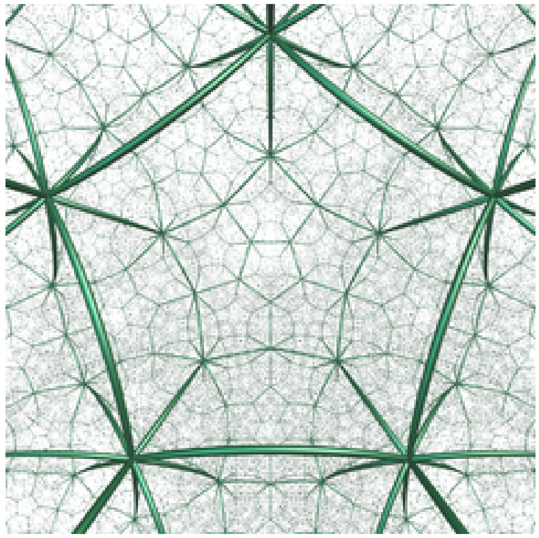


FIGURE 10. The tessellation of $\mathbb{H}^3$ into regular dodecahedra with dihedral angle $\theta = \frac{2\pi}{5}$ in the disc model.

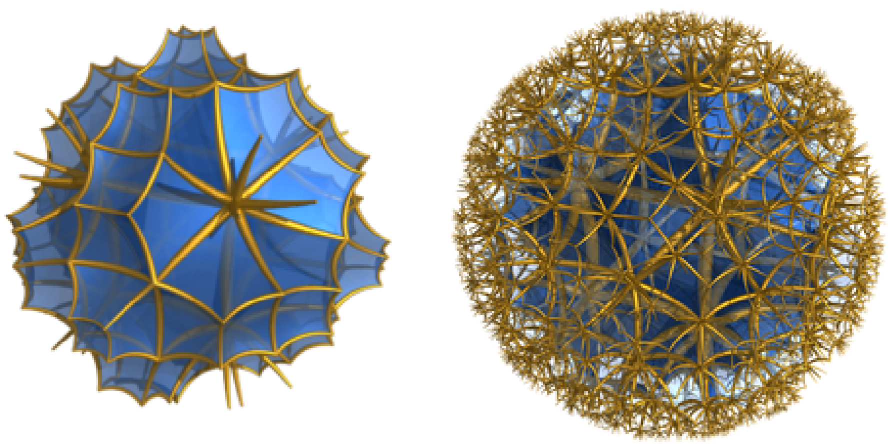


FIGURE 11. The tessellation of $\mathbb{H}^3$ into regular cubes with dihedral angle $\theta = \frac{2\pi}{5}$ in the disc model.

figure-eight knot shown in Fig. 12 and tessellates into two regular ideal tetrahedra. It is not compact but has finite volume.

## 2. Generalities on hyperbolic manifolds

We construct some basic examples and study the basic properties of hyperbolic manifolds.

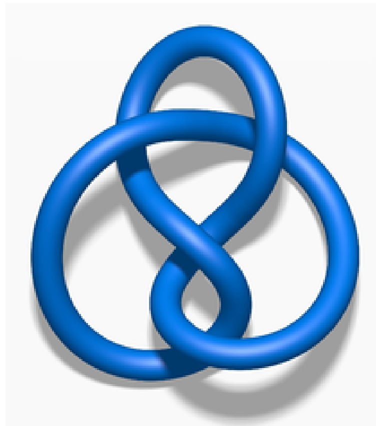


FIGURE 12. The figure-eight knot. Its complement in $S^3$ may be tessellated by two ideal regular tetrahedra.

**2.1. Tubes.** It is typically difficult to construct a hyperbolic manifold by exhibiting a discrete group $\Gamma$ of isometries of $\mathbb{H}^n$, except in some elementary cases.

Consider the cyclic group $\Gamma = \langle \varphi \rangle$ generated by a hyperbolic transformation $\varphi$ on $\mathbb{H}^n$ with axis $l$ and minimum displacement $d > 0$. The iterates $\varphi^k$ are again hyperbolic transformations with axis $l$ and displacement $kd$. Therefore $\Gamma$ acts freely on $\mathbb{H}^n$. The quotient manifold $M = \mathbb{H}^n/_\Gamma$ is called an *infinite tube*.

EXERCISE 2.1. Fix $q \in l$. Let $q_1, q_2$ be the two points in $l$ at distance $\frac{d}{2}$ from $q$ and $\pi_1, \pi_2$ the two hyperplanes orthogonal to $l$ in $q_1, q_2$. The Dirichlet domain $U(q)$ is the space comprised between $\pi_1$ and $\pi_2$.

The infinite tube $M = \mathbb{H}^n/_\Gamma$ is obtained from $U(q)$ by identifying $\pi_1$ and $\pi_2$ along $\varphi$. Its fundamental group is isomorphic to $\Gamma \cong \mathbb{Z}$. The axis $l$ projects in $M$ onto a closed geodesic $\gamma$ of length $d$.

PROPOSITION 2.2. *An infinite tube is diffemorphic to $S^1 \times \mathbb{R}^{n-1}$ or $S^1 \tilde{\times} \mathbb{R}^{n-1}$ according to whether $\varphi$ is orientation-preserving or not.*

PROOF. By projecting $\mathbb{H}^n$ orthogonally onto $l$, we give $\mathbb{H}^n$ the structure of a $\mathbb{H}^{n-1}$-bundle over $l$ which is preserved by $\varphi$ and hence descends to a structure of $\mathbb{H}^{n-1}$-bundle over $\gamma$. Therefore $M$ is diffeomorphic to the normal bundle of $\gamma$ in $M$. The conclusion follows from the classification of vector bundles over $S^1$, see Proposition 1.13 from Chapter 1. □

A *tube of radius* $R$ is the quotient $N_R(l)/_\Gamma$ where $N_R(l)$ is the $R$-neighborhood of $l$, the set of all points of distance at most $R$. It is diffeomorphic to $S^1 \times D^{n-1}$ or $S^1 \tilde{\times} D^{n-1}$: in particular it is compact. Note that the bounday of a tube is not geodesic.

Mettere formule volume? Da "volume of tubes in hyp 3-man"

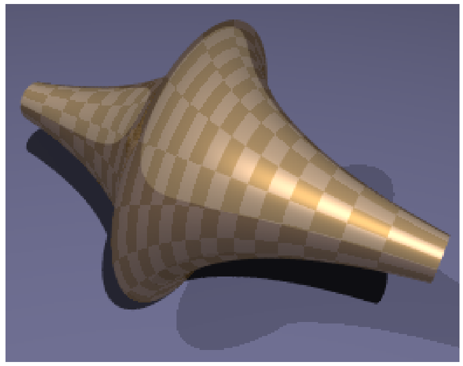


FIGURE 13. The *pseudosphere* is a surface in $\mathbb{R}^3$ isometric to the union of two truncated cusps, each with constant gaussian curvature $-1$.

**2.2. Cusps.** Let now $\Gamma < \text{Isom}(\mathbb{R}^{n-1})$ be a discrete group of euclidean isometries acting freely on $\mathbb{R}^{n-1}$: the quotient $M = \mathbb{R}^{n-1}/_\Gamma$ is a flat $(n-1)$-manifold. If we use the half-space model for $\mathbb{H}^n$ with coordinates $(x, t)$, every element $\varphi \in \Gamma$ acts as a parabolic transformation on $\mathbb{H}^n$ by sending $(x, t)$ to $(\varphi(x), t)$. The whole group $\Gamma$ is a discrete group of parabolic transformations of $\mathbb{H}^n$ fixing the point $\infty$.

The quotient $\mathbb{H}^n/_\Gamma$ is naturally identified with $M \times \mathbb{R}_{>0}$. The metric tensor on the point $(x, t)$ is

$$g_{(x,t)} = \frac{g^M_x \oplus 1}{t^2}$$

where $g^M$ is the metric tensor of $M$. The manifold $\mathbb{H}^n/_\Gamma$ is called a *cusp*.

REMARK 2.3. The coordinate $t$ may be parametrized more intrinsically using arc-length. As we have seen in Proposition 1.17 from Chapter 2, a vertical geodesic with unit speed is parametrized as $t = e^u$. Using $u$ instead of $t$ the cusp is isometric to $M \times \mathbb{R}$ with metric tensor

$$g_{(x,u)} = (e^{-2u} g_M) \oplus 1.$$

When $u$ increases, the $M$ factor shrinks exponentially fast.

A *truncated cusp* is a portion $N = M \times [a, +\infty)$, bounded by the euclidean manifold $M \times \{a\}$: note that the boundary $\partial N$ is euclidean and not geodesic. The volume of a truncated cusp is particularly simple.

PROPOSITION 2.4. *Let $N$ be truncated cusp. We have*

$$\text{Vol}\big(N\big) = \frac{\text{Vol}(\partial N)}{n-1}.$$

PROOF. It follows from Proposition 1.29. $\square$

Some hyperbolic manifold may contain a portion isometric to a truncated cusp: in that case we will call it simply a cusp.

EXAMPLE 2.5. In dimension $n = 2$ there is only one cusp. The group $\Gamma < \mathrm{Isom}(\mathbb{R})$ is the infinite cyclic group generated by a translation $x \mapsto x + b$ and up to conjugating in $\mathrm{Isom}(\mathbb{H}^2)$ we may take $b = 1$. The cusp is diffeomorphic to $S^1 \times \mathbb{R}$, the circle $S^1 \times \{u\}$ having length $e^{-2u}$. A truncated cusp (but not the whole cusp!) embeds in $\mathbb{R}^3$ as shown in Fig. 13.

Pick $p \in \mathbb{H}^2$. Note that a cusp and $\mathbb{H}^2 \setminus \{p\}$ are diffeomorphic and both hyperbolic. However, they are not isometric because the cusp is complete while $\mathbb{H}^2 \setminus \{p\}$ is not.

The injectivity radius of a cusp is zero for general reasons.

PROPOSITION 2.6. *Let $\Gamma < \mathrm{Isom}(\mathbb{H}^n)$ be a discrete group acting freely on $\mathbb{H}^n$. If $\Gamma$ contains parabolics then* $\mathrm{inj}\big(\mathbb{H}^n/_\Gamma\big) = 0$.

PROOF. A parabolic element $\gamma \in \Gamma$ has minimum displacement zero. That is, there is a sequence $x_i \in \mathbb{H}^n$ such that $d_i = d(x_i, \varphi(x_i)) \to 0$. Since $\mathrm{inj}_{\pi(x_i)}\big(\mathbb{H}^n/_\Gamma\big) < d_i$ we are done. $\square$

COROLLARY 2.7. *If $M = \mathbb{H}^n/\Gamma$ is a compact hyperbolic manifold, every non-trivial element in $\Gamma$ is hyperbolic.*

PROOF. The injectivity radius of a compact manifold is positive. $\square$

**2.3. Closed geodesics.** A *closed curve* in a manifold $M$ is a differentiable map $\alpha\colon S^1 \to M$. A (possibly closed) curve is *simple* if it is injective.

We consider $S^1$ as a subset of $\mathbb{C}$. A *closed geodesic* in a riemannian manifold $M$ is a smooth map $\alpha\colon S^1 \to M$ whose lift $\alpha \circ \pi\colon \mathbb{R} \to M$ along the universal covering $\pi(t) = e^{it}$ is a non-constant geodesic. Two closed geodesics $\alpha_1, \alpha_2$ that differ only by a rotation, *i.e.* such that $\alpha_1(z) = \alpha_2(ze^{it})$ for some fixed $t \in \mathbb{R}$, are implicitly considered equivalent. By substituting $\alpha(z)$ with $\overline{\alpha}(z) = \alpha(\bar{z})$ we change the *orientation* of the closed geodesic.

PROPOSITION 2.8. *Let $\gamma$ be a closed geodesic in a riemannian manifold $M$. Exactly one of the following holds:*

(1) *the curve $\gamma$ is simple,*
(2) *the curve $\gamma$ self-intersects transversely in finitely many points,*
(3) *the curve $\gamma$ wraps along a curve of type (1) or (2) some $k \geqslant 2$ times.*

PROOF. If the geodesic is not simple, it self-intersects. If it self-intersects only with distinct tangents, then (2) holds. Otherwise (1) holds. $\square$

The natural number $k$ in (3) is the *multiplicity* of the closed geodesic. A closed geodesic $\gamma$ of multiplicity $k$ is of type $\gamma(e^{it}) = \eta(e^{kit})$ for some geodesic $\eta$ of type (1) or (2).

EXERCISE 2.9. A closed geodesic on a riemannian manifold $M$ is determined by its support, its orientation, and its multiplicity.

**2.4. Closed geodesics in a hyperbolic manifold.** Closed geodesics in hyperbolic manifolds have a particularly nice behavior.

Let $X, Y$ be topological spaces: as usual we indicate by $[X, Y]$ the spaces of all continuous maps from $X$ to $Y$ seen up to homotopy. Let $X$ be path-connected. There is a natural map $\pi_1(X, x_0) \to [S^1, X]$, and the following is a standard exercise in topology.

Exercise 2.10. The map induces a bijection between the conjugacy classes in $\pi_1(X, x_0)$ and $[S^1, X]$.

A simple closed curve in $X$ is *homotopically trivial* if it is homotopic to a constant. As a corollary, a simple closed curve $\gamma$ is homotopically trivial if and only if it represents the trivial element in $\pi_1(X, \gamma(1))$.

On $M = \mathbb{H}^n/_\Gamma$ we get the correspondence

$$\big\{\text{conjugacy classes in } \Gamma\big\} \longleftrightarrow [S^1, M].$$

Two conjugate elements in $\Gamma$ are of the same type (trivial, parabolic, or hyperbolic) and have the same minimum displacement. Therefore every element in $[S^1, M]$ has a well-defined type and minimum displacement.

Remark 2.11. The correspondence may be described directly as follows: given $\varphi \in \Gamma$, pick any point $x \in \mathbb{H}^n$ and any arc connecting $x$ with $\varphi(x)$ and project it to get a closed curve in $M$ and hence an element in $[S^1, M]$.

Proposition 2.12. *Let $M$ be a complete hyperbolic manifold. A hyperbolic element of $[S^1, M]$ is represented by a unique closed geodesic, of length $d$ equal to its minimum displacement. Trivial and parabolic elements are not represented by closed geodesics.*

Proof. Take $M = \mathbb{H}^n/_\Gamma$. A hyperbolic isometry $\varphi \in \Gamma$ has a unique invariant geodesic in $\mathbb{H}^n$, its axis, which projects on a closed geodesic of length $d$. Conjugate isometries determine the same geodesic in $M$.

On the other hand, a closed geodesic in $M$ lifts to a segment connecting two distinct points $x_0$ and $\varphi(x_0)$ for some $\varphi \in \Gamma$ which preserves the line passing through $x_0$ and $\varphi(x_0)$: since $\varphi$ fixes a line, it is hyperbolic. □

We get a bijection

$$\big\{\text{hyperbolic conjugacy classes in } \Gamma\big\} \longleftrightarrow \big\{\text{closed geodesics in } M\big\}.$$

Corollary 2.13. *Let $M$ be a compact hyperbolic manifold. Every non-trivial element in $[S^1, M]$ is represented by a unique closed geodesic.*

Proof. Since $M$ is compact there are no parabolics. □

Proposition 2.14. *Let $M$ be a compact hyperbolic manifold. For every $L > 0$ there are finitely many closed geodesics shorter than $L$.*

Proof. Suppose there are infinitely many. Since $M = \mathbb{H}^n/_\Gamma$ is compact it has finite diameter $D$ and hence we can fix a basepoint $x_0 \in M$ and connect $x_0$ to these geodesics with arcs shorter than $D$. We use this arcs to homotope the geodesics into loops based at $x_0$ of length bounded by $L+2D$, and lift the loops to arcs in $\mathbb{H}^n$ starting from some basepoint $\widetilde{x_0} \in \mathbb{H}^n$.

If two such arcs end at the same point, then the corresponding initial geodesics are homotopic: this is excluded, hence these endpoints are all distinct. Therefore the orbit of $\widetilde{x_0}$ contains infinitely many points in the ball $B(\widetilde{x_0}, L+2D)$, a contradiction because the orbit is discrete. □

The lengths of the closed geodesics in a compact $M$ form a discrete subset of $\mathbb{R}$ called the *geodesic spectrum* of $M$. Simple closed geodesics have nice small neighborhoods, recall the notion of $R$-tube from Section 2.1.

PROPOSITION 2.15. *The $R$-neighborhood of a simple closed geodesic in a complete hyperbolic manifold is isometric to a $R$-tube, for some $R > 0$.*

PROOF. By compactness of the simple closed geodesic $\gamma$ there is a $R > 0$ such that the $R$-neighborhood of $\gamma$ lifts to disjoint $R$-neighborhoods of its geodesic lifts in $\mathbb{H}^n$. Hence their quotient is a $R$-tube. □

**2.5. Isometries that commute or generate discrete groups.** Two isometries of $\mathbb{H}^n$ that commute or generate a discrete group must be of a particular kind. We indicate by $\mathrm{Fix}(\varphi)$ the fixed points in $\overline{\mathbb{H}^n}$ of $\varphi$.

LEMMA 2.16. *Let $\varphi_1, \varphi_2 \in \mathrm{Isom}(\mathbb{H}^n)$ be two hyperbolic or parabolic isometries. If they commute then $\mathrm{Fix}(\varphi_1) = \mathrm{Fix}(\varphi_2)$.*

PROOF. If they commute, the map $\varphi_1$ acts on $\mathrm{Fix}(\varphi_2)$ and viceversa. If $\varphi_2$ is hyperbolic, then $\mathrm{Fix}(\varphi_2) = \{p, q\}$ and $\varphi_1$ fixes the line with endpoints $p$ and $q$, hence is again hyperbolic with $\mathrm{Fix}(\varphi_1) = \{p, q\}$. If $\varphi_1$ and $\varphi_2$ are parabolic then they have the same fixed point $\mathrm{Fix}(\varphi_1) = \mathrm{Fix}(\varphi_2)$. □

LEMMA 2.17. *Let $\varphi_1, \varphi_2 \in \mathrm{Isom}(\mathbb{H}^n)$ be two isometries that generate a discrete group $\Gamma < \mathrm{Isom}(\mathbb{H}^n)$ acting freely on $\mathbb{H}^n$. The following holds:*

(1) *if $\varphi_1$ is hyperbolic and $\varphi_2$ is parabolic, then $\mathrm{Fix}(\varphi_1) \cap \mathrm{Fix}(\varphi_2) = \emptyset$.*
(2) *if $\varphi_1$ and $\varphi_2$ are hyperbolic, then $\mathrm{Fix}(\varphi_1) \cap \mathrm{Fix}(\varphi_2) = \emptyset$ or $\mathrm{Fix}(\varphi_1) = \mathrm{Fix}(\varphi_2)$ and $\varphi_1, \varphi_2$ are powers of the same hyperbolic $\varphi \in \Gamma$.*

PROOF. We prove (1) using the half-space model, supposing by contradiction that $\mathrm{Fix}(\varphi_1) = \{0, \infty\}$ and $\mathrm{Fix}(\varphi_2) = \{\infty\}$.

Proposition 2.15 in Chapter 2 says that

$$\varphi_1(x,t) = \lambda(Ax, t), \quad \varphi_2(x,t) = (A'x + b, t)$$

with $A, A' \in O(n-1)$ and $\lambda \neq 1$. Hence

$$\begin{aligned} \varphi_1^n \circ \varphi_2 \circ \varphi_1^{-n}(x,t) &= \varphi_1^n\big(A'(\lambda^{-n}A^{-n}x) + b, \lambda^{-n}t\big) \\ &= \big(A^nA'A^{-n}x + \lambda^nA^nb, t\big). \end{aligned}$$

Up to interchanging $\varphi_1$ and $\varphi_1^{-1}$ we may suppose $\lambda < 1$ and get

$$\lim_{n\to\infty} \varphi_1^n \circ \varphi_2 \circ \varphi_1^{-n}(0,t) = \lim_{n\to\infty} (\lambda^nA^nb, t) = (0,t).$$

The subgroup $\Gamma$ is not discrete, a contradiction.

We prove (2). Suppose $\mathrm{Fix}(\varphi_1) = \{a, \infty\}$ and $\mathrm{Fix}(\varphi_2) = \{b, \infty\}$. The isometries $\varphi_1$ and $\varphi_2$ permute the horizontal horospheres and

$$[\varphi_1, \varphi_2] = \varphi_2 \circ \varphi_1 \circ \varphi_2^{-1} \circ \varphi_1^{-1} \in \Gamma$$

fixes every horizontal horosphere. Hence the commutator is parabolic or trivial: the first case is excluded by (1), in the second case we have $a = b$ by Lemma 2.16. Both $\varphi_1$ and $\varphi_2$ have the same axis $l$, and since they generate a discrete group $\Gamma$ they are both powers of some hyperbolic $\varphi \in \Gamma$ with that axis. To prove this, note that $\Gamma$ acts effectively on $l$ as a discrete group of translations, hence $\Gamma \cong \mathbb{Z}$. □

COROLLARY 2.18. *Let $\mathbb{H}^n/_\Gamma$ be a complete hyperbolic manifold. The axis in $\mathbb{H}^n$ of two hyperbolic isometries in $\Gamma$ are incident or ultra-parallel (not asymptotically parallel).*

**2.6. Isometry group.** We study here the isometry group $\mathrm{Isom}(M)$ of a hyperbolic manifold $M$. Recall that the *normalizer* $N(H)$ of a subgoup $H < G$ is the set of elements $g \in G$ such that $gH = Hg$. It is the biggest subgroup of $G$ containing $H$ such that $H \triangleleft N(H)$ is a normal subgroup. The isometry group $\mathrm{Isom}(M)$ has an algebraic representation.

PROPOSITION 2.19. *Let $M = \mathbb{H}^n/_\Gamma$ be a hyperbolic manifold. There is a natural isomorphism*

$$\mathrm{Isom}(M) \cong N(\Gamma)/_\Gamma.$$

PROOF. Every isometry $\varphi \colon M \to M$ lifts to an isometry $\widetilde{\varphi}$

$$\begin{tikzcd} \mathbb{H}^n \arrow[r, dotted, "\widetilde{\varphi}"] \arrow[d, "\pi"'] & \mathbb{H}^n \arrow[d, "\pi"] \\ M \arrow[r, "\varphi"'] & M \end{tikzcd}$$

such that $\widetilde{\varphi}\Gamma = \Gamma\widetilde{\varphi}$: hence $\widetilde{\varphi} \in N(\Gamma)$. The lift is uniquely determined up to left- or right- multiplication by elements in $\Gamma$, hence we get a homomorphism

$$\mathrm{Isom}(M) \to N(\Gamma)/_\Gamma$$

which is clearly surjective (every element in $N(\Gamma)$ determines an isometry) and injective (if $\widetilde{\varphi} \in \Gamma$ then $\varphi = \mathrm{id}$). □

Recall that the *centralizer* of $H < G$ is the set of elements $g \in G$ such that $gh = hg$ for all $h$. It is a subgroup of $G$.

EXERCISE 2.20. Let $M = \mathbb{H}^n/_\Gamma$ be a closed hyperbolic manifold. The centralizer of $\Gamma$ is trivial.

**2.7. Outer automorphism group.** The *automorphism group* $\mathrm{Aut}(G)$ of a group $G$ is the group of all the isomorphisms $G \to G$. The *inner automorphisms* are those isomorphisms of type $g \mapsto hgh^{-1}$ for some $h \in G$ and form a normal subgroup $\mathrm{Int}(G) \triangleleft \mathrm{Aut}(G)$. The quotient

$$\mathrm{Out}(G) = \mathrm{Aut}(G)/_{\mathrm{Int}(G)}$$

is called the *outer automorphism group* of $G$.

If $x_0, x_1$ are two points in a path-connected topological space $X$ there is a non-canonical isomorphism $\pi_1(X, x_0) \to \pi_1(X, x_1)$, unique only up to post-composing with an inner automorphism. Therefore there is a *canonical* isomorphism $\mathrm{Out}(\pi_1(X, x_0)) \to \mathrm{Out}(\pi_1(X, x_1))$. Hence $\mathrm{Out}(\pi_1(X))$ depends very mildly on the basepoint.

The group $\mathrm{Omeo}(X)$ of all homeomorphisms of $X$ does not act directly on $\pi_1(X)$ because of the inner-automorphism ambiguity, but we get a natural homomorphism

$$\mathrm{Omeo}(X) \to \mathrm{Out}(\pi_1(X))$$

which is neither injective nor surjective in general.

EXERCISE 2.21. Two homotopic self-homeomorphisms give rise to the same element in $\mathrm{Out}(\pi_1(X))$.

We turn back to our hyperbolic manifolds.

PROPOSITION 2.22. *If $M$ is a closed hyperbolic manifold the map*

$$\mathrm{Isom}(M) \to \mathrm{Out}(\pi_1(M))$$

*is injective.*

PROOF. Set $M = \mathbb{H}^n/_\Gamma$, identify $\Gamma$ with $\pi_1(M)$ and $\mathrm{Isom}(M)$ with $N(\Gamma)/_\Gamma$. With these identifications the map

$$N(\Gamma)/_\Gamma \to \mathrm{Out}(\Gamma)$$

is just the conjugacy action that sends $h \in N(\Gamma)$ to the automorphism $g \mapsto h^{-1}gh$ of $\Gamma$. This is an inner automorphism if and only if there is $f \in \Gamma$ such that $h^{-1}gh = f^{-1}gf$ for all $g \in \Gamma$, that is if $hf^{-1}$ commutes with $g$ for all $g \in \Gamma$. Exercise 2.20 shows that $h = f \in \Gamma$. The map is injective. $\square$

COROLLARY 2.23. *The isometry group* $\mathrm{Isom}(M)$ *of a closed hyperbolic manifold is finite. Two distinct isometries are not homotopically equivalent.*

PROOF. Distinct isometries have distinct images in $\mathrm{Out}(\pi_1(M))$ and are hence non-homotopic by Exercise 2.21.

The topological group $\mathrm{Isom}(M)$ is compact because $M$ is compact. To show that it is finite it suffices to prove that it is discrete: suppose that a sequence of isometries $\varphi_i$ converges to some isometry $\varphi$. By composing with $\varphi^{-1}$ we may suppose that $\varphi = \mathrm{id}$. Hence for any $\varepsilon > 0$ there is a $i_0$ such that $\varphi_i$ moves the point at most $\varepsilon$ for all $i > i_0$.

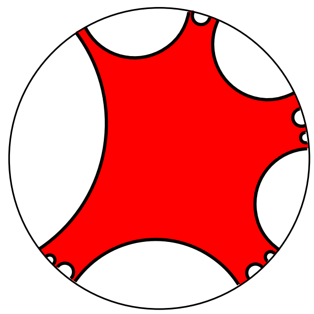


FIGURE 14. An intersection of (possibly infinitely many) half-planes. The universal cover of a hyperbolic surface with boundary is isometric to such an object.

Pick $\varepsilon < \mathrm{inj} M$: every pair of points $x$ and $\varphi_i(x)$ is connected by a unique geodesic $\gamma_x$ of length $d(x, \varphi_i(x))$. The geodesics $\gamma_x$ as $x \in M$ may be used to define a homotopy between $\varphi_i$ and id: a contradiction. □

**2.8. Hyperbolic manifolds with boundary.** The boundary version of hyperbolic (elliptic, flat) manifolds is easy to formulate.

DEFINITION 2.24. A hyperbolic (elliptic, flat) manifold $M$ *with geodesic boundary* is a riemannian manifold with boundary such that every point has an open neighborhood isometric to an open set in a half-space in $\mathbb{H}^n$ ($S^n$, $\mathbb{R}^n$).

The boundary $\partial M$ of a hyperbolic (elliptic, flat) $n$-manifold with geodesic boundary is a hyperbolic (elliptic, flat) $(n-1)$-manifold without boundary. Theorem 1.2 extends appropriately to this context.

THEOREM 2.25. *A complete simply connected hyperbolic $n$-manifold $M$ with geodesic boundary is the intersection of half-spaces in $\mathbb{H}^n$ with disjoint boundaries.*

PROOF. The proof is the same with a little variation: we construct a developing map $D\colon M \to \mathbb{H}^n$, which is a covering onto its image $D(M)$. A submanifold $D(M) \subset \mathbb{H}^n$ with geodesic boundary is necessarily the intersection of half-spaces. In particular $D(M)$ is simply connected and hence the covering $D$ is an isometry. □

An example is drawn in Fig. 14.

COROLLARY 2.26. *A connected hyperbolic $n$-manifold $M$ with geodesic boundary is contained in a unique connected hyperbolic $n$-manifold $N$ without boundary.*

PROOF. The proof of Proposition 1.8 applies and shows that $M = \widetilde{M}/_{\Gamma}$ where $\widetilde{M}$ is simply-connected and hence an intersection of half-spaces in $\mathbb{H}^n$, and $\Gamma$ is a group of isometries of $\widetilde{M}$.

Every local isometry in $\mathbb{H}^n$ extends to a global isometry and hence $\Gamma < \mathrm{Isom}(\mathbb{H}^n)$. Hence $N = \mathbb{H}^n/_{\Gamma}$ contains naturally $M$. □

Two hyperbolic manifolds with geodesic boundary can sometimes be glued along their boundary. Let $M$ and $N$ be hyperbolic manifolds with geodesic boundary and $\psi\colon \partial M \to \partial N$ an isometry. Let $M \cup_\psi N$ be the topological space obtained by quotienting the disjoint union $M \sqcup N$ by the equivalence relation that identifies $x$ to $\psi(x)$ for all $x \in \partial M$.

PROPOSITION 2.27. *The space $M \cup_\psi N$ has a natural structure of hyperbolic manifold.*

PROOF. Let $y$ be the result of gluing $x$ to $\psi(x)$. The point $y$ has two neighborhoods on both sides, both isometric to a hyperbolic half-disc of small radius $\varepsilon$. The isometry $\psi$ tells how to glue these two half-discs to a honest hyperbolic disc, which induces a hyperbolic metric near $x$. □

CHAPTER 4

# Surfaces

## 1. Geometrization of surfaces

Mettere anche Jordan liscio?

**1.1. Classification of surfaces.** We prove here the following theorem.

THEOREM 1.1 (Classification of surfaces). *A compact, connected, orientable surface is diffeomorphic to the surface $S_g$ obtained by attaching $g$ handles to the sphere $S^2$ as shown in Fig. 1-(left).*

We extend our investigation to a larger interesting class of surfaces.

DEFINITION 1.2. Let $g, b, p \geqslant 0$ be three natural numbers. The *surface of finite type* $S_{g,b,p}$ is the surface obtained from $S_g$ by removing the interior of $b$ disjoint discs and $p$ points.

See Fig. 2. We say that $S_{g,b,p}$ has genus $g$, has $b$ boundary components, and $p$ punctures. Its Euler characteristic is

$$\chi(S_{g,b,p}) = 2 - 2g - b - p.$$

FIGURE 1. A sphere with 3 handles may be represented in various ways.

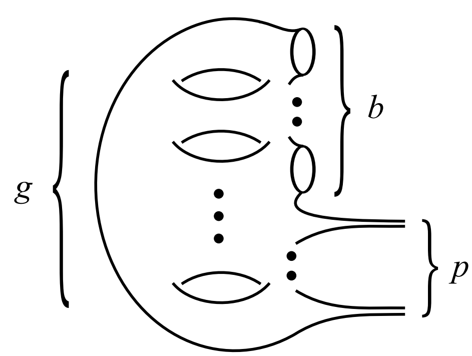



FIGURE 2. The finite type surface $S_{g,b,p}$.

We also use the notation $S_{g,b}$ to indicate $S_{g,b,0}$.

**1.2. Gauss-Bonnet theorem.** A riemannian surface $S$ is of course a surface equipped with a metric tensor. For instance, every surface in $\mathbb{R}^3$ like those shown in Fig. 1 has a metric tensor induced from the euclidean one on $\mathbb{R}^3$. Every point $p \in S$ has a gaussian curvature $K_p \in \mathbb{R}$ which varies continuously in $p \in S$. The famous Gauss-Bonnet theorem connects the curvature to the Euler characteristic of $S$:

THEOREM 1.3 (Gauss-Bonnet). *Let $S$ be a compact surface, possibly with geodesic boundary. We have*

$$\int_S K_p = 2\pi\chi(S).$$

COROLLARY 1.4. *Let $S$ be a compact riemannian orientable surface, possibly with geodesic boundary, with constant curvature $K = -1, 0$ or $1$.*

- *if $K = 1$ then $S$ is a sphere or a disc,*
- *if $K = 0$ then $S$ is an annulus or a torus,*
- *if $K < 0$ then $\chi(S) < 0$.*

*When $K = \pm 1$ we get* $\mathrm{Area}(S) = 2\pi\big|\chi(S)\big|$.

We have already constructed an elliptic metric on the sphere or the disc (take a half-sphere), and a flat metric on the torus or the annulus (take $S^1 \times [0, 1]$ with the product metric). We will construct in the next sections a hyperbolic metric for $S$ whenever $\chi(S) < 0$. We cannot do this by finding a nice embedding $S \hookrightarrow \mathbb{R}^3$ in virtue of the following.

PROPOSITION 1.5. *A compact surface without boundary in $\mathbb{R}^3$ has one point with positive curvature.*

PROOF. Consider the closed discs $D(0, R)$ of radius $R$. Let $R$ be the minimum value such that $S \subset D(0, R)$. The sphere $\partial D(0, R)$ is tangent to $S$ in some point $p$, hence all directional curvatures of $S$ in $p$ are bigger or equal than the directional curvatures $\frac{1}{R}$ on the sphere: hence $K_p > \frac{1}{R^2}$. $\square$

**1.3. Hyperbolic pair-of-pants.** We prove here the following. A *pair-of-pants* is the surface $S_{0,3}$.

PROPOSITION 1.6. *Given three real numbers $a, b, c \geqslant 0$ there is (up to isometries) a unique hyperbolic pair-of-pants with geodesic boundary, with boundary curves of length $a$, $b$, and $c$.*

When some length in $a, b, c$ is zero, we mean that the geodesic boundary is actually a puncture, hence $S_{0,3}$ degenerates to a punctured annulus $S_{0,2,1}$, a twice punctured disc $S_{0,1,2}$, or a thrice-punctured sphere $S_{0,0,3}$: see Fig. 3. We require the hyperbolic metric to be complete near the puncture.

Spiegare bene come funzia!

To prove this proposition we construct some right-angled hexagons as in Fig. 5-(left). Three *alternate sides* on a hexagon are three pairwise disjoint sides, like the $a, b, c$ shown in the figure. A *degenerate* hexagon is one where the length of some non-adjacent sides is zero as in Fig. 4.

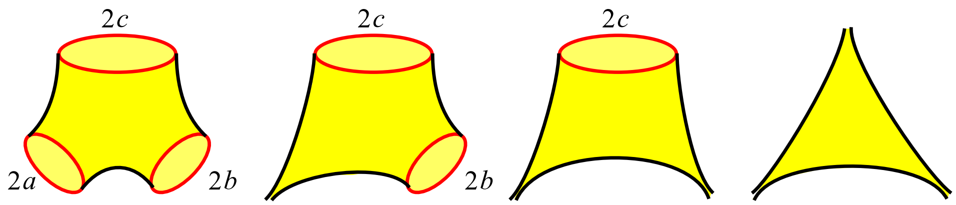



FIGURE 3. A pair-of-pants, an annulus with one puncture, a disc with two puncutres, and a thrice-punctured sphere.

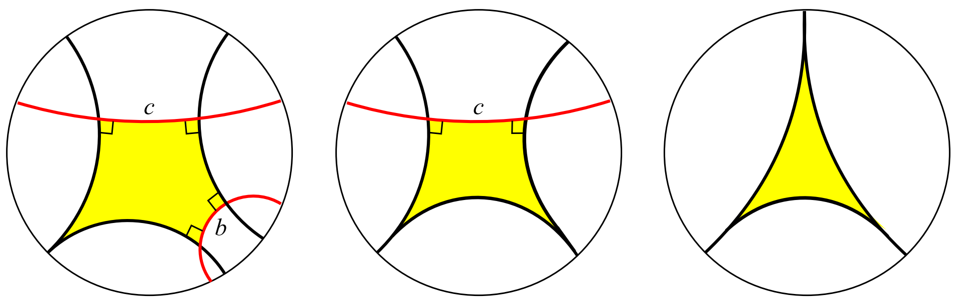



FIGURE 4. A right-angled hexagon with parameters $a, b, c \geqslant 0$ degenerates to a pentagon, quadrilateral, or triangle with ideal vertices if one, two, or three parameters are zero.

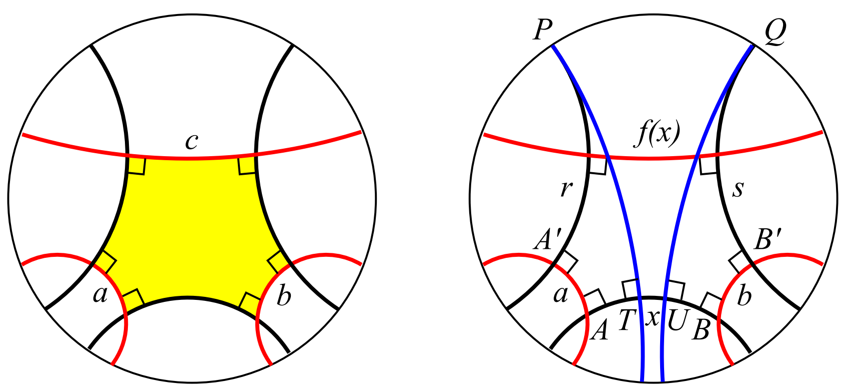



FIGURE 5. A right-angled hexagon with alternate sides of length $a$, $b$ e $c$ (left) and its construction (right), which goes as follows: take a line with two arbitrary points $A$ and $B$ in it (bottom black). Draw the perpendiculars from $A$ and $B$ (red). At distances $a$ and $b$ we find two points $A'$ e $B'$ and we draw again two perpendiculars (black) $r$ and $s$, with some points at infinity $P$ and $Q$. Draw the (unique) perpendiculars to the initial line pointing to $P$ and $Q$ (blue): they determine two points $T$ and $U$. Note that $AT$ and $UB$ have some fixed length depending only on $a$ and $b$. We can vary the parameter $x = TU$: if $x > 0$ the blue lines are ultra-parallel and there is a unique segment orthogonal to both of some length $f(x)$.

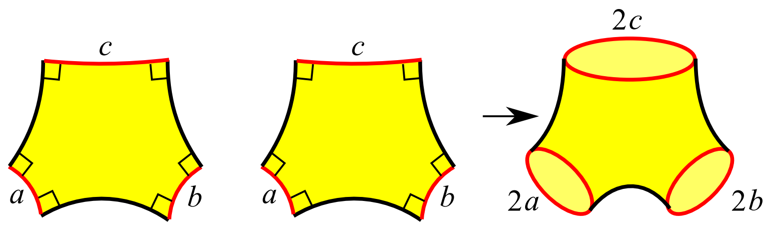



FIGURE 6. By gluing two identical right-angled hexagons along their black sides we get a hyperbolic pair-of-pants with geodesic boundary.

LEMMA 1.7. *Given three real numbers $a, b, c \geqslant 0$ there exists (up to isometries) a unique (possibly degenerate) hyperbolic right-angled hexagon with three alternate sides of length $a$, $b$, and $c$.*

PROOF. Suppose $a, b > 0$. The construction of the hexagon is depicted in Fig. 5-(right). If $x = 0$ the blue lines coincide, hence $P = Q$ and $f(0) = 0$. The function $f\colon [0, +\infty) \to [0, +\infty)$ is continuous, strictly monotonic, and with $\lim_{x\to\infty} f(x) = \infty$: therefore there is precisely one $x$ such that $f(x) = c$.

If two parameters are zero, say $a = b = 0$, then a simpler construction works: take a segment of length $c$ as in Fig. 4-(center), draw the perpendiculars at their endpoints, and a line connecting the endpoints of these.

If $a = b = c = 0$, use the half-space model and recall that $\mathbb{P}\mathrm{SL}_2(\mathbb{R})$ acts transitively on the unordered triples of points in $\partial\mathbb{H}^2$, and hence it acts transitively on ideal triangles. $\square$

The most degenerate case is so important that we single it out.

COROLLARY 1.8. *Ideal triangles in $\mathbb{H}^2$ are all isometric.*

By gluing two identical (possibly degenerate) hexagons as in Fig. 6 we construct a (possibly degenerate) hyperbolic pair-of-pants whose geodesic boundary consists of three curves of length $2a$, $2b$, and $2c$.

PROOF OF PROPOSITION 1.6. We have proved their existence, we turn to their uniqueness up to isometry. Let $P$ be a pair-of-pants whose geodesic boundaries $C_1, C_2, C_3$ have length $2a, 2b, 2c > 0$.

Since $C_1$ and $C_2$ are compact, there are points $x_1 \in C_1$ and $x_2 \in C_2$ at minimum distance $d(x_1, x_2)$. By Hopf-Rinow there is a geodesic $\gamma_3$ connecting them with $L(\gamma) = d(x_1, x_2)$. This geodesic is simple and orthogonal to $C_1$ and $C_2$: if not, some other curve connecting $x_1$ and $x_2$ would be shorter. We construct analogously two orthogeodesics $\gamma_1$ and $\gamma_2$ connecting $C_2$ to $C_3$, and $C_3$ to $C_1$ having minimal length.

The fact that $\gamma_1, \gamma_2, \gamma_3$ have minimal length easily implies that they are disjoint (if they intersect, find some shorter curve). The three geodesics subdivide $P$ into two hexagons, with alternate sides of length $L(\gamma_1)$, $L(\gamma_2)$, and $L(\gamma_3)$: by Lemma 1.7 they are isometric, and hence the three other alternating sides also have the same length $a, b$ and $c$. Hexagons are unique

Fare caso degenere!

up to isometry and hence the pair-of-pants too. $\square$

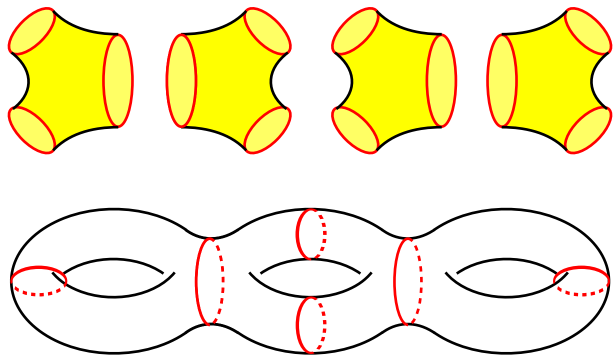


FIGURE 7. Every surface of finite type with $\chi < 0$ decomposes into pair-of-pants. We show here a decomposition of $S_3$.

**1.4. Hyperbolic surfaces.** The pair-of-pants can be used as building blocks to construct all finite type surfaces with $\chi < 0$.

PROPOSITION 1.9. *If $\chi(S_{g,b,b}) < 0$ then $S_{g,b,p}$ decomposes into $-\chi(S_{g,b,p})$ (possibly degenerate) pairs-of-pants.*

PROOF. If $b + p = 0$ then $g \geqslant 2$ and the surfaces decomposes easily in many ways, see for instance Fig. 7. If $\chi < -1$, a decomposition for $S_{g,b,p}$ may be obtained from a decomposition of $S_{g,b-1,p}$ or $S_{g,b,p-1}$ by inserting one more (possibly degenerate) pair-of-pants between two adjacent pair-of-pants. If $\chi = -1$ then the surface is either a pair-of-pants, or a torus with a puncture or boundary component, which is in turn obtained by glueing two boundary components of a pair-of-pants. □

COROLLARY 1.10. *If $\chi(S_{g,b,p}) < 0$ then $S_{g,b,p}$ admits a complete hyperbolic metric with geodesic boundaries of arbitrary length.*

PROOF. Decompose $S_{g,b,p}$ in pair-of-pants. Assign an arbitrary length to all the closed curves of the decomposition (the 6 red curves shown in Fig. 7) and hence give each pair-of-pants the hyperbolic metric determined by the three assigned boundary lengths. Everything glues to a hyperbolic metric for $S_{g,b,p}$. □

EXERCISE 1.11. Prove that a (possibly degenerate) hyperbolic pair-of-pants with geodesic boundary has area $2\pi$, thus confirming Gauss-Bonnet.

**1.5. Riemann surfaces.** Recall that a *Riemann surface* is a surface equipped with a *complex structure*, that is an atlas to open sets of $\mathbb{C}$ whose transition functions are biolomorphisms. Although defined in completely different ways, there is a dictionary translating Riemann surfaces into hyperbolic (or flat, elliptic) surfaces and viceversa. The existence of such a dictionary is quite unexpected, since a complex structure does not induce a metric tensor on the surface, and viceversa.

We indicate with $D \subset \mathbb{C}$ the open unit disc and recall the core theorem of Riemann uniformization.

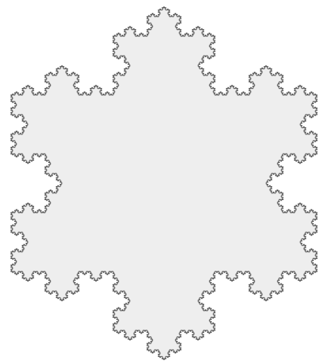


FIGURE 8. La parte interna del fiocco di neve di Koch è biolomorfa al disco aperto $D$.

THEOREM 1.12 (Riemann uniformization). *A simply connected Riemann surface is biolomorphic to $\mathbb{C}$, $\mathbb{CP}^1$, or $D$.*

COROLLARY 1.13. *A simply connected proper open set $U \subset \mathbb{C}$ is biolomorphic to $D$.*

PROOF. By uniformization $U$ is biolomorphic to $D$ or $\mathbb{C}$. Suppose there is a biolomorphism $f\colon \mathbb{C} \to U$: the singularity at infinity is not essential because $f$ is injective, hence $f$ extends to a meromorphic function $f\colon \mathbb{CP}^1 \to U \cup f(\infty) \subset \mathbb{CP}^1$. The image of $f$ is compact: since $U$ is a proper subset this is impossible. □

Note that the boundary of $U$ may be particularly wild as in Fig. 8.

COROLLARY 1.14. *A connected Riemann surface $S$ is biolomorphic to $X/_\Gamma$ where $X \in \{\mathbb{C}, \mathbb{CP}^1, D\}$ and $\Gamma$ is a discrete group of biolomorphisms acting freely on $X$.*

PROOF. The universal covering is biolomorphic to $X$ by Riemann uniformization, and a deck transformation group $\Gamma < \mathrm{Biol}(X)$ is always discrete. □

We now calculate the biolomorphism groups of the three models, recalling that $D$ is also a model for $\mathbb{H}^2$ and $\mathrm{Isom}(D)$ is its isometry group.

PROPOSITION 1.15. *The following identities hold:*

$$\begin{aligned}\mathrm{Biol}(\mathbb{CP}^1) &= \mathbb{P}\mathrm{SL}_2(\mathbb{C}),\\ \mathrm{Biol}(\mathbb{C}) &= \{z \mapsto az+b \mid a \neq 0\},\\ \mathrm{Biol}(D) &= \mathrm{Isom}^+(D).\end{aligned}$$

PROOF. The first two equalities are standard consequences of Liouville theorem, and the third is a consequence of Schwarz lemma. Note that $\mathrm{Isom}^+(D)$ consists precisely of the Möbius transformations that fix $D$, since

$\mathrm{Isom}^+(H^2) = \mathbb{P}\mathrm{SL}_2(\mathbb{R})$ and the two models are related by a Möbius anti-transformation. If $f\colon D \to D$ is a biolomorphism, up to composing with an isometry we may suppose that $f(0) = 0$, and Schwarz lemma says that $|f(z)| \leqslant |z|$ for all $z \in D$. Since the same result holds for $f^{-1}$, we get $|f(z)| = |z|$ and the map $f$ is a rotation around the origin. $\square$

We have found *a posteriori* that the three models $\mathbb{CP}^1$, $\mathbb{C}$, and $D$ for riemann surfaces are diffeomorphic to the three models $S^2$, $\mathbb{R}^2$, and $\mathbb{H}^2$ for constant curvature surfaces, and that for the most interesting model $D = \mathbb{H}^2$ biolomorphisms and orientation-preserving isometries are the same thing. This remarkable fact provides a dictionary between holomorphic and hyperbolic structures on all orientable surfaces, with very few exceptions.

Corollary 1.16. *Let $S$ be an orientable surface not diffeomorphic to the sphere, the torus, the open disc, or the open annulus. There is a natural bijection*

$$\big\{ \text{ complete hyperbolic structures on } S \big\} \longleftrightarrow \big\{ \text{ complex structures on } S \big\}.$$

Proof. If $S$ is complete hyperbolic then $S = D/_\Gamma$ for some discrete group $\Gamma < \mathrm{Isom}^+(D) = \mathrm{Biol}(D)$ acting freely, and hence inherits a complex structure. Conversely, if $S$ has a complex structure then either $S = D/_\Gamma$ and we are done, or $S = \mathbb{C}/_\Gamma$ or $\mathbb{CP}^1/_\Gamma$ for some discrete subgroup $\Gamma$ of $\mathrm{Biol}(\mathbb{C})$ or $\mathrm{Biol}(\mathbb{CP}^1)$ acting freely. In that case $\Gamma$ is fully understood and $S$ is diffeomorphic to one of the exceptions. $\square$

This dictionary can be used to prove some non-trivial facts in hyperbolic geometry, such as the following.

Proposition 1.17. *Every open subset of $\mathbb{C}$ admits a complete hyperbolic structure.*

Proof. An open set $U \subset \mathbb{C}$ has a natural complex structure and hence a corresponding complete hyperbolic structure unless it is one of the exceptions above. The two exceptions that can arise (open disc and annulus) admit a complete hyperbolic structure anyway ($\mathbb{H}^2$ and a cusp or tube). $\square$

For instance, the the complement $\mathbb{C} \setminus K$ of the Cantor set $K$ admits a complete hyperbolic structure.

Exercise 1.18. Decompose $\mathbb{C} \setminus K$ into infinitely many pair-of-pants and use them to construct a complete hyperbolic metric.

We stress the fact that the dictionary between complex and hyperbolic structure uses the universal cover $D$ and is hence not local.

## 2. Curves on surfaces

In this section we will investigate the closed curves on surfaces and use them to prove some geometric and topological theorem.

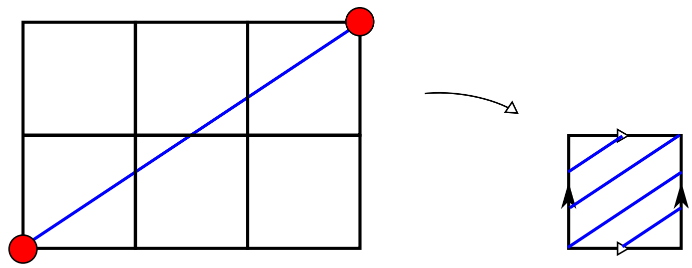


Figure 9. If $(m, n)$ are coprime the line generated by $(m, n) \in \mathbb{R}^2$ projects to a simple closed curve in $T$. Here $(p, q) = (3, 2)$.

Recall that a (possibly closed) curve $\gamma$ on a differentiable manifold $M$ is *regular* if $\gamma'(t) \neq 0$ for all $t$. The image of a regular simple closed curve is a 1-submanifold of $M$ diffeomorphic to $S^1$. All the curves will be tacitly assumed to be regular. Moreover, with a little abuse we will sometimes indicate by $\gamma$ the support of the curve $\gamma$.

**2.1. Simple closed curves on the torus.** We classify the simple closed curve on the torus $T = S^1 \times S^1$. The fundamental group is abelian $\pi(T) = \mathbb{Z} \times \mathbb{Z}$, hence a closed curve is determined up to homotopy by a pair $(m, n) \in \mathbb{Z} \times \mathbb{Z}$.

Proposition 2.1. *If $(m, n)$ are coprime the closed curve is homotopic to a simple one, unique up to isotopy. If $(m, n)$ are not coprime the closed curve is not homotopic to a simple one.*

**2.2. Preliminaries on simple curves.** Two self-diffeomorphisms of $S^1$ are *co-oriented* if they both preserve (or invert) the orientation of $S^1$.

Lemma 2.2. *Two co-oriented self-diffeomorphisms of $S^1$ are isotopic.*

Proof. The lifts $f_0, f_1 \colon \mathbb{R} \to \mathbb{R}$ to universal covers are periodic and monotone, hence $f_t = (1-t)f_0 + tf_1$ also is and descends to $S^1$. $\square$

The classification of surfaces has some non-obvious consequences.

Proposition 2.3. *There are finitely many simple closed curves in $S_g$ up to diffeomorphism.*

Proof. By cutting a simple closed curve along $\gamma$ we get a surface $S'$ with the same Euler characteristic as $S_g$, with and one or two components, and with the boundary oriented as $\gamma$: there are only finitely many diffeomorphism types for $S'$.

We prove that if $\gamma_1$ and $\gamma_2$ give two surfaces $S'_1$ and $S'_2$ of the same type then there is a self-diffeomorphism of $S_g$ sending $\gamma_1$ to $\gamma_2$. By hypothesis there is a diffeomorphism $\varphi \colon S'_1 \to S'_2$ that preserves the boundary orientations. By Lemma 2.2 we may modify $\varphi$ near the boundary so that it extends to a diffeomorphism $\varphi \colon S_g \to S_g$ sending $\gamma_1$ to $\gamma_2$. $\square$

Fare esempio/disegno di curve per $g = 3$

We use hyperbolic geometry to prove some facts on closed curves. Recall that a non-trivial element $g \in G$ in a group is *primitive* if it cannot be written as $g = h^n$ for some $n \geqslant 2$ and some $h \in G$. This condition is conjugacy-invariant, hence the following makes sense.

PROPOSITION 2.4. *Let $\gamma$ be a simple closed curve in $S_g$:*
- *if $\gamma$ is homotopically trivial, it bounds a disc;*
- *if $\gamma$ is not homotopically trivial, it is primitive in $\pi_1(S_g)$.*

PROOF. Let $S'$ be the surface obtained by cutting $S$ along $\gamma$. The surface $S'$ may have one or two components and has the same Euler characteristic of $S$. If one component of $S'$ is a disc, we are done. If $S'$ is an annulus, then $S$ is a torus, $\gamma$ is non-trivial and we are done.

Fare csc nel toro prima.

In all other cases there is a hyperbolic metric on $S$ where $\gamma$ is a geodesic: each component of $S'$ has negative Euler characteristic and hence can be given a hyperbolic structure with boundary curves of length 1; by glue them we get the metric.

This implies that $\gamma$ is not homotopically trivial. If $\gamma = \eta^k$ then $\gamma$ is also homotopic to the geodesic corresponding to $\eta$ run $k$ times: a simple closed curve cannot be homotopic to two distinct geodesics, a contradiction. □

Let the *inverse* $\gamma^*$ of a closed curve $\gamma$ be $\gamma$ run with opposite orientation.

PROPOSITION 2.5. *A homotopically non-trivial closed curve in $S_g$ is not homotopic to its inverse.*

PROOF. If $g = 1$ the curves $\gamma$ and $\gamma^*$ represent distinct elements (and hence conjugacy classes) in $\pi_1(S_1) = \mathbb{Z} \times \mathbb{Z}$. If $g \geqslant 2$, give $S_g$ a hyperbolic metric. The curve $\gamma$ is homotopic to a geodesic $\overline{\gamma}$ and hence $\gamma^*$ is homotopic to its inverse $\overline{\gamma}^*$, which is certainly distinct from $\overline{\gamma}$ as a geodesic. Distinct geodesics are not homotopic. □

Recall that the *R-neighborhood* of an object in a metric space is the set of all points of distance at most $R$ from that object. The $R$-neighborhoods of the lines in $\mathbb{H}^2$ are particularly simple.

PROPOSITION 2.6. *The $R$-neighborhood of a line $l \subset \mathbb{H}^2$ in a conformal model is bounded by two euclidean lines or circle arcs having the same endpoints as $l$ as in Fig. 10.*

PROOF. Put $l$ in the half-space model with endpoints at 0 and $\infty$. A $R$-neighborhood is invariant by the isometry $x \mapsto \lambda x$ and is hence a cone as in the figure. The other cases follow because isometries send lines and circles to lines and circles. □

We will use the $R$-neighborhoods to prove the following.

PROPOSITION 2.7. *Let $S_g$ have a hyperbolic metric. A simple closed curve is homotopic to a simple closed geodesic.*

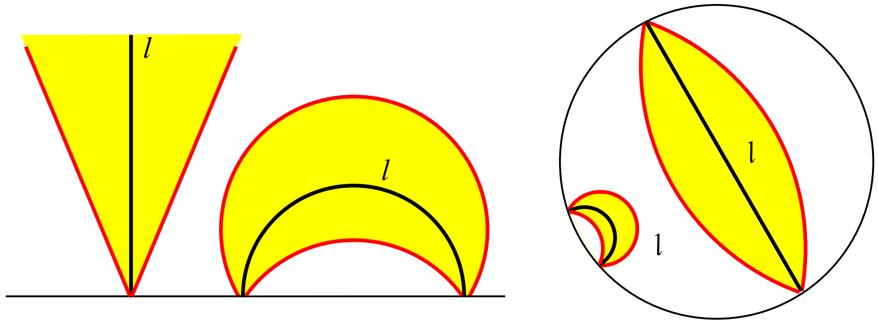



FIGURE 10. The $R$-neighborhood of a geodesic $l$ in the half-plane (left) and disc (right) models.

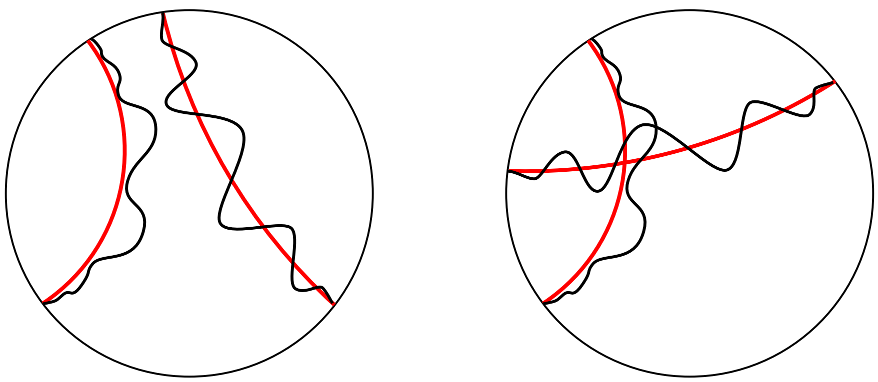


FIGURE 11. The lifts of $\gamma$ (black arcs) and its geodesic representative (red lines) have the same endpoints in $\partial\mathbb{H}^n$ (left). If two lines intersect, the corresponding arcs do (right).

PROOF. A simple curve $\gamma$ in $S_g = \mathbb{H}^2/_\Gamma$ is homotopic to a closed geodesic $\bar{\gamma}$, and we need to prove that $\bar{\gamma}$ is simple. The counterimage of $\gamma$ in $\mathbb{H}^2$ consists of disjoint simple *arcs*, while the counterimage of $\overline{\gamma}$ consists of *lines*: we prove that these lines are also disjoint.

The homotopy between $\gamma$ and $\overline{\gamma}$ lifts to a homotopy between the arcs and the lines. The homotopy between $\gamma$ and $\overline{\gamma}$ has compact support, hence there is a $R > 0$ such that every point is moved to some distance smaller than $R$. Therefore the $R$-neighborhood of each line contains entirely an arc as in Fig. 11-(sinistra).

Fare dei colori rossi più tenui per la stampa BN

This shows that lines and arcs have the same endpoints. If two lines intersects, their endpoints are linked in the circle $\partial\mathbb{H}^2$ and hence also the corresponding arcs intersect, see Fig. 11-(right): a contradiction.

Since the lifts of $\overline{\gamma}$ do not intersect, the geodesic $\overline{\gamma}$ is either simple or wraps some times along a simple geodesic, but the second possibility is excluded by Proposition 2.4. $\square$

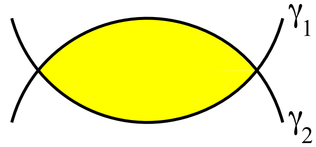



FIGURE 12. Two curves $\gamma_1$ and $\gamma_2$ are in minimal position if and only if they do not create bigons.

**2.3. Intersections of simple closed curves.** We want to study the intersections of simple closed curves. A homotopy class $g \in [S^1, M]$ is *simple* if it is represented by a simple closed curve.

DEFINITION 2.8. Let $g_1, g_2 \in [S^1, S]$ be two simple homotopy classes on an orientable surface $S$. Their *geometric intersectioon* is the number

$$i(g_1, g_2) = \min\left\{\#(\gamma_1 \cap \gamma_2) \;\middle|\; \gamma_1 \in g_1, \gamma_2 \in g_2\right\}$$

where $\gamma_1$ and $\gamma_2$ vary among the simple closed curves in the classes $g_1$ and $g_2$ that intersect transversely.

We indicate for simplicity as $i(\gamma_1, \gamma_2)$ the geometric intersection of the homotopy classes $[\gamma_1]$ and $[\gamma_2]$. Two simple closed curves $\gamma_1$ and $\gamma_2$ in $S_g$ are *parallel* if they are disjoint and cobound an annulus.

PROPOSITION 2.9. *We have $i(\gamma, \gamma) = 0$ for any simple closed curve $\gamma$.*

PROOF. A tubular neighborhood of $\gamma$ is diffeomorphic to $S^1 \times [-1, 1]$ because $S$ is orientable, hence $\gamma$ has two disjoint parallel representatives $S^1 \times \left\{-\frac{1}{2}\right\}$ and $S^1 \times \left\{\frac{1}{2}\right\}$. $\square$

**2.4. Bigon criterion.** Two simple closed curves $\gamma_1$ and $\gamma_2$ are in *minimal position* if they intersect transversely in $i(\gamma_1, \gamma_2)$ points. The complement of two transverse simple curves is the finite disjoint union of open sets with polygonal boundaries; one such set is a *bigon* if it is a disc with two edges as in Fig. 12. The following criterion is a simple and useful tool to determine the geometric intersection of two curves.

THEOREM 2.10 (Bigon criterion). *Two transverse simple closed curves $\gamma_1, \gamma_2$ in $S_g$ with $g \geqslant 2$ are in minimal position if and only if they do not create bigons.*

PROOF. If $\gamma_1$ and $\gamma_2$ create a bigon, the homotopy described in Fig. 13-(left) transforms $\gamma_1$ and $\gamma_2$ in two curves that intersect in a smaller number of points: hence $\gamma_1$ e $\gamma_2$ are not in minimal position.

Suppose now that $\gamma_1$ and $\gamma_2$ do not form bigons: we need to show that they are in minimal position. If $\gamma_1$ is trivial, it bounds a disc as in Fig. 13-(right). If $\gamma_2$ intersects $\gamma_1$, an *innermost* argument shows that they form a bigon: the curve $\gamma_2$ intersects the disc in arcs, each diving the disc into two parts; if one part contains no other arc it is a bigon, otherwise iterate.

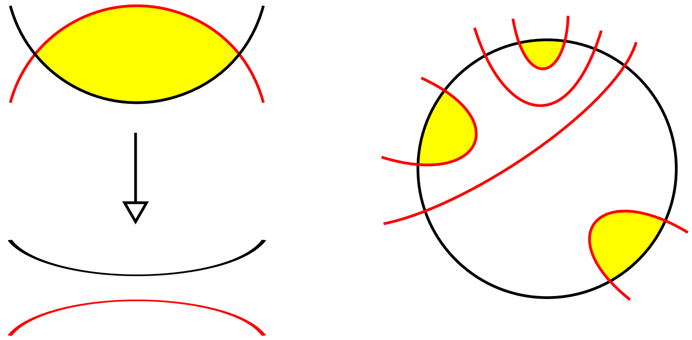


FIGURE 13. A bigon can be eliminated via homotopies (left). If $\gamma_1$ bounds a disc and $\gamma_2$ intersects $\gamma_1$, there is a bigon (right).

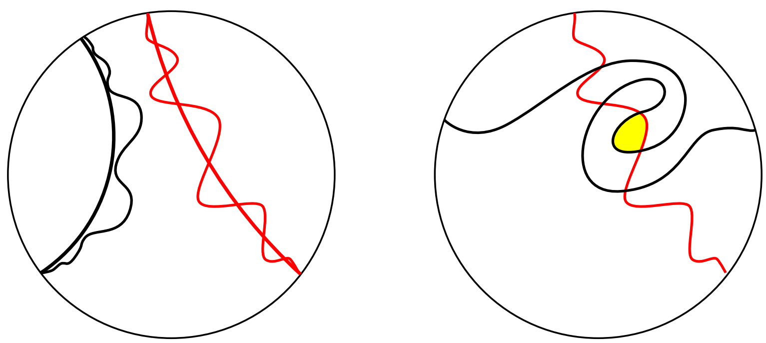


FIGURE 14. The lifts of $\gamma_i$ and $\overline{\gamma_i}$ have distance bounded by $R$ and hence have the same endpoints (left). Two curves that intersect in more than one point form a bigon (right).

Consider the case both $\gamma_1$ and $\gamma_2$ are homotopically non-trivial. Fix an arbitrary hyperbolic metric $S_g = \mathbb{H}^2/_{\Gamma}$. The two curves are homotopic to two simple geodesics $\overline{\gamma_1}$ and $\overline{\gamma_2}$. The lifts of $\gamma_i$ and $\overline{\gamma_i}$ in $\mathbb{H}^2$ are *arcs* and *lines* and there is a $R > 0$ such that every arc lies in the $R$-neighborhood of a line, see the proof of Proposition 2.7. Arcs and lines have the same endpoints at infinity as in Fig. 14-(left).

Two distinct arcs may intersect at most in one point: if they intersect more, an innermost argument shows that they form a bigon $D$ as in Fig. 14-(right), which projects to a bigon in $S$. The last assertion is actually non-immediate: the two vertices of the bigon might go to the same vertex, but this is easily excluded because $S$ is orientable.

We show how to calculate the intersections between $\gamma_1$ and $\gamma_2$ directly on the universal covering. Let $C(\gamma_i) \subset \Gamma$ be the conjugacy class of all hyperbolic transformations corresponding to $\gamma_i$. We know that the lifts of $\overline{\gamma_i}$ are the axis of the hyperbolic transformations in $C(\gamma_i)$.

By Corollary 2.18 from Chapter 3 the axis are either incident or ultra-parallel. Hence two lifts of $\gamma_1$ and $\gamma_2$ intersect (in a single point) if and only if the corresponding lifts of $\overline{\gamma_1}$ and $\overline{\gamma_2}$ intersect (in a single point), and this holds if and only if the endpoints are linked in $\partial\mathbb{H}^2$. Let $\pi\colon \mathbb{H}^2 \to \mathbb{H}^2/_\Gamma$ be the projection. We have established two bijective correspondences

$$\pi^{-1}(\gamma_1) \cap \pi^{-1}(\gamma_2) \longleftrightarrow \pi^{-1}(\overline{\gamma_1}) \cap \pi^{-1}(\overline{\gamma_2}) \longleftrightarrow X$$

with

$$X = \big\{(\varphi_1, \varphi_2) \in C(\gamma_1) \times C(\gamma_2) \;\big|\; \mathrm{Fix}(\varphi_1) \text{ and } \mathrm{Fix}(\varphi_2) \text{ are linked}\big\}.$$

The bijective correspondences are $\Gamma$-equivariant. By quotienting by the acion of $\Gamma$ we find

$$\gamma_1 \cap \gamma_2 \longleftrightarrow \overline{\gamma_1} \cap \overline{\gamma_2} \longleftrightarrow X/_\Gamma.$$

The cardinality $N$ of $X/_\Gamma$ depends only on the homotopy type of $\gamma_1$ and $\gamma_2$. Therefore two curves homotopic to $\gamma_1$ and $\gamma_2$ will have at least these $N$ intersections. Hence $\gamma_1$ and $\gamma_2$ are in minimal position. □

In the last part of the proof we have implicitly proved the following.

Corollary 2.11. *Let $S_g$ be a hyperbolic surface. Two simple closed geodesics with distinct supports are always in minimal position.*

Proof. Two geodesics do not create bigons: if they do, the bigon lifts to a bigon between two lines in $\mathbb{H}^2$, but lines may intersect at most once. □

Corollary 2.12. *If two closed curves $\gamma, \eta$ intersect transversely in one point, we have $i(\gamma, \eta) = 1$. In particular, they are homotopically non-trivial.*

Exercise 2.13. If $\gamma$ is not homotopically trivial there is $\eta$ such that $i(\gamma, \eta) > 0$.

Hint. Use Proposition 2.3 to transform $\gamma$ into a comfortable curve and draw an $\eta$ which intersects $\gamma$ in at most 2 points without bigons. □

Exercise 2.14. Prove the bigon crierion for the torus. Deduce that

$$i\big((p,q),(r,s)\big) = \left|\det\begin{pmatrix} p & r \\ q & s \end{pmatrix}\right|.$$

**2.5. Homotopy and isotopy of curves.** We show here that two simple closed curves are homotopic if and only if they are ambient isotopic. We start with a particular case.

Lemma 2.15. *Let $\gamma_1$ and $\gamma_2$ be two homotopically non-trivial curves in $S_g$. If they are disjoint and homotopically equivalent, they are parallel.*

Proof. Cut $S_g$ along $\gamma_1 \cup \gamma_2$. We don't obtain discs because the curves are homotopically non-trivial and if we obtain an annulus the two curves are parallel. In all other cases we obtain surfaces with negative curvature and hence we may assign a hyperbolic metric where both $\gamma_1$ and $\gamma_2$ are geodesics: hence they are not homotopic. □

We now turn to the general case.

Proposition 2.16 (Homotopy implies isotopy). *Two homotopically non-trivial simple closed curves in $S_g$ are homotopically equivalent if and only if they are ambiently isotopic.*

Proof. Recall that isotopy implies ambient isotopy here because $S^1$ is compact: hence we only need to prove that they are isotopic.

Let $\gamma_1$ and $\gamma_2$ be the two curves. Up to perturbing with a little isotopy we may suppose they intersect transversely. Since $i(\gamma_1, \gamma_2) = i(\gamma_1, \gamma_1) = 0$ the two curves are disjoint or form a bigon. If they form a bigon, we can eliminate it via isotopies as in Fig. 13-(left) and after finitely many steps we get two disjoint curves.

The curves $\gamma_1$ and $\gamma_2$ are parallel by Lemma 2.15, and we use the annulus they cobound to move $\gamma_2$ isotopically over $\gamma_1$. The two curves now have the same support and the same orientation by Proposition 2.5: by Lemma 2.2 they are isotopic. □

This fact is not true in higher dimensions: two homotopically equivalent simple closed curves may be knotted differently and hence are not isotopic; the *knot theory* studies precisely this phenomenon.

Corollary 2.17. *Let $S_g$ be equipped with a hyperbolic metric. A homotopically non-trivial simple closed curve is ambiently-isotopic to a geodesic.*

**2.6. Multicurves.** We introduce the following objects.

Definition 2.18. A *multicurve* $\mu$ in $S_g$ is a finite set of disjoint homotopically non-trivial simple closed curves.

See an example in Fig. 15. A multicurve is *essential* if it has no parallel components. By cutting $S_g$ along an essential multicurve $\mu$ we get finitely many surfaces $S^1, \ldots, S^k$ of negative Euler characteristic. If each such surface is a pair-of-pants, then $\mu$ is called a *pants decomposition.*

Proposition 2.19. *An essential multicurve $\mu$ in $S_g$ with $g \geqslant 2$ has at most $3g-3$ components, and it has $3g-3$ if and only if it is a pants decomposition.*

Proof. By cutting $S_g$ along $\mu$ we get some surfaces $S^1, \ldots, S^k$ of negative Euler characteristic such that $\chi(S_g) = \chi(S^1) + \ldots + \chi(S^k)$. If each $S^i$ is a pair-of-pants then $\chi(S^i) = -1$ and $k = -\chi(S_g) = 2g-2$; the curves are $\frac{3}{2}(2g-2) = 3g-3$ because each boundary curve is counted twice. If some $S^i$ is not a pair-of-pants it can be further subdivided into pair-of-pants. □

Let $m_1$ and $m_2$ be two isotopy classes of multicurves. We define their geometric intersection as

$$i(m_1, m_2) = \min \big\{ \#(\mu_1 \cap \mu_2) \bigm| \mu_1 \in m_1, \mu_2 \in m_2 \big\}.$$

where $\mu_1$ and $\mu_2$ varies among all multicurves in the classes $m_1$ and $m_2$ intersecting transversely. This definition extends the geometric intersection

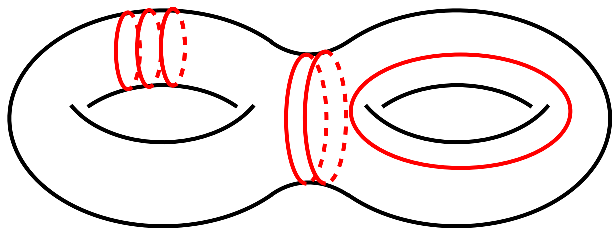


FIGURE 15. A multicurve in a surface of genus two.

of simple closed curves by Proposition 2.16. We still indicate for simplicity by $i(\mu_1, \mu_2)$ the geometric intersection of the classes $[\mu_1]$ and $[\mu_2]$. Two transverse multicurves $\mu_1$ and $\mu_2$ are in *minimal position* if they intersect in $i(\mu_1, \mu_2)$ points: the bigon criterion easily extends to this context.

PROPOSITION 2.20. *Let $\mu_1, \mu_2 \subset S_g$ be transverse multicurves with $g \geqslant 2$. The following equality holds:*

$$i(\mu_1, \mu_2) = \sum_{\substack{\gamma_1 \subset \mu_1 \\ \gamma_2 \subset \mu_2}} i(\gamma_1, \gamma_2)$$

*where the sum is taken on all components $\gamma_1$, $\gamma_2$ of $\mu_1$, $\mu_2$. The multicurves $\mu_1$ and $\mu_2$ are in minimal position if and only if they do not form bigons.*

PROOF. If $\mu_1$ and $\mu_2$ form no bigons, then $\gamma_1$ and $\gamma_2$ have no bigons too, and are therefore in minimal position. This proves the equality and that $\mu_1$ and $\mu_2$ are in minimal position. $\square$

Note again that $i(\mu, \mu) = 0$. We extend Proposition 2.16 to essential multicurves.

PROPOSITION 2.21 (Homotopy implies isotopy). *Let*

$$\mu^1 = \{\gamma_1^1, \ldots, \gamma_n^1\}, \quad \mu^2 = \{\gamma_1^2, \ldots, \gamma_n^2\}$$

*be essential multicurves in $S_g$. If $\gamma_j^1$ is homotopically equivalent to $\gamma_j^2$ for all $j$ then there is an ambient isotopy moving $\mu^1$ to $\mu^2$.*

PROOF. We adapt the proof of Proposition 2.16. Since $i(\gamma_j^1, \gamma_k^2) = i(\gamma_j^1, \gamma_k^1) = 0$ we get $i(\mu_1, \mu_2) = 0$ and after an isotopy $\mu_1 \cap \mu_2 = \emptyset$. The Lemma 2.15 implies that $\gamma_j^1$ and $\gamma_j^2$ are parallel and can be superposed separately for each $j$. $\square$

COROLLARY 2.22. *Let $S_g$ have a hyperbolic metric. An essential multicurve can be isotoped to a (unique) geodesic essential multicurve.*

**2.7. Uniqueness of the minimal position.** We show that the minimal position of two essential multicurve is in fact unique up to isotopy. Given two multicurves $\mu_1, \mu_2$, we indicate by $\mu_1 \cup \mu_2$ the union of their supports.

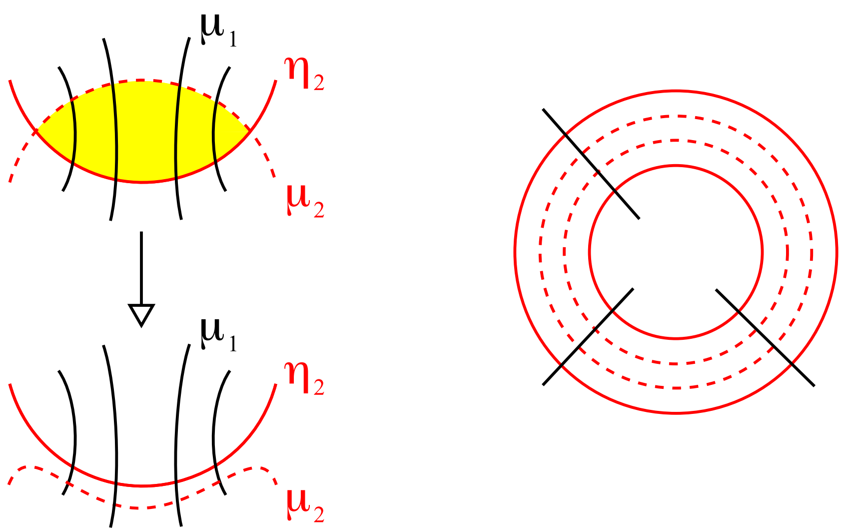



FIGURE 16. A bigon between $\mu_2$ and $\eta_2$ intersects $\mu_1$ into vertical arcs and can be removed via an ambient isotopy that preserves the support of $\mu_1$ (left). Per ogni tipo di omotopia esistono $k$ curve in $\mu_2$ e $k$ in $\eta_2$ e tutte queste intersecano $\mu_1$ in archi come in figura (destra).

PROPOSITION 2.23 (Uniqueness of the minimal position). *Let $(\mu_1, \mu_2)$ and $(\eta_1, \eta_2)$ be two pairs of essential multicurves in minimal position in $S_g$ with $g \geqslant 2$. If $\mu_i$ and $\eta_i$ are isotopic for all $i = 1, 2$, there is an ambient isotopy that carries $\mu_1 \cup \mu_2$ to $\eta_1 \cup \eta_2$.*

PROOF. In what follows we will be concerned only with the support of the multicurves, not their parametrizations.

By hypothesis there is an ambient isotopy carrying $\mu_1$ to $\eta_1$, hence we can suppose $\mu_1 = \eta_1$. We now construct an ambient isotopy that fixes $\mu_1$ and carries $\mu_2$ to $\eta_2$. Up to a little ambient isotopy fixing $\mu_1$, we may suppose that $\mu_2$ and $\eta_2$ intersect transversely.

If $\mu_2 \cap \eta_2 \neq \emptyset$ then $\mu_2$ and $\eta_2$ produce a bigon as in Fig. 16-(left): the multicurve $\mu_1$ intersects the bigon in arcs that join distinct edges as in the figure (otherwise $\mu_1 = \eta_1$ would forf a bigon with $\mu_2$ or $\eta_2$, which is excluded by their minimal position). We can eliminate the bigon by an ambient isotopy that fixes $\mu_1$ as shown in Fig. 16-(left).

We now have $\mu_1 = \eta_1$ and $\mu_2 \cap \eta_2 = \emptyset$. Since $\mu_2$ and $\eta_2$ are disjoint and isotopic, every component of $\mu_2$ is parallel to a component of $\eta_2$ through an annulus which may intersect $\mu_1 = \eta_1$ only by arcs as in Fig. 16-(right). A radial ambient isotopy overlaps the two components keeping $\mu_1$ fixed. By performing this on each component we get $\mu_2 = \eta_2$. □

Cambiare figura e aggiungere caso $\mu_1$ parallelo o cambiare enunciato!

As an example, consider two homotopically non-trivial simple closed curves $\gamma_1, \gamma_2$ in $S_g$. A hyperbolic metric on $S_g$ gives two geodesic representatives $\overline{\gamma_1}$ and $\overline{\gamma_2}$ for them, and the following holds:

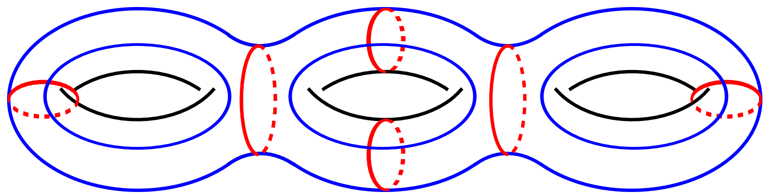


FIGURE 17. Two multicurves (blue and red) in minimal position.

COROLLARY 2.24. *The support $\overline{\gamma_1} \cup \overline{\gamma_2}$ in $S_g$ does not depend (up to ambient isotopy) on the hyperbolic metric chosen.*

PROOF. The geodesics $\overline{\gamma_1}$ and $\overline{\gamma_2}$ coincide or are in minimal position for any metric. $\square$

**2.8. The Alexander trick.** We have proved that homotopy implies homotopy for simple closed curves in $S_g$, now we want to prove an analogous result for diffeomorphisms of $S_g$. We start with the disc.

PROPOSITION 2.25 (Alexander trick). *Two diffeomorphisms $\varphi, \psi\colon D^n \to D^n$ that coincide on $\partial D^n$ are linked by an isotopy that fixes $\partial D^n$ at each time $t$.*

PROOF. We take $f = \varphi \circ \psi^{-1}$ and $\mathrm{id}_{D^n}$ and construct an isotopy that sends $f$ to $\mathrm{id}_{D^n}$ fixing $\partial D^n$. The following function does the job:

$$F(x,t) = \begin{cases} x & \text{if } \|x\| \geqslant t, \\ tf\left(\frac{x}{t}\right) & \text{if } \|x\| \leqslant t. \end{cases}$$

$\square$

When $n = 2$ there is also a *smooth* isotopy that links $\varphi$ and $\psi$, but the proof is more complicate; in higher dimension the existence of a smooth isotopy is an open problem (for $n = 4$) and is often false (for infinite values of $n$, starting from $n = 7$: this fact is connected to the existence of *exotic spheres*, differentiable manifolds that are homeomorphic but not diffeomorphic to $S^n$).

**2.9. Homotopy and isotopy between diffeomorphisms.** We conclude the chapter with this result.

PROPOSITION 2.26 (Homotopy implies isotopy). *Two diffeomorphisms $\varphi, \psi\colon S_g \to S_g$ are hopotopic if and only if they are isotopic.*

PROOF. Fix two multicurves $\mu_1$ and $\mu_2$ as in Fig. 10. The complement of $\mu_1 \cup \mu_2$ consists of polygons with at least 4 sides: since there are no bigons, they are in minimal position.

The multicurves $\varphi(\mu_1)$ and $\psi(\mu_1)$ are homotopic and hence isotopic by Proposition 2.21, and so are $\varphi(\mu_2)$ and $\psi(\mu_2)$. The pairs $\big(\varphi(\mu_1), \varphi(\mu_2)\big)$ and $\big(\psi(\mu_1), \psi(\mu_2)\big)$ are in minimal position (because $\mu_1$ and $\mu_2$ are), hence by

Proposition 2.23 there is an ambient isotopy that carries $\varphi(\mu_1) \cup \varphi(\mu_2)$ to $\psi(\mu_1) \cup \psi(\mu_2)$.

Up to composing with this isotopy we may suppose that $\varphi(\mu_1 \cup \mu_2) = \psi(\mu_1 \cup \mu_2)$. Note that $\mu_1 \cup \mu_2$ is a graph: the maps $\varphi$ and $\psi$ on $\mu_1 \cup \mu_2$ may only differ by different parametrizations on the edges, and with an isotopy these differences disappear. Now $\varphi = \psi$ pointwise on $\mu_1 \cup \mu_2$.

Pick now a polygon $P$ in $S_g \setminus \mu_1 \cup \mu_2$. The maps $\varphi$ and $\psi$ send $P$ to the same polygon $Q$ in $S_g \setminus \varphi(\mu_1 \cup \mu_2)$, because $P$ is determined by the cyclic order of its edges. They coincide on $\partial P$ and by Alexander trick they are linked by an isotopy on $P$. By applying this isotopy on each polygon we obtain an isotopy transforming $\varphi$ into $\psi$. $\square$

CHAPTER 5

# Teichmüller space

We study in this chapter the hyperbolic metrics that can be assigned to a fixed surface $S_g$ of genus $g \geqslant 2$.

## 1. Generalities

We introduce two important definitions.

DEFINITION 1.1. Take $g \geqslant 2$. The *moduli space* of $S_g$ is the set of all the hyperbolic metrics on $S_g$ considered up to isometry.

The *Teichmüller space* $\mathrm{Teich}(S_g)$ of $S_g$ is the set of all hyperbolic metrics on $S_g$ considered up to isometries isotopic to the identity.

At a first sight, the moduli space seems a more natural object to study. It turns out however that the second space is homeomorphic (for some natural topology) to an open ball, while the moduli space is topologically more complicated. It is then better to define and study Teichmüller space first, and then consider the moduli space as a quotient of Teichmüller space.

Ricordiamo che l'operazione di riscalamento di una metrica riemanniana consiste nel sostituire il tensore metrico $g$ con $\lambda g$ per qualche $\lambda > 0$: in geometria iperbolica questa operazione cambia la curvatura se $\lambda \neq 1$ e quindi non è permessa, in geometria piatta invece la curvatura nulla resta invariata e questa operazione ha senso (in altre parole, le omotetie esistono solo nella geometria euclidea e non nelle geometrie iperbolica e sferica). Nel riscalamento le lunghezze variano di un fattore $\sqrt{\lambda}$ e l'area del toro varia di un fattore $\lambda$: a meno di riscalamento si può quindi sempre chiedere che una metrica piatta sul toro abbia area unitaria.

**1.1. Teichmüller space of the torus.** On a torus a flat metric $g$ can be rescaled by any constant $\lambda > 0$ to give another flat metric $\lambda g$. The rescaling changes the lengths by a factor $\sqrt{\lambda}$ and the area by a factor $\lambda$. Up to rescaling, we may ask that the torus have unit area.

The *moduli space* of $T$ is the set of all the flat metrics on $T$ considered up to isometry and rescaling, and the *Teichmüller space* is the set of all flat metrics on $T$ up to isometries isotopic to the identity and rescaling.

The flat metrics on $T$ are easily classified. A *lattice* is a discrete subgroup $\Gamma < \mathbb{R}^2$ isomorphic to $\mathbb{Z}^2$. The group $\mathbb{R}^2$ acts as translations to $\mathbb{R}^2$, hence a lattice is naturally a discrete subgroup of $\mathrm{Isom}^+(\mathbb{R}^2)$.

PROPOSITION 1.2. *A flat torus is isometric to* $\mathbb{R}^2/_\Gamma$ *for some lattice* $\Gamma$.

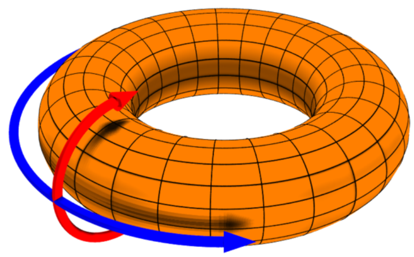


FIGURE 1. Meridiano (in rosso) e longitudine (in blu) sono due generatori di $\pi_1(T) = \mathbb{Z} \times \mathbb{Z}$.

PROOF. As a complete flat orientable surface, a flat torus is isometric to $\mathbb{R}^n/_\Gamma$ for some discrete group $\Gamma$ of orientation-preserving isometries of $\mathbb{R}^n$ acting freely. An orientation-preserving isometry of $\mathbb{R}^2$ without fixed points is a translation. $\square$

REMARK 1.3. Every translation $x \mapsto x + b$ in $\mathbb{R}^2$ commutes with $\Gamma$ and hence descends to an isometry on the flat torus $T = \mathbb{R}^2/_\Gamma$. Hence the isometry group $\mathrm{Isom}^+(T)$ is not discrete. Moreover, the flat torus is *homogeneus*, *i.e.* for every pair of points $x, y \in T$ there is an isometry sending $x$ to $y$.

EXERCISE 1.4. Let $T$ be a flat torus. Every non-trivial element $\gamma \in \pi_1(T)$ is represented by a closed geodesic, unique up to translations. The geodesic is simple if and only if $\gamma$ is primitive.

We have $\pi_1(T) = \mathbb{Z} \times \mathbb{Z}$, generated by a *meridian* $m = (1, 0)$ and *longitude* $l = (0, 1)$ as in Fig. 1. A flat metric $h$ on $T$ identifies $m$ and $l$ with two translations $w, z \in \mathbb{C} = \mathbb{R}^2$: the pair $(w, z)$ is well-defined up to an isometry of $\mathbb{C} = \mathbb{R}^2$. Hence the ratio $\frac{z}{w} \in \mathbb{C}$ depends only on $h$.

PROPOSITION 1.5. *We get a bijection:*

$$\begin{aligned} \mathrm{Teich}(T) &\longrightarrow H^2 \\ h &\longmapsto \frac{z}{w}. \end{aligned}$$

PROOF. The map is well-defined: an isometry relating $h$ and $h'$ isotopic to the identity fixes $m$ and $l$ up to isotopy and hence we get the same pair $(w, z)$, if we rescale the metric we get $(\lambda w, \lambda z)$ and hence the same $\frac{z}{w}$.

An inverse $H^2 \to \mathrm{Teich}(T)$ is as follows: for $z \in H^2$ take $\Gamma = \langle 1, z \rangle$ and assign to $T$ the metric of $\mathbb{H}^2/_\Gamma$ with $m$ and $n$ corresponding to 1 and $z$. $\square$

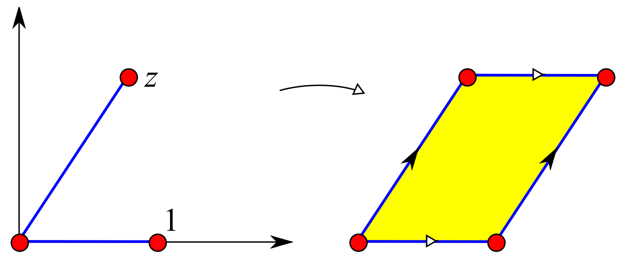



FIGURE 2. The flat metric on the torus $T$ determined by $z \in H^2$ may be constructed by identifying the opposite sides of the parallelogram with vertices $0, 1, z, z+1$. The lattice $\Gamma$ is generated by 1 and $z$ and the parallelogram is a fundamental domain.

The flat metric corresponding to $z \in H^2$ may be constructed by identifying the opposite sides of a parallelogram as in Fig. 2.

**1.2. Mapping class group.** A diffeomorphism $\varphi\colon S_g \to S_g$ transforms the metric $h$ into the metric $\varphi_* h$, defined as

$$(\varphi_* h)_{\varphi(x)}\big(d\varphi_x(v), d\varphi_x(w)\big) = h_x(v, w).$$

If $h$ varies through an isotopy, the metric $\varphi_*$ varies through a corresponding isotopy: therefore $\varphi$ acts on $\mathrm{Teich}(S_g)$ as follows:

$$\begin{aligned} \mathrm{Teich}(S_g) &\longrightarrow \mathrm{Teich}(S_g) \\ [h] &\longmapsto [\varphi_* h] \end{aligned}$$

If we vary $\varphi$ by an isotopy the action is unaffected. It is then natural to define the following important group.

DEFINITION 1.6. The *mapping class group* of $S_g$ is the group

$$\mathrm{MCG}(S_g) = \mathrm{Diffeo}^+(S_g)/_{\sim}$$

where $\mathrm{Diffeo}^+(S_g)$ indicates the group of all orientation-preserving diffeomorphisms $S_g \to S_g$ and $\varphi \sim \psi$ if $\varphi$ and $\psi$ are isotopic.

We have seen that $\mathrm{MCG}(S_g)$ acts on $\mathrm{Teich}(S_g)$: by definition, the quotient

$$\mathrm{Teich}(S_g)/_{\mathrm{MCG}(S_g)}$$

is the moduli space of $S_g$.

The group $\mathrm{MCG}(S_g)$ acts on the first homology group $H_1(S_g, \mathbb{Z})$ di $S_g$, since homotopic functions induce the same maps in homology. We get a group homomorphism

$$\mathrm{MCG}(S_g) \longrightarrow \mathrm{Aut}\big(H_1(S_g, \mathbb{Z})\big) = \mathrm{Aut}\big(\mathbb{Z}^{2g}\big) = \mathrm{GL}_{2g}(\mathbb{Z})$$

which is neither injective nor surjective in general. Its kernel is called the *Torelli group* of $S_g$.

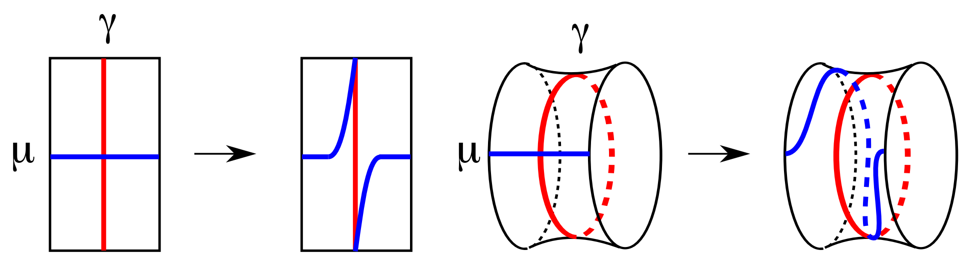



FIGURE 3. A Dehn twist along a curve $\gamma$ maps a transverse arc $\mu$ onto an arc which makes a complete left turn.

**1.3. The mapping class group and moduli space of the torus.** As usual, everything is simple on the torus $T$. Let $\mathrm{Aut}^+(H_1(T)) \cong \mathrm{SL}_2(\mathbb{Z})$ denote the automorphisms with positive (hence unit) determinant.

PROPOSITION 1.7. *The Torelli group of the torus $T$ is trivial and*

$$\mathrm{MCG}(T) \cong \mathrm{Aut}^+(H_1(T)).$$

PROOF. Consider the meridian $m$ and longitude $l$ of $T$. A diffeomorphism $\varphi$ of $T$ that acts trivially on $H_1(T) = \pi_1(T) = \mathbb{Z}^2$ sends $m$ and $l$ to two simple closed curves $m'$ e $l'$ homotopic and hence isotopic to $m$ and $l$: the proof of Proposition 2.26 from Capitolo 4 applies in this simple case to prove that $\varphi$ is isotopic to the identity.

A diffeomorphism $\varphi$ is orientation-preserving and hence acts on the group $H_1(T, \mathbb{Z}) = Z^2$ with positive determinant. Conversely, a matrix $A \in \mathrm{SL}_2(\mathbb{Z})$ acts linearly on $\mathbb{R}^2$ preserving the lattice $\mathbb{Z}^2$ and hence descends to $T = \mathbb{R}^2/_{\mathbb{Z}^2}$. $\square$

PROPOSITION 1.8. *The action of* $\mathrm{MCG}(T)$ *on* $\mathrm{Teich}(T)$ *is the following action of* $\mathrm{SL}_2(\mathbb{Z})$ *on* $H^2$ *as Möbius transformations:*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} : z \longmapsto \frac{az-b}{-cz+d}.$$

PROOF. The metric $z$ assigns to $T$ the structure $\mathbb{R}^2/_\Gamma$ with $\Gamma = \langle 1, z\rangle$ and $(m, l)$ corresponding to $(1, z)$.

Pick $\varphi = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z}) = \mathrm{MCG}(T)$. Since $\varphi^{-1} = \left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right)$, in the new metric $\varphi_*$ the pair $(m, l)$ corresponds to $(d - cz, -b + az)$. $\square$

The kernel of the action is $\{\pm I\}$: two matrices $A$ and $-A$ act in the same way on $\mathrm{Teich}(T)$.

Add description of moduli space of the torus.

**1.4. Dehn twist.** Let $\gamma$ be a homotopically non-trivial simple closed curve in an oriented surface $S_g$. The *Dehn twist* along $\gamma$ is an element $T_\gamma \in \mathrm{MCG}(S_g)$ defined as follows.

Pick a tubular neighborhood of $\gamma$ diffeomorphic to $S^1 \times [-1, 1]$ where $\gamma$ lies as $S^1 \times \{0\}$. Let $f : [-1, 1] \to \mathbb{R}$ be a smooth function which is zero in

$\left[-1,-\frac{1}{2}\right]$ and $2\pi$ on $\left[\frac{1}{2},1\right]$. Let

$$T_\gamma\colon S_g \longrightarrow S_g$$

be the diffeomorphism that acts on the tubular neighborhood as $T_\gamma(e^{i\alpha},t) = (e^{i(\alpha+f(t))},t)$ and on its complementary set in $S_g$ as the identity. We may visualize $T_\gamma$ by noting that it gives a complete left turn to any arc that intersects $\gamma$ as in Fig. 5.

PROPOSITION 1.9. *The element $T_\gamma \in \mathrm{MCG}(S_g)$ depends only on the homotopy class of $\gamma$.*

PROOF. When defining $T_\gamma$ we have chosen a tubular neighborhood for $\gamma$ and a smooth function $f$. Tubular neighborhoods are ambiently isotopic, and increasing functions with fixed extremes are isotopic too: this facts imply easily that $T_\gamma$ is well-defined up to isotopy. A homotopy of $\gamma$ can be promoted to an ambient isotopy by Proposition 2.16. $\square$

REMARK 1.10. To define $T_\gamma$ we needed the orientation of $S_g$ to distinguish from "turning left" and "turning right", but not an orientation for $\gamma$. If we change the orientation of $\gamma$ the element $T_\gamma$ is unaffected.

PROPOSITION 1.11. *The Dehn twists along $m$ and $l$ in $\mathrm{MCG}(T) = \mathrm{SL}_2(\mathbb{Z})$ are*

$$\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

PROOF. In homology we find

$$T_m(m) = m, \quad T_m(l) = l - m, \qquad T_l(l) = l, \quad T_l(m) = m + l.$$

$\square$

We have constructed a bijection between $\mathrm{Teich}(T)$ and $H^2$. We now want to construct for $g \geqslant 2$ some analogous identifications between $\mathrm{Teich}(S_g)$ and some open set of $\mathbb{R}^N$ for some $N$ depending on $g$. To this purpose we need to introduce some concepts.

**1.5. Earthquakes.** Hyperbolic metrics may be *twisted* along simple closed curves: this operation is called an *earthquake*.

Let $h$ be a complete hyperbolic metric on an oriented surface $S_g$ and $\gamma$ a simple closed geodesic. Fix an *angle* $\theta \in \mathbb{R}$. Informally, a new metric $h_\theta$ is constructed by sliding one of the two components near $\gamma$ on the left by the $\theta$. Formally, the new metric is defined as follows.

Recall from Proposition 2.15 in Chapter 3 that $\gamma$ has a $R$-neighborhood isometric to a $R$-tube for some $R > 0$. A $R$-tube here is a $R$-annulus as in Fig. 4, defined by quotienting a $R$-neighborhood of a line $l$ by a hyperbolic transformation. The $R$-annulus is naturally parametrized as $\gamma \times [-R,R] \cong S^1 \times [-R,R]$, where $x \times [-R,R]$ are geodesic segments orthogonal to $\gamma$.

We choose a diffeomorphism $\varphi$ of $S^1 \times [-R,R]$ that curves the segments by an angle $\theta$ as in Fig. 5-(right). More precisely, let $f\colon [-R,R] \to \mathbb{R}$ be

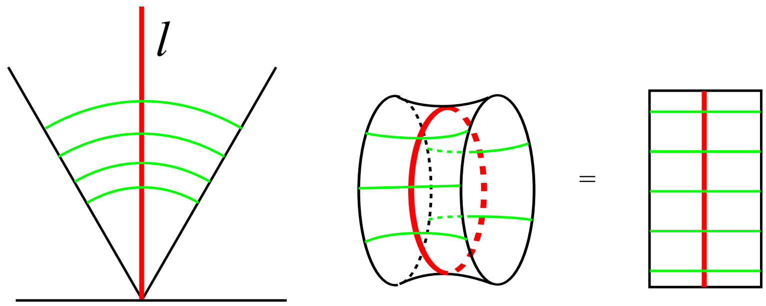



FIGURE 4. A $R$-annulus around a geodesic $\gamma$ is the quotient of a $R$-neighborhood of a line $l$ by a hyperbolic transformation. The orthogonal (green) geodesic segments are parametrized by arc-length as $[-R, R]$, hence the $R$-annulus is naturally parametrized as $\gamma \times [-R, R]$.

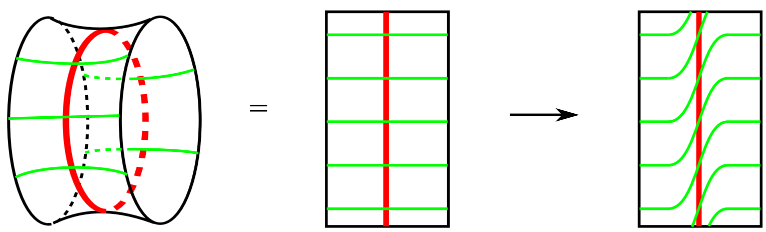



FIGURE 5. To define the earthquake we pick a diffeomorphism of the $R$-annulus that modifies the orthogonal segments as shown here.

a strictly increasing smooth function which is zero on $\left[-R, -\frac{R}{2}\right]$ and $\theta$ on $\left[\frac{R}{2}, R\right]$. We set $\varphi(e^{i\alpha}, t) = (e^{i(\alpha + f(t))}, t)$.

We define a new metric $h_\theta$ on $S_g$ as follows: the metric tensor $h_\theta$ coincides with $\varphi_* h$ on the $R$-annulus and coincides with $h$ on the complement of the $\frac{R}{2}$-annulus $\left[\frac{R}{2}, \frac{R}{2}\right] \times S^1$.

PROPOSITION 1.12. *The metric tensor $h_\theta$ is well-defined and gives a complete hyperbolic metric on $S_g$.*

PROOF. It is well-defined because $h$ and $h_\theta$ coincide on $S^1 \times \left[\frac{R}{2}, R\right]$, because $(e^{i\alpha}, t) \mapsto (e^{i(\alpha+\theta)}, t)$ is an isometry of the $R$-annulus. □

REMARK 1.13. In the new metric $h_\theta$ the curve $\gamma$ is still a geodesic of the same length as before, and its $R$-neighborhood is also unchanged.

As for Dehn twists, earthquakes define an action on Teichmüller space. If $\gamma$ is a simple closed curve we let $h_\theta^\gamma$ be the result of an earthquake of angle $\theta$ performed along the unique geodesic homotopic to $\gamma$ in the metric $h$.

PROPOSITION 1.14. *The map*

$$\begin{aligned} E_\gamma \colon \mathbb{R} \times \mathrm{Teich}(S_g) &\longrightarrow \mathrm{Teich}(S_g) \\ (\theta, h) &\longmapsto h_\theta^\gamma \end{aligned}$$

*is an action of* $\mathbb{R}$ *on* $\mathrm{Teich}(S_g)$*, determined only by the homotopy class of* $\gamma$*.*

PROOF. We prove that $E_\gamma$ is well-defined. The only ambiguity in the definition of $h_\theta$ is the choice of the function $f$. If we use another function $f'$ the resulting metric changes only by an isotopy: the diffeomorphism of $S_g$ which is the identity outside the $R$-annulus and sends $(e^{i\alpha}, t)$ to $(e^{i(\alpha+f(t)-f'(t))}, t)$ is an isometry between the two metrics, and is isotopic to the identity.

To prove that $E_\gamma$ is an action we need to check that

$$h^\gamma_{\theta+\theta'} = \left(h^\gamma_{\theta'}\right)^\gamma_\theta .$$

By Remark 1.13 we can take the same $R$-annulus to compose two earthquakes and hence the equality follows. $\square$

As for Dehn twists, to define the action $E_\theta$ we needed the orientation of $S_g$ but not one for $\gamma$. The earthquake action extends continuously the discrete action of Dehn twists:

PROPOSITION 1.15. *We have* $T_\gamma(h) = E_\gamma(2\pi, h)$*.*

PROOF. It follows directly from their definitions. $\square$

**1.6. Length functions.** A homotopically non-trival closed curve $\gamma$ in $S_g$ with $g \geqslant 2$ defines a *length function*

$$\ell^\gamma \colon \mathrm{Teich}(S_g) \to \mathbb{R}_{>0}$$

which assigns to a metric $h \in \mathrm{Teich}(S_g)$ the length $\ell^\gamma(h)$ of the unique geodesic isotopic to $\gamma$.

PROPOSITION 1.16. *The function* $\ell^\gamma$ *is well-defined.*

PROOF. If $h' = \varphi_* h$ for some isometry $\varphi$ then $\ell^\gamma(h') = \ell^{\varphi(\gamma)}(h)$. If $\varphi$ is isotopic to the identity the curves $\varphi(\gamma)$ and $\gamma$ are isotopic and hence $\ell^{\varphi(\gamma)}(h) = \ell^\gamma(h)$ by definition. $\square$

**1.7. Earthquakes and length functions on the torus.** Most of what we said extends to the torus case and can be nicely described. On a flat torus a closed geodesic is unique in its homotopy class only up to translations, and we fix the convention that the metric $h \in \mathrm{Teich}(T)$ is always rescaled to have unit area. Earthquakes and length functions are hence well-defined, and can be written explicitly as we now see. As usual we denote simple closed curves as coprime pairs $(p, q)$ and identify $\mathrm{Teich}(T)$ with $H^2 \subset \mathbb{C}$.

PROPOSITION 1.17. *The formula holds:*

$$\ell^{(p,q)}(z) = \frac{\|p + qz\|}{\sqrt{\Im z}}$$

*for any curve* $(p, q)$ *and any metric* $z \in \mathrm{Teich}(T)$*.*

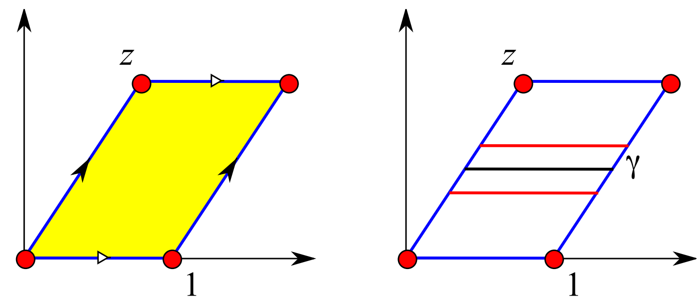



FIGURE 6. A torus with metric $z$ (left) twisted along the curve $\gamma$ (right).

PROOF. Up to rescaling we have $T = \mathbb{R}^2/_\Gamma$ with $\Gamma = \langle 1, z\rangle$. The translation in $\Gamma$ corresponding to $(p,q)$ is $p \cdot 1 + q \cdot z$ and the closed geodesic it produces has length $\|p+qz\|$. The area of the torus $T$ is $\Im z$ and hence we must rescale it by $1/\sqrt{\Im z}$. □

We can write the earthquake action along the meridian $m$ of $T$.

EXERCISE 1.18. We have:

$$E_m(\theta)\colon z \mapsto z + \frac{\theta}{2\pi}.$$

HINT. Draw $T$ and $\gamma$ as in Fig. 2. □

COROLLARY 1.19. *The earthquake action $E_{(p,q)}$ is the 1-parameter family of parabolic transformations with fixed point $-\frac{p}{q} \in \partial H^2$.*

PROOF. We know the case $(p,q) = (1,0) = m$, and the general case follows from Proposition 1.8. □

The orbits of $E_{(p,q)}$ are of course the horospheres centered at $-\frac{p}{q}$.

COROLLARY 1.20. *If $i(\gamma, \gamma') > 0$ then $\ell^\gamma$ is strictly convex along the orbits of $E_{\gamma'}$*

PROOF. We may take $\gamma' = m$ and note that the function in Proposition 1.17 is strictly convex on the horospheres $\Im z = k$. □

**1.8. Convexity of the length functions.** In higher genus there is no nice explicit formula for $\ell^\gamma$, but we will generalize Corollary 1.20 anyway. We will use the following.

EXERCISE 1.21. Let $f\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}_{\geqslant 0}$ be strictly convex and proper. The function

$$\begin{aligned} F\colon \mathbb{R}^n &\longrightarrow \mathbb{R}_{\geqslant 0} \\ y &\longmapsto \min\left\{ f(x,y) \;\middle|\; x \in \mathbb{R}^m \right\} \end{aligned}$$

is well-defined, strictly convex and proper.

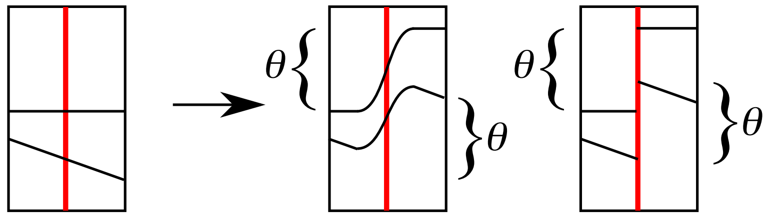



FIGURE 7. A geodesic $\overline{\eta}$ in $h^\gamma_\theta$ is a geodesic outside the annulus and deviates on the left by an angle $\theta$ when crossing it (center) since $h^\gamma_\theta$ is obtained by twisting a hyperbolic metric (left). We may simplify the picture by describing $\overline{\eta}$ as a broken geodesic line (right).

Length functions are either constant or strictly convex on orbits of earthquakes:

PROPOSITION 1.22. *Let $\eta$ and $\gamma$ be two homotopically non-trivial simple closed curves in and $h$ be a hyperbolic metric on $S_g$. The function*

$$\begin{aligned} \mathbb{R} &\longrightarrow \mathbb{R}_{\geqslant 0} \\ \theta &\longmapsto \ell^\eta(h^\gamma_\theta) \end{aligned}$$

*is*

- *constant if $i(\eta,\gamma)=0$,*
- *strictly convex and proper if $i(\eta,\gamma)>0$.*

PROOF. We can suppose that $\gamma$ is geodesic with respect to $h$. If $i(\eta,\gamma) = 0$ the curves $\eta$ and $\gamma$ are disjoint geodesics and the length of $\eta$ is not affected by the earthquakes we perform near $\gamma$.

Consider the case $n = i(\eta,\gamma) > 0$ e $g = 1$. Denote by $\overline{\eta}^\theta$ the geodesic isotopic to $\eta$ in the twisted metric $h^\gamma_\theta$: it intersects $\gamma$ transversely in $n$ points.

Fix a sufficiently small $R$-annulus around $\theta$ and recall that the geodesics in $h^\gamma_\theta$ are curves that are geodesic outside the $R$-annulus and deviate smoothly on the left when crossing it as in Fig. 7-(center). We may represent efficiently a geodesic by substituting every smooth deviation with a broken jump as shown in Fig. 7-(right). We get a bijection

$$\left\{ \begin{array}{c} \text{closed geodesics} \\ \text{with respect to } h^\gamma_\theta \end{array} \right\} \longleftrightarrow \left\{ \begin{array}{c} \text{broken geodesics} \\ \text{with respect to } h \end{array} \right\}$$

where a *broken geodesic* is a geodesic which at every crossing of $\gamma$ jumps to the left at distance $\frac{\theta l(\gamma)}{2\pi}$. This correspondence is useful because it preserves the lengths: the length of the closed geodesic for $h^\gamma_\theta$ is equal to the length of the broken geodesic (which is the sum of the lengths of its components), because the segments in Fig. 7-(left) and (right) are isometric.

We lift this description to the universal cove $\mathbb{H}^2$ and fix a lift $l$ of $\overline{\eta}^0$. Pick $n+1$ consecutive intersections $r_1,\ldots,r_{n+1}$ of $l$ with the lifts of $\gamma$ as in Fig. 7. The hyperbolic transformation $\tau$ with axis $l$ corresponding to $\eta$ sends $r_1$ to $r_{n+1}$.

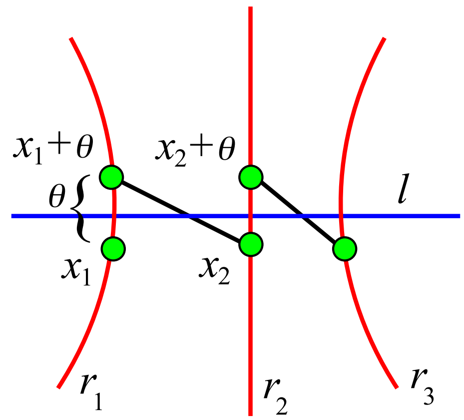



FIGURE 8. The (blue) line $l$ is a lift of $\overline{\eta}$. Consider $n$ consecutive intersections with lifts of $\gamma$ (in red): the $(n+1)$-th lift is equivalent to the first via the hyperbolic transformation $\tau$ having axis $l$ (here $n = 2$).

The closed geodesic $\overline{\eta}^\theta$, represented as a broken geodesic, lifts to a geodesic which starts from some point $x_1 + \theta \in r_1$ and arrives to some point $x_2 \in r_2$, then jumps on the left at distance $\frac{\theta l(\gamma)}{2\pi}$ and start again from $x_2 + \theta$, and so on until it reaches the point $\tau(x_1) \in \tau(r_1) = r_{n+1}$. As $x_1 \in r_1, \ldots, x_n \in r_n$ vary we get various broken paths, but only one starts and arrive at the same $r_i$ with the same angle, thus representing a closed geodesics (because it is unique). The other broken paths represent piecewise-geodesic curves in $h_\theta^\gamma$ and therefore are longer than $\overline{\eta}^\theta$. Hence

$$l(\overline{\eta}^\theta) = \min\left\{ \sum_{i=1}^n d(x_i + \theta, x_{i+1}) \;\middle|\; (x_1, \ldots, x_n) \in \mathbb{R}^n \right\}$$

where $x_{n+1} = \tau(x_1)$. It remains to prove that the function $\theta \mapsto l(\overline{\eta}^\theta)$ is proper and strictly convex. The function

$$\psi\colon \mathbb{R}^{2n} \longrightarrow \mathbb{R}$$
$$(x_1, y_1, \ldots, x_n, y_n) \longmapsto \sum_{i=1}^n d(y_i, x_{i+1})$$

where $x_{n+1} = \tau(x_1)$ is stricly convex and proper by Proposition 3.4. The auxiliary function

$$\phi\colon \mathbb{R}^{2n} \times \mathbb{R} \longrightarrow \mathbb{R}$$
$$(x, \theta) \longmapsto \psi(x)$$

is only convex, but its restriction to the subspace

$$H = \{y_i = x_i + \theta\}$$

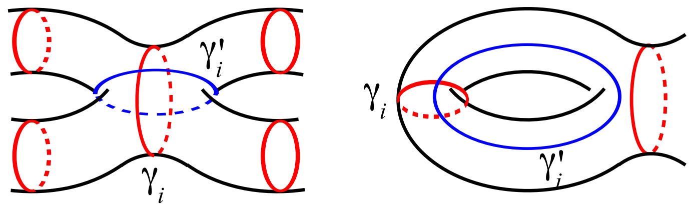



FIGURE 9. Choose for any component $\gamma_i$ of a pants decomposition $\mu$ a curve $\gamma_i'$ that intersect $\gamma_i$ in one or two points and is disjoint from the other components. There are two cases to consider, depending on whether the two pairs-of-pants adjacent to $\gamma_i$ are distinct (left) or not (right).

is strictly convex and proper, because $H$ is not parallel to the direction $(0,\ldots,0,1)$. The coordinates $x_i$ and $\theta$ identify $H$ with $\mathbb{R}^n\times\mathbb{R}$. The restriction $f=\psi|_H$ is hence a function $f\colon\mathbb{R}^n\times\mathbb{R}\to\mathbb{R}$ and we obtain

$$l(\overline{\eta}^\theta)=\min\left\{f(x,\theta)\;\middle|\;x\in\mathbb{R}^n\right\}.$$

By Exercise 1.21 the function $\theta\mapsto l(\overline{\eta}^\theta)$ is strictly convex and proper. $\square$

**1.9. Earthquakes and pants decomposition.** The convexity of the length functions imply easily the following.

COROLLARY 1.23. *For any simple closed curve $\gamma$, the earthquake action $E_\gamma$ is faithful.*

PROOF. Suppose by contradiction that $h=h^\gamma_\theta$, then $h=h^\gamma_{n\theta}$ for any $n\in\mathbb{N}$. Let $\eta$ be a simple closed curve with $i(\eta,\gamma)>0$, see Exercise 2.13 from Chapter 4; the function $\theta\mapsto\ell^\eta(h^\gamma_\theta)$ is strictly convex and periodic: a contradiction. $\square$

The earthquake action may be define more generally for essential multicurves. An essential multicurve $\mu=\gamma_1\sqcup\cdots\sqcup\gamma_k$ of $S_g$ determine an action

$$\begin{aligned}E_\mu\colon\mathbb{R}^k\times\mathrm{Teich}(S_g)&\longrightarrow\mathrm{Teich}(S_g)\\(\theta,h)&\longmapsto h^\mu_\theta\end{aligned}$$

where $\theta=(\theta_1,\ldots,\theta_k)$ and $h^\mu_\theta=h^{\gamma_1}_{\theta_1}\circ\cdots\circ h^{\gamma_k}_{\theta_k}$. This action is again faithful:

COROLLARY 1.24. *For every essential multicurve $\mu$, the earthquake action is faithful.*

PROOF. We may complete $\mu$ to a pants-decomposition. Pick for any $i=1,\ldots,3g-3$ a curve $\gamma_i'$ as in Fig. 9 such that $i(\gamma_i,\gamma_i')>0$ for all $i$ and $i(\gamma_i,\gamma_j')=0$ per ogni $i\neq j$. Suppose by contradiction that $h=h^\mu_\theta$ for some $\theta\neq0$: hence $h^\mu_{n\theta}=h$ for all $n\in\mathbb{Z}$. There is a $i$ such that $\theta_i\neq0$. The length function $\ell^{\gamma_i'}$ depends only on $\theta_i$ and not on all coordinates of $\theta$, because $\gamma_i'\cap\gamma_j=\emptyset$ for all $i\neq j$. This function is convex and periodic in $\theta_i$: a contradiction. $\square$

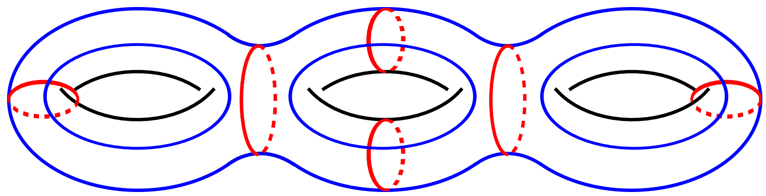
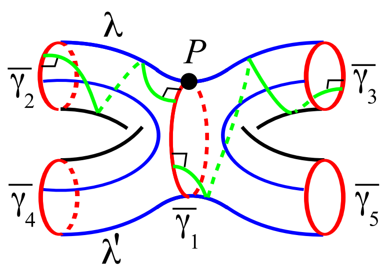


FIGURE 10. A frame for Fenchel-Nielsen coordinates consists of a (red) pants decomposition $\mu$ and a (blue) transverse multicurve $\eta$ that cuts each pants into two hexagons. The number of components of $\mu$ is $3g-3$, that of $\eta$ can vary.



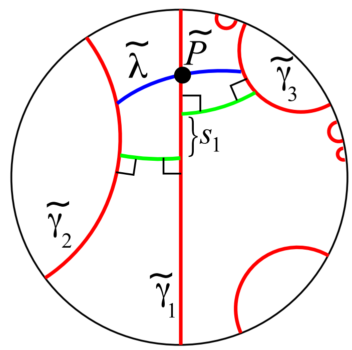




FIGURE 11. A geodesic $\gamma_1$ and the two adjacent pairs-of-pants. The torsion parameter $t_1$ measures the distance (in the universal covering) between the two orthogeodesics.

## 2. Fenchel-Nielsen coordinates

**2.1. The coordinates.** We want to fix a bijection between Teichmüller space Teich$(S_g)$ and $\mathbb{R}^{6g-6}$ for all $g \geqslant 2$; as for vector spaces, this bijection depends on the choice of a "frame", which consists here of an orientation for $S_g$ and two essential multicurves $\mu$ e $\nu$ in minimal position, such that:

(1) the multicurve $\mu$ is a pants-decomposition,
(2) the multicurve $\nu$ decomposes every pair-of-pants in two hexagons.

A frame is shown in Fig. 10. A induces a map

$$\begin{aligned} \mathrm{FN}\colon \mathrm{Teich}(S_g) &\longrightarrow \mathbb{R}^{3g-3}_{>0} \times \mathbb{R}^{3g-3} \\ h &\longmapsto (l_1, \ldots, l_{3g-3}, \theta_1, \ldots, \theta_{3g-3}) \end{aligned}$$

as follows. The *length parameters* $l_i = \ell^{\gamma_i}(h)$ are the length functions on the curves of $\mu = \gamma_1 \sqcup \ldots \sqcup \gamma_{3g-3}$. The *torsion angles* $\theta_i$ are defined as follows. Let $\overline{\gamma_i}$ be the geodesic homotopic to $\gamma_i$ in the metric $h$. The geodesic multicurve $\overline{\mu} = \overline{\gamma_1} \sqcup \ldots \sqcup \overline{\gamma_{3g-3}}$ decomposes $S_g$ into geodesic pair-of-pants.

To definire $\theta_i$ we need the second multicurve $\nu$. Fig. 11-(left) shows the two geodesic pants adjacent to $\overline{\gamma_1}$ (which might coincide). The second multicurve $\eta$ intersects these pants in four blue arcs, two of which $\lambda$, $\lambda'$

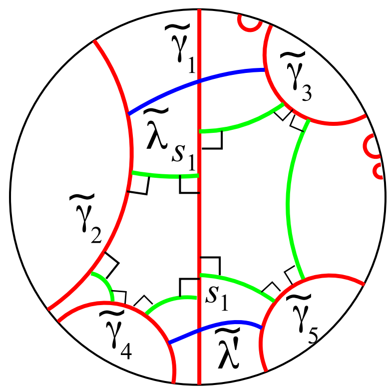



FIGURE 12. If we pick $\lambda'$ instead of $\lambda$ we find a segment of the same length $s_1$, since the two right-angled hexagons shown are isometric.

intersect $\overline{\gamma_1}$: we pick one, say $\lambda$. We fix a lift $\widetilde{P} \in \mathbb{H}^2$ of $P = \overline{\gamma_1} \cap \lambda$ and we lift from $\widetilde{P}$ the curve $\overline{\gamma_1}$ to a line $\widetilde{\gamma_1}$ and $\lambda$ to a (non-geodesic) curve $\widetilde{\lambda}$ that connects two coverings $\widetilde{\gamma_2}$ and $\widetilde{\gamma_3}$ of the closed geodesics $\overline{\gamma_2}$ e $\overline{\gamma_3}$. See Fig. 11-(right).

We draw as in the figure the orthogeodesics connecting $\widetilde{\gamma_1}$ to $\widetilde{\gamma_2}$ and $\widetilde{\gamma_3}$ and we denote by $s_1$ the signed length of the segment in $\widetilde{\lambda}_i$ comprised between the two orhogeodesics, with positive sign if (as in the figure) an observer walking on a orthogeodesic towards $\widetilde{\gamma}_i$ sees the other orthogeodesics on the left (here we use the orientation of $S_g$).

By repeating this construction for each $\overline{\gamma_i}$ we find some real numbers $s_i$. The torsion parameter $\theta_i$ is

$$\theta_i = \frac{2\pi s_i}{l_i}.$$

THEOREM 2.1 (Fenchel-Nielsen coordinates). *The map* FN *is well-defined and a bijection.*

PROOF. We first note that while defining the torsion parameters we could have chosen $\lambda'$ instead of $\lambda$, but we would have obtained the same length $s_i$ as shown in Fig. 12. Moreover if $h'$ is a hyperbolic metric isometric to $h$ through a diffeomorphism $\varphi$ isotopic to the identity, the parameters $l_i$ and $\theta_j$ depend only on the isotopy class of $\mu$ and $\eta$ and hence do not vary. Therefore FN is well-defined on $\mathrm{Teich}(S_g)$.

We prove that FN is surjective. For any vector $(l_1, \ldots, l_{3g-3}) \in \mathbb{R}^{3g-3}_{>0}$ we may construct a metric on $S_g$ by assigning to each pants of the pants-decomposition $\mu$ the hyperbolic metric with boundary lengths $l_i$. We get some twisted parameters, which can be changed arbitrarily by an earthquake along $\mu$: an earthquake of angle $\theta'$ changes them from $\theta$ to $\theta + \theta'$, hence any torsion parameter can be obtained.

We prove that FN is injective. Suppose $\mathrm{FN}(h) = \mathrm{FN}(h')$. Up to acting via earthquakes we suppose that $\mathrm{FN}(h) = \mathrm{FN}(h') = (l_1, \ldots, l_{3g-3}, 0, \ldots, 0)$.

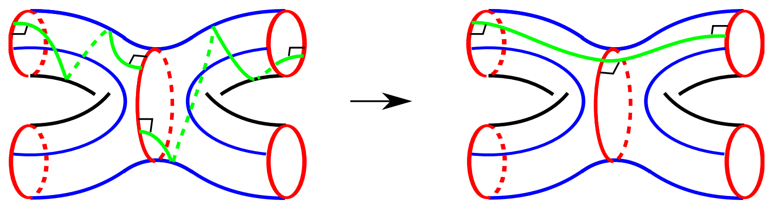


FIGURE 13. Una torsione di un opportuno angolo $\theta$ permette di far combaciare le curve verdi minimizzando le intersezioni con le curve blu.

Since the torsion parameter is zero, the orthogeodesics in Fig. 11-(right) match, and project in $S_g$ to a geodesic multicurve $\overline{\nu}$ isotopic to $\nu$ and orthogonal to $\overline{\mu}$. Therefore $S_g \setminus (\overline{\mu} \cup \overline{\nu})$ is a tessellation of $S_g$ into right-angled hexagons, determined by the lengths $l_i$. Both metrics $h$ and $h'$ have the same tessellation and are hence isometric, via an isometry which is isotopic to the identity. $\square$

REMARK 2.2. As shown in the proof, the torsion parameters for $h$ are zero if and only if the geodesic representatives $\overline{\nu}$ and $\overline{\mu}$ of $\nu$ and $\mu$ are everywhere orthogonal.

**2.2. Length functions of $9g-9$ curves.** We show here that a finite number $9g-9$ of length functions suffice to determine a point in Teichmüller space.

Let $\mu = \gamma_1 \sqcup \ldots \sqcup \gamma_{3g-3}$ be a pants-decomposition for $S_g$: for any $\gamma_i$ we choose a curve $\gamma_i'$ as in Fig. 9. We indicate by $\gamma_i'' = T_{\gamma_i}(\gamma_i')$ the curve obtained by Dehn-twisting $\gamma_i'$ along $\gamma_i$.

PROPOSITION 2.3. *The map*

$$\begin{aligned} L\colon \mathrm{Teich}(S_g) &\longrightarrow \mathbb{R}^{9g-9}_{>0} \\ h &\longmapsto \big(\ell^{\gamma_i}(h), \ell^{\gamma_i'}(h), \ell^{\gamma_i''}(h)\big) \end{aligned}$$

*is injective.*

PROOF. We compose $L$ with $\mathrm{FN}^{-1}$ and obtain a map

$$\begin{aligned} L \circ \mathrm{FN}^{-1} \colon \mathbb{R}^{3g-3}_{>0} \times \mathbb{R}^{3g-3} &\longrightarrow \mathbb{R}^{9g-9}_{>0} \\ (l_i, \theta_i) &\longmapsto (l_i, l_i', l_i'') \end{aligned}$$

We prove that it is injective. It suffices to consider the case where the values $l_i$ are fixed and $\theta_i$ vary. Note that $\gamma_i'$ and $\gamma_i''$ intersect $\gamma_j$ if and only if $i = j$: hence $l_i'$ and $l_i''$ depend only on $\theta_i$ and not on the other torsion parameters $\theta_j$. We know that $l_i' = f(\theta_i)$ is strictly convex and $l_i'' = f(\theta_i + 2\pi)$ by Proposition 1.15. A proper convex function is not injective, but the function

$$\begin{aligned} \mathbb{R} &\longrightarrow \mathbb{R} \times \mathbb{R} \\ \theta_i &\longmapsto \big(f(\theta_i), f(\theta_i + 2\pi)\big) \end{aligned}$$

is injective. Hence $L$ is injective. $\square$

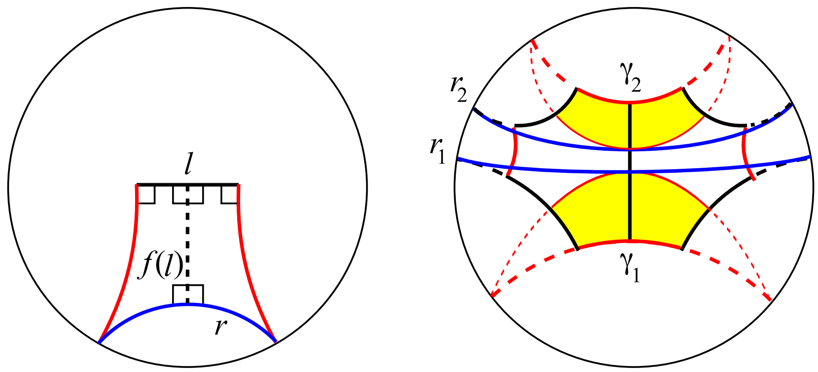



FIGURE 14. For every $l$ there is a unique right-angled hexagons with alternate sides of length $l, 0, 0$. Let $f(l)$ be the distance between the opposite sides $l$ and $r$ (left). We draw a geodesic pair-of-pants as the union of two isometric hexagons: the closed geodesics $\gamma_1$ and $\gamma_2$ have disjoint $f(l_1)$ and $f(l_2)$-neighborhoods, colored here in yellow (right).

An analogous result holds for the torus $T$. Let $\gamma, \gamma'$ be two simple closed curves with $i(\gamma, \gamma') = 1$ and $\gamma'' = T_\gamma(\gamma')$.

PROPOSITION 2.4. *The map*

$$\begin{aligned} L\colon \operatorname{Teich}(T) &\longrightarrow \mathbb{R}^3_{>0} \\ h &\longmapsto \left(\ell^\gamma(h), \ell^{\gamma'}(h), \ell^{\gamma''}(h)\right) \end{aligned}$$

*is injective.*

PROOF. Up to the action of $\mathrm{MCG}(T) = \mathrm{SL}_2(\mathbb{Z})$ we take $\gamma = (1,0)$, $\gamma' = (0,1)$, and $\gamma'' = (-1,1)$. Exercise 1.17 gives

$$L(z) = \left(\frac{1}{\sqrt{\Im z}}, \frac{\|z\|}{\sqrt{\Im z}}, \frac{\|z-1\|}{\sqrt{\Im z}}\right)$$

which is injective. □

**2.3. Collar lemma.** Every simple closed geodesic $\gamma$ has a $R$-annulus neighborhood. We show here that we can take an arbitrarily big $R$ if $\gamma$ is sufficiently short.

For any $l > 0$, draw the unique right-angled hexagons with alternate sides of length $l, 0, 0$ as in Fig. 14-(left). Let $f(l)$ be the distance between the side $l$ and the opposite side $r$.

EXERCISE 2.5. The function $f\colon \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ is strictly increasing and $\lim_{l\to 0} f(l) = \infty$.

LEMMA 2.6 (Collar lemma). *Let $S_g$ have a hyperbolic metric. Disjoint simple closed geodesics $\gamma_1, \ldots, \gamma_k$ of length $l_1, \ldots, l_k$ have disjoint $f(l_i)$-annular neighborhoods.*

Proof. We may suppose that the closed geodesics form a pants deomposition, and it suffices to consider the $f(l_i)$-neighborhoods of two curves $\gamma_1$ and $\gamma_2$ incident to the same pair-of-pants $P$.

The geodesic pair-of-pants $P$ subdivide into two isometric hexagons. We lift a hexagon to the universal cover $\mathbb{H}^2$ as in Fig. 14-(right). The $l(\gamma_i)$-neighborhoods of $\gamma_1$ and $\gamma_2$ are drawn in yellow and are disjoint. $\square$

The collar lemma has various consequences.

Corollary 2.7. *Let $S_g$ have a hyperbolic metric. Let $\gamma$ and $\eta$ be two simple closed geodesics in $S_g$. The inequality holds:*

$$l(\eta) \geqslant 2i(\gamma,\eta)\cdot f(l(\gamma)).$$

Proof. The geodesic $\gamma$ has an $f(l(\gamma))$-annular neighborhood. The geodesic $\eta$ intersecs $\gamma$ in $i(\eta,\gamma)$ points and hence crosses the annular neighborhood $i(\eta,\gamma)$ times, each with a segment of length at least $2f(l(\gamma))$. $\square$

**2.4. Topology of Teichmüller space.** We indicate by $\mathscr{S} = \mathscr{S}(S_g)$ the set of all homotopically non-trivial simple closed curves in $S_g$ with $g \geqslant 2$, seen up to isotopy and change of orientation. Each element $\gamma \in \mathscr{S}$ induces a length function

$$\ell^\gamma \colon \mathrm{Teich}(S_g) \longrightarrow \mathbb{R}_{>0}.$$

We indicate as usual with $\mathbb{R}^{\mathscr{S}}$ the set of all functions $\mathscr{S} \to \mathbb{R}$ and give it the usual product topology (the weakest one where all projections are continuous). The natural map

$$\begin{aligned}\mathrm{Teich}(S_g) &\longrightarrow \mathbb{R}^{\mathscr{S}}\\ h &\longmapsto \big(\gamma \longmapsto \ell^\gamma(h)\big)\end{aligned}$$

is injective by Proposition 2.3. We may hence consider $\mathrm{Teich}(S_g)$ as a subspace $\mathbb{R}^{\mathscr{S}}$ and give it the subspace topology. This topology on $\mathrm{Teich}(S_g)$ is the weakest one where the length functions $\ell^\gamma$ are continuous.

Proposition 2.8. *The space $\mathbb{R}^{\mathscr{S}}$ is Hausdorff and has a countable base.*

Proof. Product of Hausdorff spaces is Hausdorff, and product of spaces with countable bases has a countable base. $\square$

We recall the following topological fact.

Proposition 2.9. *Let $f\colon X \to Y$ be a continuous and proper map between topological spaces. If $Y$ is Hausdorff and has a countable base then $f$ is closed.*

Corollary 2.10. *Let $f\colon X \to Y$ be a continuous, injective, and proper map between topological spaces. If $Y$ is Hausdorff and has countable base then $f$ is a homeomorphism onto its image.*

Recall that with the half-space model every isometry in $\mathrm{Isom}^+(H^2) = \mathbb{P}\mathrm{SL}_2(\mathbb{R})$ is a $2\times 2$ matrix determined up to sign.

PROPOSITION 2.11. *Let $S_g = \mathbb{H}^2/_\Gamma$ be a hyperbolic surface. A hyperbolic transformation $\varphi \in \Gamma$ determines a closed geodesic $\gamma$ in $S_g$ with*

$$2|\mathrm{tr}\varphi| = \cosh \frac{l(\gamma)}{2}.$$

PROOF. Up to conjugacy we have $\varphi(z) = e^{l(\gamma)}z$. The matrix is

$$\varphi = \begin{pmatrix} e^{\frac{l(\gamma)}{2}} & 0 \\ 0 & e^{-\frac{l(\gamma)}{2}} \end{pmatrix}$$

hence $2|\mathrm{tr}\varphi| = \cosh \frac{l(\gamma)}{2}$. □

In particular, the length of $\gamma$ depends continuously on the transformation $\varphi$. We will use this to prove the following.

PROPOSITION 2.12. *The Fenchel-Nielsen coordinates*

$$\mathrm{FN}\colon \mathrm{Teich}(S_g) \longrightarrow \mathbb{R}^{3g-3}_{>0} \times \mathbb{R}^{3g-3}$$

*are a homeomorphism.*

PROOF. We consider $\mathrm{Teich}(S_g)$ inside $\mathbb{R}^{\mathscr{S}}$ and examine the inverse map

$$\mathrm{FN}^{-1}\colon \mathbb{R}^{3g-3}_{>0} \times \mathbb{R}^{3g-3} \longrightarrow \mathbb{R}^{\mathscr{S}}.$$

We prove that $\mathrm{FN}^{-1}$ is continuous. The map $\mathrm{FN}^{-1}$ assigns to the parameters $(l_i, \theta_i)$ a metric on $S_g$ constructed by attaching right-angled hexagons. Both the hexagons and the attaching maps depend continuously on the parameters $(l_i, \theta_i)$ and lift to a tessellation of $\mathbb{H}^2$ into hexagons. Since the tessellation varies continuously, its deck transformations vary continuously (in the matrix topology of $\mathbb{P}\mathrm{SL}_2(\mathbb{R})$) and hence the length functions too by Proposition 2.11. Therefore $\mathrm{FN}^{-1}$ is continuous.

We prove that $\mathrm{FN}^{-1}$ is proper. Take a diverging sequence of parameters $(l_i, \theta_i)$ (that is, without converging subsequences) in $\mathbb{R}^{3g-3}_{>0} \times \mathbb{R}^{3g-3}$: we need to show that its image is also a diverging subsequence. This thesis is equivalent to show that the length function of some curve goes to infinity. If $l_i \to +\infty$ for some $i$ we are done. If $l_i \to 0$, the length of any curve intersecting essentially $\gamma_i$ goes to infinity by Corollary 2.7. Suppose then by contradiction that every length $l_i$ is bounded above and below but some $\theta_j$ goes to infinity: in that case the length of any curve intersecting $\gamma_j$ goes to infinity by Proposizione 1.22.

Finally, the map $\mathrm{FN}^{-1}$ is a homeomorphism onto its image by Corollary 2.10. □

Recall that the action of a topological group $G$ on a topological space $X$ is *continuous* if the action map $G \times X \to X$ is continuous. This implies that $G$ acts on $X$ by homeomorphisms.

PROPOSITION 2.13. *Earthquakes and mapping class group actions on* $\mathrm{Teich}(S_g)$ *are continuous.*

Proof. The mapping class group acts on $\mathscr{S}$ by permutations, hence its action on the whole $\mathbb{R}^{\mathscr{S}}$ is continuous. Concerning earthquakes, on Fenchel-Nielsen coordinates the earthquake action sends $\theta$ to $\theta + \theta'$. $\square$

The immersion in $\mathbb{R}^{9g-9}$ from Proposition 2.3 is also topologically faithful.

Proposition 2.14. *The injective representation* $\mathrm{Teich}(S_g) \to \mathbb{R}^{9g-9}$ *is a homeomorphism onto its image.*

Proof. Using Fenchel-Nielsen coordinates the map is clearly continuous. The proof that it is proper is as in Proposition 2.12. $\square$

CHAPTER 6

# Orbifolds

An orbifold is an object locally modeled on finite quotients of $\mathbb{R}^n$. It naturally arises as a quotient of a manifold by a discrete group, whose action is not free. Orbifolfds behave like manifolds on many aspects.

## 1. Generalities

**1.1. Definition.** We introduce the following notion.

DEFINITION 1.1. A $n$-dimensional *orbifold* is a topological Hausdorff space $O$ covered by a collection of open sets $\{U_i\}_{i\in I}$ closed by finite intersection and equipped with the following structure. For every $i \in I$ there is:

(1) a finite subgroup $\Gamma_i < O(n)$ and a $\Gamma_i$-invariant open set $V_i \subset \mathbb{R}^n$;
(2) a $\Gamma_i$-invariant continuous map $\varphi_i\colon V_i \to U_i$ which induces a homeomorphism $V_i/_{\Gamma_i} \to U_i$.

The charts must fulfill this compatibility condition:

- for every inclusion $U_i \subset U_j$ there is an injective homomorphism $f_{ij}\colon \Gamma_i \hookrightarrow \Gamma_j$ and a $\Gamma_i$-equivariant diffeomorfism $\psi_{ij}$ from $V_i$ and an open set in $V_j$ compatible with the charts, that is $\varphi_j \circ \psi_{ij} = \varphi_i$.

Fare figura per $\psi_{ij}$ definita solo a meno di composizione: usare punto conico fra due specchi.

REMARK 1.2. One should think at the maps $\psi_{ij}$ e $f_{ij}$ as defined only up to the action of $\Gamma_j$ (which acts on $\psi_{ij}$ by composition and on $f_{ij}$ by conjugation). In particolar, if $U_i \subset U_j \subset U_k$ then we can verify that the equalities $\psi_{ik} = \psi_{ij} \circ \psi_{jk}$ and $f_{ik} = f_{jk} \circ f_{ij}$ hold only up to this ambiguity.

The *isotropy group* of a point $x \in O$ is the stabilizer of $x$ with respect to the action of $\Gamma_i$ on any chart $U_i$ containing $x$. By definition $\Gamma_x$ is a finite subgroup of $O(n)$. A point $x$ is *regular* if its isotropy group is trivial, and *singular* otherwise.

EXAMPLE 1.3. A differentiable manifold is an orbifold whose points are all regular. A differentiable manifold with boundary may be interpreted as an orbifold whose boundary points have the local structure of type $\mathbb{R}^n/\Gamma$ where $\Gamma \cong \mathbb{Z}_2$ is generated by a reflection along a hyperplane. The boundary should now be interpreted as a *mirror*.

PROPOSITION 1.4. *The regular points in an orbifold form a dense subset.*

PROOF. A singular point is locally the fixed point locus of a finite group in $O(n)$ and is hence closed and contained in a hyperplane. $\square$

Many notions extend from manifolds to orbifolds.

DEFINITION 1.5. A *continuous map* from a topological space $X$ to an orbifold $O$ is a continuous map $\alpha\colon X \to O$ together with an explicit lift on $V_i$ at every chart $V_i \to V_i/_{\Gamma_i}$; these lift must be compatible: if $U_i \subset U_j$ the two lifts $\alpha_i$ and $\alpha_j$ to $V_i$ and $V_j$ must fulfill $\alpha_i = \alpha_j \circ \psi_{ij}$ through a transition map $\psi_{ij}$ (which we recall is defined only up to post-composing with $\Gamma_j$).

Two maps whose lift to $V_i$ change only by an action of the groups $\Gamma_i$ are considered equivalent.

DEFINITION 1.6. Pick a basepoint $x_0 \in O$. The *fundamental group* $\pi_1(O, x_0)$ is defined as usual with loops and homotopies using the notion of continuous map just introduced.

Vedere se la definizione di rivestimento può essere migliorata

DEFINITION 1.7. A *covering* $\pi\colon \widetilde{O} \to O$ of orbifolds is a continuous map where every point $x \in O$ has a neighborhood $U$ of type $V/_{\Gamma}$ for some $V \subset \mathbb{R}^n$, whose counterimage $\pi^{-1}(U)$ is a disjoint union of pieces of type $V/_{\Gamma_i}$ for some subgroups $\Gamma_i \subset \Gamma$. (We mean here that $U$ is isomorphic to $V/_{\Gamma}$, that $\pi^{-1}(U) = \cup_i U_i$ with $U_i$ isomorphic to $V/_{\Gamma_i}$, and that all isomorphisms commute with $\pi$.)

An orbifold is *good* if it is finitely covered by a manifold. A covering $\pi\colon \widetilde{O} \to O$ is *universal* if it satisfies the following universal property: for any covering $\pi'\colon O' \to O$ there is a covering $q\colon \widetilde{O} \to O'$ such that $\pi = q \circ \pi'$

THEOREM 1.8. *An orbifold $O$ has a universal covering, unique up to isomorphisms. The universal covering is a manifold if a only if $O$ is good.*

As for manifolds, the automorphism group of the universal covering is isomorphic to the fundamental group of the orbifold.

DEFINITION 1.9. The *tangent space* $T_xO$ is an orbifold defined as usual as classes of curves passing through $x$, and is hence isomorphic to $\mathbb{R}^n/_{\Gamma_x}$.

An orbifold is *oriented* if $\Gamma_i < SO(n)$ for all $i$ and the diffeomorphisms $\psi_{ij}$ preserve the orientation of $\mathbb{R}^n$. In an orientable orbifold the reflections are not admitted in $\Gamma_i$ and hence the singular locus has codimension at least 2.

**1.2. Riemannian orbifold.** A *riemannian orbifold* is an orbifold with a positive scalar product on each tangent space $T_x$ which changes smoothly with $x$. That is, there is a $\Gamma_i$-invariant riemannian metric on each $V_i$ and they are compatible.

On a riemannian orbifold we have the notion of length of curves, and hence of distance between points, of geodesic (a curve whose lift in $V_i$ is a geodesic) and volume form (the open dens set formed by regular points is a riemannian manifold).

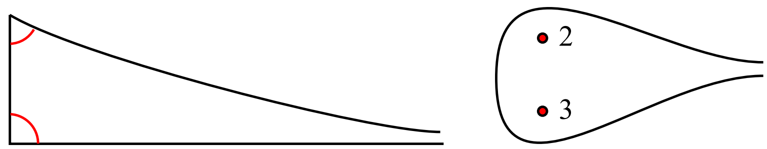



FIGURE 1. The orbifold $\mathbb{H}^2/_{\mathbb{P}\mathrm{GL}_2(\mathbb{Z})}$ has three mirrors, two of which are infinite and converge to a cusp, and two cones of angle $\frac{\pi}{2}$ and $\frac{\pi}{3}$. The orbifold has area $\frac{\pi}{6}$. The index-two subgroup $\mathbb{P}\mathrm{SL}_2(\mathbb{Z}) < \mathbb{P}\mathrm{GL}_2(\mathbb{Z})$ that contains only orientation-preserving transformations produce the orbifold $\mathbb{H}^2/_{\mathbb{P}\mathrm{SL}_2(\mathbb{Z})}$ (right) of area $\frac{\pi}{3}$, obtained by mirroring the one on the left. It has a cusp and two cone points with rotational isotropy $\mathbb{Z}_2$ e $\mathbb{Z}_3$.

PROPOSITION 1.10. *If $M$ a riemannian manifold and $\Gamma < \mathrm{Isom}(M)$ is a discrete subgroup, the quotient $M/_\Gamma$ has a natural orbifold structure.The projection $\pi\colon M \to M/_\Gamma$ is an orbifold covering.*

PROOF. Take a point $x \in M/_\Gamma$ and $\tilde{x} \in M$ a lift. Since $\Gamma$ is discrete, the stabilizer $\Gamma_{\tilde{x}}$ of $\tilde{x}$ is finite and there is $r > 0$ such that $\exp_{\tilde{x}}(B_r(0)) = B_r(\tilde{x})$ and $g(B_r(\tilde{x}))$ intersects $B_r(\tilde{x})$ if and only if $g \in \Gamma_{\tilde{x}}$. The ball $B_r(\tilde{x})$ is clearly $\Gamma_x$-invariant.

We define an orbifold structure on $M/_\Gamma$ by taking for each $x$ the open set $U_x = B_r(\tilde{x})/_{\Gamma_x}$ with chart $V_x = B_r(\tilde{x})$ and finite group $\Gamma_x$ for some $\tilde{x} \in p^{-1}(x)$. We extend the covering $\{U_x\}$ thus obtained by taking all the non-empty intersections. The projection is a covering by construction. $\square$

More generally, if $\Gamma < \Gamma'$ are discrete groups of isometries for some riemannian manifold $M$ then $M/_\Gamma \to M/_{\Gamma'}$ is an orbifold covering. In particular the lattice of discrete groups in $\mathrm{Isom}(M)$ transforms into a lattice of orbifold coverings. A *hyperbolic, flat*, or *elliptic orbifold* is the quotient of $\mathbb{H}^n$ ($\mathbb{R}^n$, $S^n$) by a discrete group $\Gamma$ of isometries.

**1.3. Examples.** The triangular group defines an orbifold which is topolgoically a triangle. By Selberg lemma, these orbifolds are all good. Interior points have trivial isotropy, those on the sides have $\mathbb{Z}_2$ generated by a reflection, the vertices have $\mathbb{Z}_{2a}, \mathbb{Z}_{2b}, \mathbb{Z}_{2c}$ generated by rotations.

The index-two subgroup of the triangular group consisting of orientation-preserving transformations gives an orbifold that double-covers the triangle and consists of a sphere with three points with rotation isotopy groups $\mathbb{Z}_a, \mathbb{Z}_b, \mathbb{Z}_c$. Fare figura con sfera e tre angoli conici.

CHAPTER 7

# Hyperbolic 3-manifolds

We construct here various hyperbolic 3-manifolds.

## 1. Cusped 3-manifolds

**1.1. Ideal tetrahedra.** Ideal tetrahedra play a fundamental role in the construction of hyperbolic 3-manifolds. They can be described up to isometry by a single complex parameter $z \in H^2$, as we now see.

An ideal tetrahedron is determined by its four ideal vertices $v_1, v_2, v_3, v_4 \in \partial\mathbb{H}^3$. We use the half-space model $H^3$ and recall that $\text{Isom}^+(H^3) = \mathbb{P}\text{SL}_2(\mathbb{C})$, hence there is a unique orientation-preserving isometry of $H^3$ that sends $v_1, v_2, v_3, v_4$ respectively to $0, 1, \infty, z$ for some $z$. Up to mirroring with the orientation-reversing reflection $z \mapsto \bar{z}$ we can suppose that $\Im z > 0$.

Remark 1.1. By definition the number $z$ is the *cross-ratio* of the four numbers $v_1, v_2, v_3, v_4$.

A horosphere centered at the vertex at $\infty$ is a euclidean plane and intersects the ideal tetrahedron in a euclidean triangle as in Fig. 1. The triangle

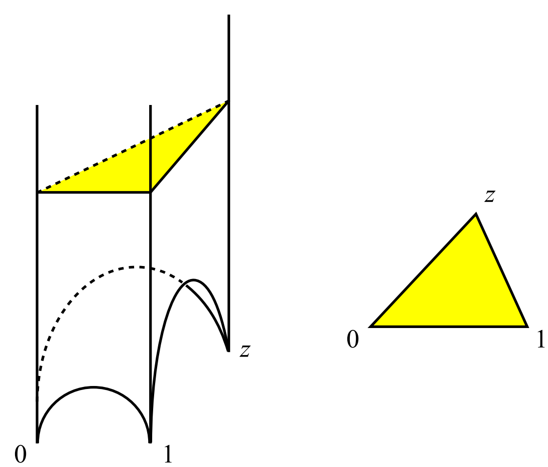



Figure 1. An ideal tetrahedron with three vertices in $0, 1, \infty$ in the half-space mode is determined by the position $z \in \mathbb{C} \cup \{\infty\}$ of the fourth vertex. A little horosphere centered in the ideal vertex intersects the tetrahedron in a euclidean triangle uniquely determined up to similarities.

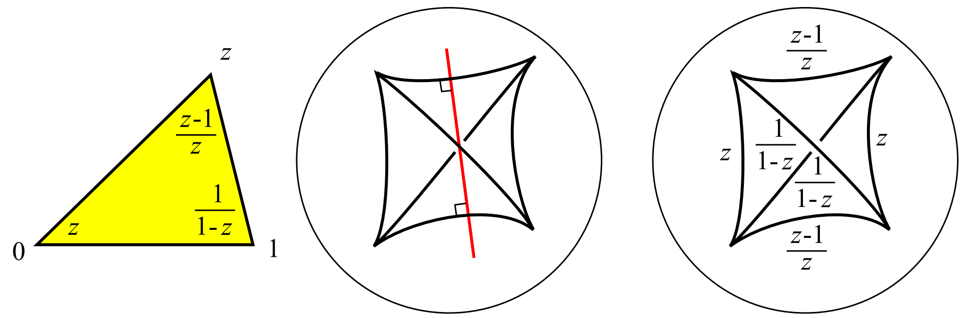



FIGURE 2. At each ideal vertex we have a euclidean triangle defined up to similarities: each vertex of the triangle has a well-defined complex angle (left). A pair of opposite edges in the ideal tetrahedron has an axis orthogonal to both which is a symmetry axis for the tetrahedron (center). We can assign the complex angles directly to the edges of the tetrahedron (right). The argument is the dihedral angle of the edge.

is well-defined up to orientation-preserving similarities since changing the horosphere results only in a dilation. Hence it has itself a unique representation as a triangle in $\mathbb{C} = \mathbb{R}^2$ with vertices at 0, 1, and $z$ as in Fig. 2. The *complex angle* of a vertex of the triangle is the ratio of the two adjacent sides, taken with counterclockwise order and seen as complex numbers. The three complex angles are shown in Fig. 2 and are

$$z, \frac{z-1}{z}, \frac{1}{1-z}.$$

The argument is the usual angle, and the modulus is the ratio of the two lengths of the adjacent sides. We now note that the tetrahedron has some non-trivial symmetries.

PROPOSITION 1.2. *For any pair of opposite edges in an ideal tetrahedron $T$ there is a symmetry axis $r$ orthogonal to both as in Fig. 2-(center) such that $T$ is symmetric with respect to a $\pi$ rotation around $r$.*

PROOF. Since $T$ is non-degenerate the opposite sides $e$ ed $e'$ are ultraparallel lines in $\mathbb{H}^3$ and hence have a common perpendicular $r$. A $\pi$ rotation around $r$ permutes the vertices of $e$ and $e'$ but preserves the 4 ideal vertices of $T$, hence $T$ itself. $\square$

The symmetries of the tetrahedron then act transitively on its vertices. It follows that every vertex has the same triangular section as in Fig. 2-(sinistra), and that all these sections can be recovered by assigning the complex numbers directly to the edges of $T$ as showin in Fig. 2-(right). These labels on the edges determine the ideal tetrahedron up toisometries of $\mathbb{H}^3$.

PROPOSITION 1.3. *The argument of the complex label is the dihedral angle of the edge.*

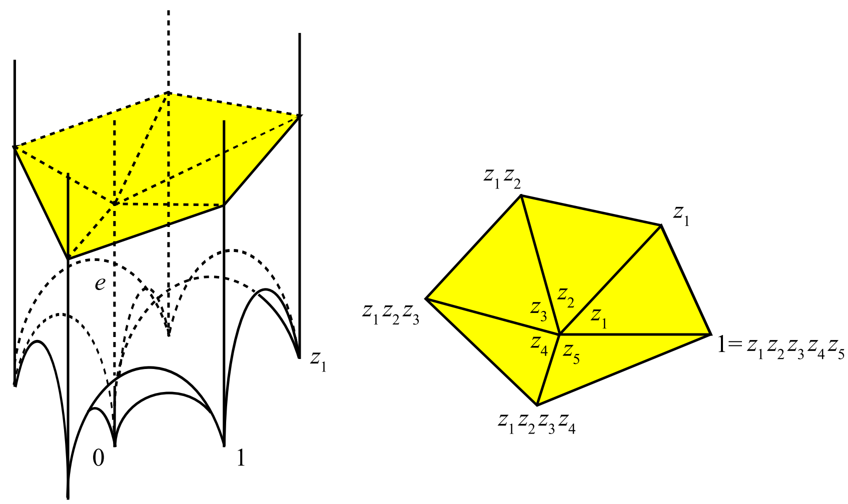



FIGURE 3. If we manage to glue all the tetrahedra incident to an edge $e$ inside $\mathbb{H}^3$ as shown in the left, the hyperbolic structure is defined also in $e$. Let $z_1, \ldots, z_h$ be the complex numbers assigned to the sides of the $h$ incident tetrahedra (here $h = 5$). This can be done if $z_1 \cdots z_h = 1$ and the arguments sum to $2\pi$ (right).

PROOF. The dihedral angle of an edge $e$ in a polyhedron is calculated by intersecting the polyhedron with a hypersurface orthogonal to $e$, and calculate the angle in the two-dimensional picture there: one may take a horocusp as a hypersurface. □

**1.2. Ideal triangulations and completeness equations.** Let now $\Delta_1, \ldots, \Delta_k$ be oriented ideal hyperbolic tetrahedra.

DEFINITION 1.4. a *face-pairing* is a partition of the $4k$ triangular faces of the tetrahedra into pairs, and for each pair a bijection between their triples of ideal vertices.

The face-pairing may be realized isometrically in a unique way in virtue of the following.

PROPOSITION 1.5. *Given two ideal triangles $\Delta$ and $\Delta'$, every bijection between the ideal vertices of $\Delta$ and of $\Delta'$ is realized by a unique isometry.*

PROOF. We see the ideal triangles in $H^2$ and recall that for any two triples of points in $\partial H^2$ there is a unique isometry sending pointwise the first triple to the second. □

The face-pairing is *orientable* if all the resulting isometries between triangles are orientation-reversing. Suppose now that we identify all the tetrahedra by an orientable face-pairing. In the complement of the edges, we have constructed an oriented hyperbolic 3-manifold, and we now try to extend its hyperbolic structure to the edges. We can do this if we manage to glue all the $h$ tetrahedra around an edge $e$ inside $\mathbb{H}^3$ as in Fig. 3. Let $z_1, \ldots, z_h$ be the complex numbers associated to the edges of the $h$ tetrahedra incident to $e$. As shown in the figure, if $z_1 \cdots z_h = 1$ and the sum of their argument is $2\pi$ (and not some other multiple of $2\pi$) then all tetrahedra can be glued simultaneously in $\mathbb{H}^3$ and a hyperbolic structure extends also to $e$.

We can now pick an arbitrary edge for every tetrahedron $\Delta_i$ and assign to it the complex variable $z_i$: the other sides of $\Delta_i$ are automatically labeled by one of the variables $z_i$, $\frac{z_i-1}{z_i}$, or $\frac{1}{1-z_i}$ as in Fig. 2. As we have seen, for every edge we obtain an equation of type $w_1 \cdots w_h = 1$ (to which we must add the condition that the sum of the arguments is $2\pi$), where each $w_j$ equals $z_i$, $\frac{z_i-1}{z_i}$, or $\frac{1}{1-z_i}$ for some $i$.

We have thus obtained a system of equations called *compatibility equations*, with a variable for each tetrahedron and an equation for each edge. Our discussion proves the following. A *positive solution* to these equations is a solution $(z_1, \ldots, z_k)$ where $\Im z_i > 0$ for all $i$.

PROPOSITION 1.6. *A positive solution to the compatibility equation identifies a hyperbolic manifold obtained by gluing the ideal tetrahedra.*

The hyperbolic structure is however not necessarily complete: the manifold is not compact because we are employing ideal tetrahedra. To get a complete hyperbolic manifold we must add some more equations.

### 1.3. Completeness equations.

**1.4. Volumes of ideal tetrahedra.** We will express the volume of an ideal tetrahedron in terms of its complex modulus $z$, using the following function.

DEFINITION 1.7. The *Lobachevsky function* is the function

$$\Lambda(\theta) = -\int_0^\theta \log|2\sin t|dt.$$

Draw function.

The function $\log|2\sin t|$ is $-\infty$ on $\pi\mathbb{Z}$ but is integrable, hence $\Lambda$ is well-defined and continuous on $\mathbb{R}$. Its first derivatives are

$$\Lambda'(\theta) = -\log|2\sin\theta|, \qquad \Lambda''(\theta) = -\cot\theta.$$

The function $\Lambda$ has derivative $+\infty$ on $\pi\mathbb{Z}$ and is an odd function, because its derivative is even.

PROPOSITION 1.8. *The function $\Lambda$ is $\pi$-periodic. We have $\Lambda(0) = \Lambda\left(\frac{\pi}{2}\right) = \Lambda(\pi) = 0$. The function $\Lambda$ is strictly positive on $\left(0, \frac{\pi}{2}\right)$, strictly negative on $\left(\frac{\pi}{2}, \pi\right)$, and has absolute maximum and minimum at $\frac{\pi}{6}$ and $\frac{5}{6}\pi$. For all $m \in \mathbb{N}$ the following holds:*

$$\Lambda(m\theta) = m\sum_{k=0}^{m-1}\Lambda\left(\theta + \frac{k\pi}{m}\right).$$

PROOF. We prove the equality for $m = 2$:

$$\begin{aligned}
\frac{\Lambda(2\theta)}{2} &= -\frac{1}{2}\int_0^{2\theta} \log|2\sin t|dt = -\int_0^{\theta} \log|2\sin 2t|dt \\
&= -\int_0^{\theta} \log|2\sin t|dt - \int_0^{\theta} \log\left|2\sin\left(t+\frac{\pi}{2}\right)\right| dt \\
&= \Lambda(\theta) - \int_{\frac{\pi}{2}}^{\frac{\pi}{2}+\theta} \log|2\sin t|dt \\
&= \Lambda(\theta) + \Lambda\left(\theta+\frac{\pi}{2}\right) - \Lambda\left(\frac{\pi}{2}\right).
\end{aligned}$$

By setting $\theta = \frac{\pi}{2}$ we get $\Lambda(\pi) = 0$. Since the derivative $\Lambda'$ is $\pi$-periodic and $\Lambda(\pi) = 0$, also $\Lambda$ is $\pi$-periodic. Since $\Lambda$ is $\pi$-periodic and odd, we have $\Lambda\left(\frac{\pi}{2}\right) = 0$. We have also proved the formula for $m = 2$.

To prove the formula for generic $m$ we use a generalization of the duplication formula for the sinus. From the equality

$$z^m - 1 = \prod_{k=0}^{m-1}\left(z - e^{-\frac{2\pi i k}{m}}\right)$$

we deduce

$$2\sin(mt) = \prod_{k=0}^{m-1} 2\sin\left(t + \frac{k\pi}{m}\right)$$

and hence

$$\begin{aligned}
\frac{\Lambda(m\theta)}{m} &= -\frac{1}{m}\int_0^{m\theta} \log|2\sin t|dt = -\int_0^{\theta} \log|2\sin(mt)|dt \\
&= -\sum_{k=0}^{m-1}\int_0^{\theta} \log\left|2\sin\left(t+\frac{k\pi}{m}\right)\right| dt \\
&= -\sum_{k=0}^{m-1}\left(\int_0^{\theta+\frac{k\pi}{m}} \log|2\sin t|dt - \int_0^{\frac{k\pi}{m}} \log|2\sin t|dt\right) \\
&= -\sum_{k=0}^{m-1}\Lambda\left(\theta+\frac{k\pi}{m}\right) + C(m)
\end{aligned}$$

where $C(m)$ is a constant independent of $\theta$. By integrating both sides we get

$$\frac{1}{m}\int_0^{\pi}\Lambda(m\theta) = -\sum_{k=0}^{m-1}\int_0^{\pi}\Lambda\left(\theta+\frac{k\pi}{m}\right) + C(m)\pi.$$

Since $\Lambda$ is odd and $\pi$-periodic, we have

$$\int_0^{\pi}\Lambda(m\theta) = 0$$

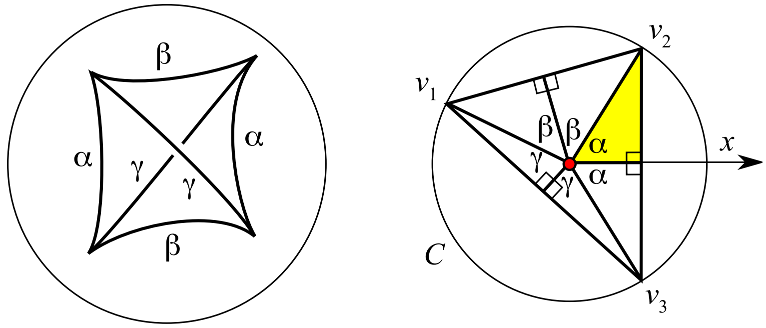



FIGURE 4. The dihedral angles $\alpha, \beta, \gamma$ of an ideal tetrahedron. Opposite edges have the same angle and $\alpha + \beta + \gamma = \pi$.

for any integer $m$. Hence $C(m) = 0$ and the formula is proved. Finally we note that $\Lambda''(\theta) = -\cot\theta$ is strictly negative in $(0, \frac{\pi}{2})$ and strictly positive in $(\frac{\pi}{2}, \pi)$, hence $\Lambda$ is strictly positive in $(0, \frac{\pi}{2})$ and strictly negative in $(\frac{\pi}{2}, \pi)$. $\square$

PROPOSITION 1.9. *An ideal tetrahedron is determined up to isometry by its dihedral angles $\alpha, \beta, \gamma$ as in Fig. 4. The relation $\alpha + \beta + \gamma = \pi$ holds.*

PROOF. An ideal tetrahedron is determined by its complex angles as in Fig. 2, determined by a triangle as in Fig. 2-(left), unique up to similarities. The triangle is also determined by its inner angles. $\square$

The regular ideal tetrahedron has of course equal angles $\alpha = \beta = \gamma = \frac{\pi}{3}$.

THEOREM 1.10. *Let $\Delta$ be an ideal tetrahedron with dihedral angles $\alpha$, $\beta$ and $\gamma$. We have*

$$\mathrm{Vol}(\Delta) = \Lambda(\alpha) + \Lambda(\beta) + \Lambda(\gamma).$$

PROOF. We represent $\Delta$ in $H^3$ with one vertex $v_0$ at infinity and three $v_1, v_2, v_3$ in $\mathbb{C}$. Let $C$ be the circle containing $v_1$, $v_2$ e $v_3$: up to composing with elements in $\mathbb{P}\mathrm{SL}_2(\mathbb{C})$ we can suppose that $C = S^1$. The euclidean triangle $T \subset \mathbb{C}$ with vertices $v_1$, $v_2$ e $v_3$ has interior angles $\alpha$, $\beta$, and $\gamma$.

We first consider the case $0 \in T$, that is $\alpha, \beta, \gamma \leqslant \frac{\pi}{2}$. We decompose $T$ into six triangles as in Fig. 4: the tetrahedron $\Delta$ decomposes accordingly into six tetrahedra lying above them, and we prove that the one $\Delta_\alpha$ lying above the yellow triangle has volume $\frac{\Lambda(\alpha)}{2}$. This proves the theorem.

This tetrahedron is the intersection of four half-spaces: three vertical ones bounded by the hyperplanes $y = 0$, $x = \cos\alpha$, and $y = x\tan\alpha$, and one bounded by the half-sphere $z^2 = x^2 + y^2$. Therefore

$$\begin{aligned}
\mathrm{Vol}(\Delta_\alpha) &= \int_0^{\cos\alpha} dx \int_0^{x\tan\alpha} dy \int_{\sqrt{1-x^2-y^2}}^{\infty} \frac{1}{z^3} dz \\
&= \int_0^{\cos\alpha} dx \int_0^{x\tan\alpha} dy \left[-\frac{1}{2z^2}\right]_{\sqrt{1-x^2-y^2}}^{\infty} \\
&= \frac{1}{2}\int_0^{\cos\alpha} dx \int_0^{x\tan\alpha} \frac{1}{1-x^2-y^2} dy.
\end{aligned}$$

To solve this integral we use the relation

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

and hence $\mathrm{Vol}(\Delta_\alpha)$ equals

$$\begin{aligned}
&\frac{1}{4}\int_0^{\cos\alpha}\frac{dx}{\sqrt{1-x^2}}\left(\Big[-\log(\sqrt{1-x^2}-y)\Big]_0^{x\tan\alpha} + \Big[\log(\sqrt{1-x^2}+y)\Big]_0^{x\tan\alpha}\right) \\
=&\frac{1}{4}\int_0^{\cos\alpha}\frac{dx}{\sqrt{1-x^2}}\Big(-\log(\sqrt{1-x^2}-x\tan\alpha) + \log(\sqrt{1-x^2}+x\tan\alpha)\Big).
\end{aligned}$$

By writing $x = \cos t$ and hence $dx = -\sin t\, dt$ we obtain

$$\begin{aligned}
\mathrm{Vol}(\Delta_\alpha) &= \frac{1}{4}\int_{\frac{\pi}{2}}^{\alpha}\frac{-\sin t}{\sin t}\left(-\log\frac{\sin t\cos\alpha - \cos t\sin\alpha}{\sin t\cos\alpha + \cos t\sin\alpha}\right)dt \\
&= -\frac{1}{4}\int_{\frac{\pi}{2}}^{\alpha}\log\frac{\sin(t+\alpha)}{\sin(t-\alpha)}dt = -\frac{1}{4}\int_{\frac{\pi}{2}}^{\alpha}\log\frac{|2\sin(t+\alpha)|}{|2\sin(t-\alpha)|}dt \\
&= \frac{1}{4}\int_{2\alpha}^{\frac{\pi}{2}+\alpha}\log|2\sin t|dt - \frac{1}{4}\int_0^{\frac{\pi}{2}-\alpha}\log|2\sin t|dt \\
&= \frac{1}{4}\left(-\Lambda\left(\frac{\pi}{2}+\alpha\right) + \Lambda(2\alpha) + \Lambda\left(\frac{\pi}{2}-\alpha\right)\right) \\
&= \frac{1}{4}\left(-\Lambda\left(\frac{\pi}{2}+\alpha\right) + 2\Lambda(\alpha) + 2\Lambda\left(\frac{\pi}{2}+\alpha\right) - \Lambda\left(\frac{\pi}{2}+\alpha\right)\right) = \frac{1}{2}\Lambda(\alpha)
\end{aligned}$$

using Proposition 1.8.

If $0 \notin T$ the triangle $T$ may be decomposed analogously into triangles, some of which contribute negatively to the volume, and we obtain the same formula. □

COROLLARY 1.11. *The regular ideal tetrahedron is the hyperbolic tetrahedron of maximum volume.*

PROOF. It is easy to prove that every hyperbolic tetrahedron is contained in an ideal tetrahedron: hence we may consider only ideal tetrahedra. Consider the triangle $T = \{0 \leqslant \alpha, \beta, \alpha+\beta \leqslant \pi\}$ and

$$\begin{aligned}
f\colon \qquad T &\longrightarrow \mathbb{R} \\
(\alpha,\beta) &\longmapsto \Lambda(\alpha) + \Lambda(\beta) + \Lambda(\pi-\alpha-\beta).
\end{aligned}$$

The continuous function $f$ is null on $\partial T$ and strictly positive on the interior of $T$ because it measures the volume of the ideal tetrahedron of dihedral angles $\alpha, \beta, \gamma = \pi - \alpha - \beta$. Hence $f$ has at least a maximum on some interior point $(\alpha, \beta)$. The gradient

$$\nabla f = \begin{pmatrix} \Lambda'(\alpha) - \Lambda'(\pi - \alpha - \beta) \\ \Lambda'(\beta) - \Lambda'(\pi - \alpha - \beta) \end{pmatrix} = \begin{pmatrix} -\log|2\sin\alpha| + \log|2\sin(\pi - \alpha - \beta)| \\ -\log|2\sin\beta| + \log|2\sin(\pi - \alpha - \beta))\end{pmatrix}$$

must vanish there, and this holds if and only if $\sin\alpha = \sin(\pi-\alpha-\beta) = \sin\beta$, *i.e.* if and only if the tetrahedron has all dihedral angles equal to $\frac{\pi}{3}$. $\square$

CHAPTER 8

# Mostow rigidity theorem

We have defined in Chapter 5 the Teichmüller space $\mathrm{Teich}(\Sigma_g)$ as the space of all hyperbolic metrics on $\Sigma_g$, seen up to isometries isotopic to the identity; we have then proved that $\mathrm{Teich}(\Sigma_g) \cong \mathbb{R}^{6g-6}$.

This definition of $\mathrm{Teich}(M)$ extends to any closed differentiable manifold $M$: we show here that if $\dim M \geqslant 3$ then $\mathrm{Teich}(M)$ is either empty or consists of a single point. This strong result is known as *Mostow rigidity.* Thanks to this theorem, every geometric information of a hyperbolic manifold $M$ of dimension $\geqslant 3$ (volume, geodesic spectrum, etc.) is actually a topological invariant of $M$. In its stronget version, Mostow rigidity says that it depends only on $\pi_1(M)$.

## 1. Simplicial volume

**1.1. Generalities.** Gromov has introduced a measure of "volume" of a closed manifold $M$ which makes use only of the homology of $M$. Quite surprisingly, this notion of volume coincides (up to a factor) with the riemannian one when $M$ is hyperbolic.

Let $X$ be a topological space and $R$ a ring. Recall that a *singular $k$-simplex* is a continuous map $\alpha\colon \Delta_k \to X$ from the standard $k$-dimensional simplex $\Delta_k$ in $X$. A *$k$-chain* is an abstract linear combination

$$\lambda_1\alpha_1 + \ldots + \lambda_h\alpha_h$$

of singular $k$-simplexes $\alpha_1, \ldots, \alpha_h$ with coefficients $\lambda_1, \ldots, \lambda_h \in R$. The set $C_k(X,R)$ of all $k$-chains is a $R$-module. There is a linear boundary map $\partial_k\colon C_k(X,R) \to C_{k-1}(X,R)$ such that $\partial_{k-1} \circ \partial_k = 0$. The *cycles* and *boundaries* are the elements of the submodules

$$Z_k(X,R) = \ker \partial_k, \quad B_k(X,R) = \mathrm{Im}\,\partial_{k+1}.$$

The $k$-th *homology group* is the quotient

$$H_k(X,R) = Z_k(X,R)/_{B_k(X,R)}.$$

Consider now the case $A = \mathbb{R}$. We define the *norm* of a cycle $\alpha = \lambda_1\alpha_1 + \ldots + \lambda_h\alpha_h$ as follows:

$$|\alpha| = |\lambda_1| + \ldots + |\lambda_h|.$$

DEFINITION 1.1. The *norm* of a class $a \in H_k(X,\mathbb{R})$ is the infimum of the norms of its elements:

$$|a| = \inf\big\{|\alpha| \bigm| \alpha \in Z_k(X,\mathbb{R}), [\alpha] = a\big\}.$$

Recall that a *seminorm* on a real vector space $V$ is a map $|\cdot|\colon V \to \mathbb{R}_{\geqslant 0}$ such that

- $|\lambda v| = |\lambda||v|$ for any scalar $\lambda \in \mathbb{R}$ and vector $v \in V$,
- $|v + w| \leqslant |v| + |w|$ for any pair of vectors $v, w \in V$.

A *norm* is a seminorm where $|v| = 0$ implies $v = 0$. The following is immediate.

Proposition 1.2. *The norm $|\cdot|$ induces a seminorm on $H_k(X, \mathbb{R})$.*

Although it is only a seminorm, the function $|\cdot|$ is called a norm for simplicity. Let now $M$ be an oriented closed connected manifold. We know that $H_n(M, \mathbb{Z}) \cong \mathbb{Z}$ and the orientation of $M$ determines one of the two generators of $H_n(M, \mathbb{Z})$, called *fundamental class* and denoted by $[M]$. Moreover $H_n(M, \mathbb{R}) \cong \mathbb{R}$ and there is a natural inclusion

$$\mathbb{Z} \cong H_n(M, \mathbb{Z}) \hookrightarrow H_n(M, \mathbb{R}) \cong \mathbb{R}$$

hence the fundamental class $[M]$ is also an element of $H_n(M, \mathbb{R})$ and has a norm.

Definition 1.3. The *simplicial volume* $\|M\| \in \mathbb{R}_{\geqslant 0}$ of a closed oriented connected $M$ is the norm of its fundamental class:

$$\|M\| = |[M]|$$

Since $|[M]| = |-[M]|$ the simplicial volume actually does not depend on the orientation. When $M$ is non-orientable we set $\|M\| = \|\widetilde{M}\|/2$ where $\widetilde{M}$ is the orientable 2-cover of $M$. The definition of $\|M\|$ is relatively simple but has various non-obvious consequences.

A continuous map $f\colon M \to N$ between closed oriented $n$-manifolds induces a homomorphism $f_*\colon H_n(M, \mathbb{Z}) \to H_n(N, \mathbb{Z})$. Recall that the *degree* of $f$ is the integer $\deg f$ such that

$$f_*([M]) = \deg f \cdot [N].$$

Proposition 1.4. *Let $f\colon M \to N$ be a continuous map between closed oriented manifolds. The following inequality holds:*

$$\|M\| \geqslant |\deg f| \cdot \|N\|.$$

Proof. Every description of $[M]$ has a cycle $\lambda_1\alpha_1 + \ldots \lambda_h\alpha_h$ induces a description of $f_*([M]) = \deg f[M]$ as a cycle $\lambda_1 f \circ \alpha_1 + \ldots \lambda_h f \circ \alpha_h$ with the same norm. □

Corollary 1.5. *If $M$ and $N$ are closed orientable and homotopically equivalent $n$-manifolds then $\|M\| = \|N\|$.*

Proof. A homotopic equivalence consists of two maps $f\colon M \to N$ and $g\colon N \to M$ whose compositions are both homotopic to the identiy. In particular both $f$ and $g$ have degree $\pm 1$. □

Corollary 1.6. *If $M$ admits a continuous self-map $f\colon M \to M$ of degree $\geqslant 2$ then $\|M\| = 0$.*

Corollary 1.7. *A sphere $S^n$ has norm zero. More generally we have $\|M \times S^n\| = 0$ for any $M$ and any $n \geqslant 1$.*

Proof. A sphere $S^n$ admits self-maps of non-zero degree, which extend to $M \times S^n$. □

Among the surfaces $S_g$, the sphere and the torus have hence simplicial volume zero. We will see soon that every surface of genus $g \geqslant 2$ has positive simplicial volume. When the continuous map is a covering the inequality from Proposition 1.4 promotes to an equality.

Proposition 1.8. *If $f\colon M \to N$ is a covering of degree $g$ we have*

$$\|M\| = d \cdot \|N\|.$$

Proof. The reason for this equality is that cycles can be lifted and projected along the covering. More precisely, we already know that $\|M\| \geqslant d \cdot \|N\|$. Conversely, let $\alpha = \lambda_1\alpha_1 + \ldots + \lambda_h\alpha_h$ represent $[N]$; each $\alpha_i$ is a map $\Delta_n \to N$. Since $\Delta_n$ is simply connected, the map $\alpha_i$ lfts to $d$ distinct maps $\alpha_i^1, \ldots, \alpha_i^d\colon \Delta_n \to N$. The chain $\tilde\alpha = \sum_{ij} \lambda_i\alpha_i^j$ is a cycle in $M$ and $f_*(\tilde\alpha) = d\alpha$. Hence $\|M\| \leqslant d \cdot \|N\|$. □

**1.2. Simplicial and hyperbolic volume.** In the next pages we will prove the following theorem only for the dimension $n = 3$ which is of interest for us. Let $v_n$ be the volume of the regular ideal $n$-simplex in $\mathbb{H}^n$.

Theorem 1.9. *Let $M$ be a closed hyperbolic $n$-manifold. We have*

$$\mathrm{Vol}(M) = v_n\|M\|.$$

The theorem furnishes in particular some examples of manifolds with positive simplicial volume and shows that $\mathrm{Vol}(M)$ depends only on the topology of $M$, thus generalizing Gauss-Bonnet theorem to all dimensions. Mostow rigidity will then strengthen this result in dimension $n \geqslant 3$, showing that the hyperbolic metric itself depends only on the topology.

Both quantities $\mathrm{Vol}(M)$ and $\|M\|$ are multiplied by $d$ if we substitute $M$ with a degree-$d$ covering. In particular, up to substituting $M$ with its orientable 2-cover, we can suppose that $M$ is orientable.

The proof for general $n$ makes use of the following result:

Theorem 1.10. *The regular ideal simplex is the simplex of maximum volume among all simplexes in $\mathbb{H}^n$.*

For the sake of clarity we will prove Theorem 1.9 assuming this result: our proof is however complete only for $n = 2, 3$, where Theorem 1.10 was proved in Chapter 7 as Corollary 1.11 for $n = 3$ and follows from the formula for the area of a triangle for $n = 2$.

**1.3. Cycle straightening.** The *straight singular $k$-simplex* with vertices $v_1, \ldots, v_{k+1} \in I^n$ is the map

$$\begin{aligned} \alpha\colon \quad \Delta_k &\longrightarrow \mathbb{H}^n \\ (t_1, \ldots, t_n) &\longmapsto t_1 v_1 + \ldots + t_{k+1} v_{k+1} \end{aligned}$$

defined using convex combinations. If the $k+1$ vertices are not contained in a $(k-1)$-plane the singular $k$-simplex is *non-degenerate* and its image is a hyperbolic $k$-simplex.

The *straightening* $\alpha^{\mathrm{st}}$ of a singular simplex $\alpha\colon \Delta_k \to \mathbb{H}^n$ is the straight singular simplex with the same vertices $\alpha(e_1), \ldots, \alpha(e_k)$. The straightening $\alpha^{\mathrm{st}}$ of a singular simplex $\alpha\colon \Delta_k \to M$ in a hyperbolic manifold $M = \mathbb{H}^n/_{\Gamma}$ is defined by lifting the singular simplex in $\mathbb{H}^n$, straightening it, and projecting it back to $M$ by composition with the covering map. Different lifts produce the same straightening in $M$ because they are related by isometries of $\mathbb{H}^n$.

The straightening extends by linearity to a homomorphism

$$\mathrm{st}\colon C_k(M, R) \to C_k(M, R)$$

which commutes with $\partial$ and hence induces a homomorphism in homology

$$\mathrm{st}_*\colon H_k(M, R) \to H_k(M, R).$$

PROPOSITION 1.11. *The map* $\mathrm{st}_*$ *is the identity.*

PROOF. We may define a homotopy between a singular simplex $\sigma$ and its straightening $\sigma^{\mathrm{st}}$ using the convex combination

$$\sigma^t(x) = t\sigma(x) + (1-t)\sigma^{\mathrm{st}}(x).$$

This defines a chain homotopy between $\mathrm{st}_*$ and id via the same technique used to prove that homotopic maps induce the same maps in homology. $\square$

The *abstract volume* of a straightened singular simplex $\alpha\colon \Delta_n \to M$ is the volume of its lift in $\mathbb{H}^n$ and may also be calculated as

$$\left| \int_\alpha \omega \right|$$

where $\omega$ is the volume form on $M$ pulled back along $\alpha$. The abstract volume is smaller than $v_n$. If $\alpha$ is non-degenerate, its *sign* is *positive* if $\alpha$ is orientation-preserving and *negative* otherwise: equivalently, it is the sign of $\int_\alpha \omega$.

We can prove one inequality.

PROPOSITION 1.12. *Let $M$ be a closed hyperbolic $n$-manifold. We have*

$$\mathrm{Vol}(M) \leqslant v_n \|M\|.$$

PROOF. As we said above, we can suppose $M$ is orientable. Take a cycle $\alpha = \lambda_1\alpha_1 + \ldots + \lambda_k\alpha_k$ that represents $[M]$. By straightening it we get another cycle $\alpha^{\rm st} = \lambda_1\alpha_1^{\rm st} + \ldots + \lambda_k\alpha_k^{\rm st}$ that represents $[M]$. Let $\omega$ be the volume form on $M$. We get

$$\mathrm{Vol}(M) = \int_M \omega = \int_\alpha \omega = \lambda_1 \int_{\alpha_1} \omega + \ldots + \lambda_k \int_{\alpha_k} \omega.$$

The quantity $\left|\int_{\alpha_i}\omega\right|$ is the abstract volume of $\alpha_i$. Hence $\left|\int_{\alpha_i}\omega\right| < v_n$ and

$$\mathrm{Vol}(M) < \big(|\lambda_1| + \ldots + |\lambda_k|\big)v_n.$$

This holds for all $\alpha$, hence $\mathrm{Vol}(M) \leqslant v_n\|M\|$. □

**1.4. Efficient cycles.** Let $M = \mathbb{H}^n/_\Gamma$ be a closed oriented hyperbolic manifold. An *$\varepsilon$-efficient cycle* for $M$ is a straightened cycle

$$\alpha = \lambda_1\alpha_1 + \ldots \lambda_k\alpha_k$$

representing $[M]$ where the abstract volume of $\alpha_i$ if bigger than $v_n - \varepsilon$ and the sign of $\alpha_i$ is coherent with the sign of $\lambda_i$, for all $i$.

We will construct an $\varepsilon$-efficient cycle for every $\varepsilon > 0$. This will conclude the proof of Theorem 1.9 in virtue of the following:

Lemma 1.13. *If for any $\varepsilon > 0$ the manifold $M$ admits an $\varepsilon$-efficient cycle we have* $\mathrm{Vol}(M) \geqslant v_n\|M\|$.

Proof. Let $\alpha = \lambda_1\alpha_1 + \ldots \lambda_k\alpha_k$ be an $\varepsilon$-efficient cycle and $\omega$ be the volume form on $M$. Coherence of signs gives $\lambda_i \int_{\alpha_i}\omega > 0$ for all $i$. We get

$$\begin{aligned}\mathrm{Vol}(M) &= \int_M \omega = \int_\alpha \omega = \lambda_1 \int_{\alpha_1} \omega + \ldots + \lambda_k \int_{\alpha_k} \omega \\ &\geqslant \big(|\lambda_1| + \ldots + |\lambda_k|\big)\cdot(v_n - \varepsilon).\end{aligned}$$

Therefore $\mathrm{Vol}(M) \geqslant \|M\|\cdot(v_n - \varepsilon)$ for all $\varepsilon > 0$. □

It remains to construct $\varepsilon$-efficient cycles.

Proposition 1.14. *If $\Delta^i$ is a sequence of simplexes in $\mathbb{H}^n$ whose vertices tend to the vertices of a regular ideal simplex in $\partial\mathbb{H}^n$, then*

$$\mathrm{Vol}(\Delta^i) \to v_n.$$

Da dimostrare, anche solo per $n = 2, 3$.

For any $t > 0$, let $\Delta(t)$ be a regular simplex obtained as in Section 1.10 from Chapter 3 as follows. Pick a point $x \in \mathbb{H}^n$ and a regular simplex in the euclidean $T_x$, centered at the origin with vertices ot distance $t$ from it, and project the vertices in $\mathbb{H}^n$ using the exponential map.

A *$t$-simplex* is a simplex isometric to $\Delta(t)$ equipped with an ordering of its vertices. The ordering allows to consider it as a straigthened singular simplex. Let $S(t)$ be the set of all $t$-simplexes in $\mathbb{H}^n$.

Exercise 1.15. The group $\mathrm{Isom}(\mathbb{H}^n)$ acts on $S(t)$ freely and transitively.

Mettere da qualche parte che $\mathrm{Isom}(\mathbb{H}^n)$ è unimodulare.

Therefore the Haar measure on $\mathrm{Isom}(\mathbb{H}^n)$ induces a measure on $S(t)$ invariant by the action of $\mathrm{Isom}(\mathbb{H}^n)$.

Let $M = \mathbb{H}^n/_\Gamma$ be a closed hyperbolic manifold and $\pi\colon \mathbb{H}^n \to M$ the covering projection. Fix a point $x_0 \in \mathbb{H}^n$ and consider its orbit $O = \Gamma x_0$. Consider the set

$$\Sigma = \Gamma^{n+1}/\Gamma$$

of the $(n+1)$-uples $(g_0, \ldots, g_n)$ seen up to the diagonal action of $\Gamma$:

$$g \cdot (g_0, \ldots, g_n) = (gg_0, \ldots, gg_n).$$

An element $\sigma = (g_0, \ldots, g_n) \in S$ determines a singular simplex $\widetilde{\sigma}$ in $\mathbb{H}^n$ with vertices $g_0(x_0), \ldots, g_n(x_0) \in O$ only up to translations by $g \in \Gamma$, hence gives a well-defined singular simplex $\sigma = \pi \circ \widetilde{\sigma}_s$ in $M$, which we still denote by $\sigma$. We now introduce the chain

$$\alpha(t) = \sum_{\sigma \in \Sigma} \lambda_\sigma(t) \cdot \sigma$$

for some appropriate real coefficients $\lambda_\sigma(t)$ that we now define. Recall that $x$ determines the Dirichlet tessellation of $\mathbb{H}^n$ into domains $D(g(x))$, $g \in \Gamma$. For $\sigma = (g_0, \ldots, g_n)$ we let $S^+_\sigma(t) \subset S(t)$ be the set of all positive $t$-simplexes whose $i$-th vertex lies in $D(g_i(x))$ for all $i$. The number $\lambda^+_\sigma(t)$ is the measure of $S^+_\sigma(t)$. We define analogously $\lambda^-_\sigma(t)$ and set

$$\lambda_\sigma(t) = \lambda_\sigma(t)^+ - \lambda_\sigma(t)^-.$$

Lemma 1.16. *The chain $\alpha(t)$ has finitely many addenda and is a cycle. If $t$ is sufficiently big the cycle $\alpha(t)$ represents a positive multiple of $[M]$ in $H_n(M, \mathbb{R})$.*

Proof. We prove that the sum is finite. Let $d, T$ be the diameters of $D(x)$ and of a $t$-simplex. We write $\sigma = (\mathrm{id}, g_1, \ldots, g_n)$ for all $\sigma \in \Sigma$: that is, all simplexes have their first vertex at $x$. If $\lambda_\sigma(t) \neq 0$ then $d(g_i x, x) < 2d + T$ for all $i$: therefore $\alpha(t)$ has finitely many addenda (because $O$ is discrete).

We prove that $\alpha(t)$ is a cycle. The boundary $\partial\alpha(t)$ is a linear combination of straight $(n-1)$-simplexes with vertices in $(g_0 x, g_1 x, \ldots, g_{n-1} x)$ as $g_1, \ldots, g_{n-1}$ varies. The coefficient one such $(n-1)$-simplex is

$$\sum_{j=0}^{n} (-1)^{n-j} \sum_{g \in \Gamma} \lambda_{(g_0, \ldots, g_{j-1}, g, g_j, \ldots, g_{n-1})}(t).$$

We prove that each addendum in the sum over $j$ is zero; for simplicity we take the case $j = n$ and get

$$\sum_{g \in \Gamma} \lambda_{(g_0, \ldots, g_{n-1}, g)}(t) = \sum_{g \in \Gamma} \lambda_{(g_0, \ldots, g_{n-1}, g)}(t)^+ - \sum_{g \in \Gamma} \lambda_{(g_0, \ldots, g_{n-1}, g)}(t)^-.$$

The first addendum measures the positive $t$-simplexes whose first $n$ vertices lie in $D(g_0(x)), \ldots, D(g_n(x))$, the second measures the negative $t$-simplexes with the same requirement. These two subsets have the same volume in

$S(t)$ because they are related by the involution $r\colon S(t) \to S(t)$ that mirrors a simplex with respect to its first facet.

We show that for sufficiently big $t$ the cycle is a positive multiple of $[M]$. Let $t$ be sufficiently big so that two vertices in a $t$-simplex have distance bigger than $2d$. This condition implies that if there is a positive $t$-simplex with vertices in $D(g_0(x)), \ldots, D(g_n(x))$, then any straight simplex with vertices in $D(g_0(x)), \ldots, D(g_n(x))$ is positive. Therefore in the expression

$$\alpha(t) = \sum_{\sigma\in\Sigma} \lambda_\sigma(t) \cdot \sigma$$

the signs of $\lambda_\sigma(t)$ and $\sigma$ are coherent and

$$\int_{\alpha(t)} \omega = \sum_{\sigma\in\Sigma} \lambda_\sigma(t) \cdot \int_\sigma \omega > 0.$$

Therefore $\alpha(t)$ is a positive multiple of $[M]$. □

For sufficiently big $t$ we have $\alpha(t) = k_t[M]$ in homology for some $k_t > 0$. The rescaled $\bar\alpha(t) = \alpha(t)/k_t$ hence represents $[M]$. We have found our $\varepsilon$-efficient cycles.

LEMMA 1.17. *For any $\varepsilon > 0$ there is a $t_0 > 0$ such that $\bar\alpha(t)$ is $\varepsilon$-efficient for all $t > t_0$.*

PROOF. Let $d$ be the diameter of the Dirichlet domain $D(x)$. Let a *quasi $t$-simplex* be a simplex whose vertices are at distance $< d$ from those of a $t$-simplex. By construction $\bar\alpha(t)$ is a linear combination of quasi $t$-simplexes.

We now show that for any $\varepsilon > 0$ there is a $t_0 > 0$ such that for all $t > t_0$ every quasi $t$-simplex has volume bigger than $v_n - \varepsilon$. By contradiction, let $\Delta^t$ be a sequence of quasi $t$-simplexes of volume smaller than $v_n - \varepsilon$ with $t \to \infty$. The vertices of $\Delta^t$ are $d$-closed to a $t$-simplex $\Delta^t_*$, and we move the pair $\Delta^t, \Delta^t_*$ isometrically so that the $t$-simplexes $\Delta^t_*$ have the same barycenter. Now both the vertices of $\Delta^t$ and $\Delta^t_*$ tend to the vertices of an ideal regular simplex and Proposition 1.14 gives a contradiction. □

The previous lemmas together prove the second half of Theorem 1.9.

COROLLARY 1.18. *Let $M$ be a closed hyperbolic $n$-manifold. We have*

$$\mathrm{Vol}(M) \geqslant v_n \|M\|.$$

Theorem 1.9 has some non-trivial consequences.

COROLLARY 1.19. *Let $M, N$ be closed orientable hyperbolic $n$-manifolds If there is a map $f\colon M \to N$ of degree $d$ then $\mathrm{Vol}(M) \geqslant |d| \cdot \mathrm{Vol}(N)$.*

In particular, if there is a map $f\colon \Sigma \to \Sigma'$ of degree $d$ between closed orientable surfaces of genus $g \geqslant 2$ then $-\chi(\Sigma) \geqslant -d \cdot \chi(\Sigma')$.

COROLLARY 1.20. *Two homotopically equivalent closed hyperbolic manifold have the same volume.*

We now strengthen the last corollary in dimension $n \geqslant 3$.

## 2. Mostow rigidity

**2.1. Introduction.** We want to prove the following.

THEOREM 2.1 (Mostow rigidity). *Let $M$ and $N$ be two closed connected orientable hyperbolic manifolds of dimension $n \geqslant 3$. Every isomorfism $\pi_1(M) \stackrel{\sim}{\to} \pi_1(N)$ between fundamental groups is induced by a unique isometry $M \stackrel{\sim}{\to} N$.*

To estimate how powerful is this theorem, note the following chain of implications:

$$\text{isometry} \Longrightarrow \text{diffeo} \Longrightarrow \text{homeo} \Longrightarrow \begin{array}{c}\text{homotopic}\\ \text{equivalence}\end{array} \Longrightarrow \begin{array}{c}\text{isomorphism}\\ \text{on } \pi_1\end{array}$$

Such implications cannot be reversed in general:

- two riemannian diffeomorphic manifolds are non isometric in general, even if they have constant curvature: consider for instance hyperbolic surfaces, or flat $n$-tori;
- in dimension 2 and 3 indeed a homeomorphism implies a diffeomorphism, but this is false in dimension 4, where a closed topological manifold like the $K3$ surface can have infinitely many non-equivalent smooth structures; it is also false in higher dimensions: sometimes a manifold homeomorphic to $S^n$ may not be diffeomorphic to it, starting from $n = 7$;
- the lens spaces $L(7, 1)$ e $L(7, 2)$ are homotopically equivalent but not homeomorphic closed 3-manifolds;
- the closed 4-manifolds $S^4$, $\mathbb{CP}^2$, and $S^2 \times S^2$ are simply connected but non homotopically equivalent because their second homology group is respectively $\{e\}$, $\mathbb{Z}$, and $\mathbb{Z}^2$.

Closed hyperbolic manifolds are aspherical because their universal covering $\mathbb{H}^n$ is contractible. For such manifolds every isomorphism $\pi_1(M) \to \pi_1(N)$ is induced by a homotopy equivalence, unique up to homotopy: see Corollary 4.4 from Chapter 1. To prove Mostow theorem we need to promote this homotopy equivalence to an isometry: we already know that $\mathrm{Vol}(M) = \mathrm{Vol}(N)$ by Corollary 1.20.

**2.2. Quasi and pseudo-isometries.** We introduce the following.

DEFINITION 2.2. A map $F\colon X \to Y$ between metric spaces is a *quasi-isometry* if there are two constants $C_1 > 0$, $C_2 \geqslant 0$ such that

$$\frac{1}{C_1} d(x_1, x_2) - C_2 \leqslant d(F(x_1), F(x_2)) \leqslant C_1 d(x_1, x_2) + C_2$$

for all $x_1, x_2 \in X$ and if $d(F(X), y) \leqslant C_2$ for all $y \in Y$.

A quasi-isometry is an isometry up to some error: note that $F$ may neither be continuous nor injective. Two metric spaces are *quasi-isometric* if there is a quasi-isometry $F\colon X \to Y$ (which implies the existence of a quasi-isometry $G\colon Y \to X$) and quasi-isometry is an equivalence relation between metric spaces. Intuitively, looking at a space up to quasi-isometries is like watching it from some distance: compact metric spaces are obviously quasi-isometric to a point.

This notion is an important ingredient in *geometric group theory*: one may for instance give any finitely-presented group $G$ a canonical metric (through a *Cayley graph*), uniquely determined up to quasi-isometries.

Let $f\colon M \to N$ be a homotopic equivalence of closed hyperbolic $n$-manifolds. Every continuous function is homotopic to a smooth one, hence we suppose that $f$ is smooth. The map lifts to a map $\widetilde{f}\colon \mathbb{H}^n \to \mathbb{H}^n$. We will prove that $\widetilde{f}$ is a quasi-isometry. Actually, the map $\widetilde{f}$ is also continuous and Lipschitz: it will be useful for us to retain this information on $\widetilde{f}$ to simplify some arguments and we hence introduce the following strengthened (but less natural) version of a quasi-isometry:

Definition 2.3. A map $F\colon X \to Y$ between metric spaces is a *pseudo-isometry* if there are two positive constants $C_1, C_2 > 0$ such that

$$\frac{1}{C_1} d(x_1, x_2) - C_2 \leqslant d(F(x_1), F(x_2)) \leqslant C_1 d(x_1, x_2)$$

for any $x_1, x_2 \in X$.

In particular a pseudo-isometry is $C_1$-Lipschitz and hence continuous. Let $f\colon M \to N$ be a smooth map between riemannian $n$-manifolds; the *maximum dilatation* of $f$ at a point $x \in M$ is the maximum ratio $\frac{|df_x(v)|}{|v|}$ where $v$ varies among the unitary vectors $T_x$. The maximum dilatation of $f$ is the supremum of all maximum dilatations as $x \in M$ varies.

Exercise 2.4. If $f\colon M \to N$ has maximum dilatation $C$ the map $f$ is $C$-Lipschitz.

Proposition 2.5. *Let $f\colon M \to N$ be a smooth homotopy between closed hyperbolic $n$-manifolds. The lift $\widetilde{f}\colon \mathbb{H}^n \to \mathbb{H}^n$ is a pseudo-isometry.*

Proof. Since $M$ is compact, the map $f$ has finite maximum dilatation $C$. Since $\widetilde{f}$ is locally like $f$, it also has maximum dilatation $C$ and is hence $C$-Lipschitz. The same holds for the homotopic inverse $g\colon N \to M$. Therefore there is a $C_1 > 0$ such that

$$\begin{aligned} d\big(\widetilde{f}(x_1), \widetilde{f}(x_2)\big) &\leqslant C_1 \cdot d(x_1, x_2) \quad \forall x_1, x_2 \in \mathbb{H}^n, \\ d\big(\widetilde{g}(y_1), \widetilde{g}(y_2)\big) &\leqslant C_1 \cdot d(y_1, y_2) \quad \forall y_1, y_2 \in \mathbb{H}^n. \end{aligned}$$

Being a composition of lifts, the map $\widetilde{g} \circ \widetilde{f}$ commutes with $\Gamma$ and has maximum displacement bounded by some $K > 0$, equal to the maximum displacement of the points belonging to a (compact) Dirichlet domain. Hence

$$d(x_1, x_2) - 2K \leqslant d\big(\widetilde{g}(\widetilde{f}(x_1)), \widetilde{g}(\widetilde{f}(x_2))\big) \leqslant C_1 \cdot d\big(\widetilde{f}(x_1), \widetilde{f}(x_2)\big)$$

for all $x_1, x_2 \in \mathbb{H}^n$. Therefore $\widetilde{f}$ is a pseudo-isometry with $C_2 = 2K/C_1$. $\square$

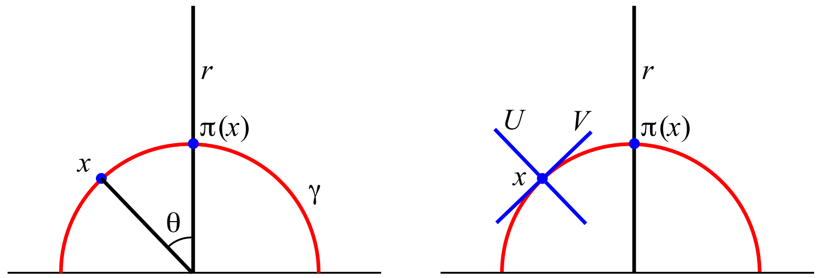



FIGURE 1. The hyperbolic cosinus of the distance between $x$ e $\pi(x)$ is the inverse of the cosinus of $\theta$ (left). To determine the maximum dilatation we decompose the tangent space $T_x$ (right).

**2.3. Boundary extension of a pseudo-isometry.** We dedicate this section to showing the following.

THEOREM 2.6. *A pseudo-isometry $F\colon \mathbb{H}^n \to \mathbb{H}^n$ extends to a continuous map $F\colon \overline{\mathbb{H}^n} \to \overline{\mathbb{H}^n}$ that injects $\partial\mathbb{H}^n$ to itself.*

We separate the proof in some lemmas.

LEMMA 2.7. *Consider the picture in Fig. 1. We have*

$$\cosh d(x,\pi(x)) = \frac{1}{\cos\theta}.$$

PROOF. We can suppose $\pi(x) = i$. The geodesic $r$ is parametrized as $ie^t$. The Möbius transformation $z \mapsto \frac{z+1}{-z+1}$ sends $r$ to $\gamma$ and fixes $i$, hence $\gamma(t) = \frac{ie^t+1}{-ie^t+1}$. Set $s = d(x,\pi(x))$. We get $x = \frac{ie^s+1}{-ie^s+1}$ and

$$\cos\theta = \Im x = \Im\frac{(ie^s+1)^2}{e^{2s}+1} = \frac{2e^s}{e^{2s}+1} = \frac{2}{e^s+e^{-s}} = \frac{1}{\cosh s}.$$

□

LEMMA 2.8. *Let $r \subset \mathbb{H}^n$ be a line and $\pi\colon \mathbb{H}^n \to r$ the orthogonal projection to $r$. The maximum dilatation of $\pi$ at $x \in \mathbb{H}^n$ is*

$$d = \frac{1}{\cosh s}$$

*where $s = d(x,r)$.*

PROOF. We use the half-space model with $r$ and $x$ as in Fig. 1-(left): we know that $\cosh s = \frac{1}{\cos\theta}$. We have $T_x = U \oplus V$ as in Fig. 1-(right) with $V = \ker d\pi_x$. A generator $u$ of $U$ is just rotated by $d\pi_x$ with respect to the euclidean metric; with respect to the hyperbolic metric we have

$$\frac{|d\pi_x(u)|}{|u|} = \frac{x_n}{\pi(x)_n} = \cos\theta = \frac{1}{\cosh s}.$$

□

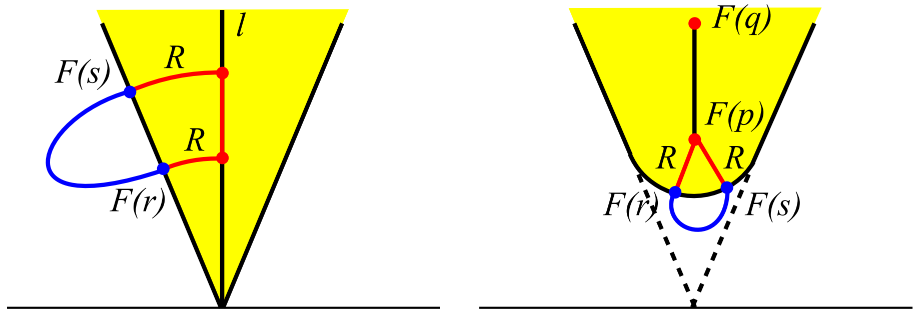



FIGURE 2. The red paths give better estimates for the distance between $F(r)$ and $F(s)$. On the left: since $F$ is $C_1$-Lipschitz, the blue path has length at most $C_1 d(r,s)$. Its projection onto $l$ has dilatation at most $1/\cosh R$ by Lemma 2.8, hence the red path in $l$ has length at most $C_1 d(r,s)/\cosh R$. Therefore $d(F(r),F(s)) \leqslant C_1 \frac{d(r,s)}{\cosh R} + 2R$. On the right we get $d(F(r),F(s)) \leqslant 2R$.

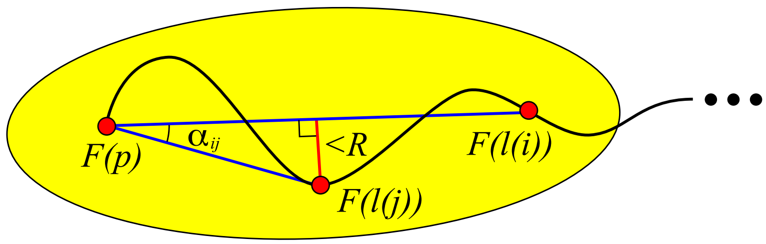
We denote by $\overline{pq}$ the segment from $p$ to $q$ and by $N_r(A)$ the $r$-neighborhood of $A$.

LEMMA 2.9. *Let $F\colon \mathbb{H}^n \to \mathbb{H}^n$ be a pseudo-isometry. There is a $R > 0$ such that*
$$F(\overline{pq}) \subset N_R\big(\overline{F(p)F(q)}\big)$$
*for all distinct points $p, q \in \mathbb{H}^n$.*

PROOF. Let $C_1, C_2$ be the pseudo-isometry constants of $F$. Fix a sufficiently big $R$ so that $\cosh R > C_1^2$. Let $l$ be the line containing $F(p)$ and $F(q)$. We show that $F(\overline{pq})$ can exit from $N_R(l)$, but only for a limited amount of time. Let $\overline{rs} \subset \overline{pq}$ be a maximal segment where $F(\overline{rs})$ is disjoint from the interior of $N_R(l)$, as the blue arc in Fig. 2-(left). We have
$$\frac{1}{C_1} d(r,s) - C_2 \leqslant d(F(r),F(s)) \leqslant C_1 d(r,s).$$
We can strengthen the right hand-side as in Fig. 2-(left) and get
$$\frac{1}{C_1} d(r,s) - C_2 \leqslant d(F(r),F(s)) \leqslant C_1 \frac{d(r,s)}{\cosh R} + 2R$$
Since $\cosh R > C_1^2$ we get $d(r,s) < M$ for some constant $M$ that depends only on $C_1$ and $C_2$. We have proved that $F(\overline{pq})$ may exit from $N_R(l)$ only on subsegments of length $< M$. Since $F$ is $C_1$-Lipchitz the curve $F(\overline{pq})$ lies entirely in $N_{R+C_1M/2}(l)$, and we replace $R$ with $R + C_1M/2$.

It remains to prove that $F(\overline{pq})$ lies entirely (up to taking a bigger $R$) in the bounded set $N_R(\overline{F(p)F(q)})$: the proof is analogous and easier, since Fig. 2-(right) shows that $d(F(r),F(s)) \leqslant 2R$. $\square$

In the previous and following lemmas, the constant $R$ depends only on the pseudo-isometry constants $C_1$ and $C_2$.




PROOF.

FIGURE 3. For any $0 < u < t$, the point $F(l(u))$ is contained in the (yellow) $R$-neighborhood of $\overline{F(p)F(l(t))}$. If $u$ is big, the blue segment $\overline{F(p)F(l(u))}$ is long, while the red one is bounded by $R$: hence the angle $\alpha_{tu}$ between $v_t$ and $v_u$ is small. Therefore $v_t$ is a Cauchy sequence.

LEMMA 2.10. *Let $F\colon \mathbb{H}^n \to \mathbb{H}^n$ be a pseudo-isometry. There is a $R > 0$ such that for all $p \in \mathbb{H}^n$ and any half-line $l$ starting from $p$ there is a unique half-line $l'$ starting from $F(p)$ such that*

$$F(l) \subset N_R(l').$$

Cambiare $i, j$ in $t, u$ nella figura.

We parametrize $l$ as a geodesic $l\colon [0, +\infty) \to \mathbb{H}^n$ with unit speed. Since $F$ is a pseudo-isometry we get

$$\lim_{t\to\infty} d\big(F(p), F(l(t))\big) \to \infty.$$

Let $v_t \in T_{F(p)}$ be the unitary tangent vector pointing towards $F(l(t))$: Fig. 3 shows that $\{v_t\}_{t\in\mathbb{N}}$ is a Cauchy sequence, that converges to a unitary vector $v \in T_{F(p)}$. Let $l'$ be the half-line starting from $F(p)$ with direction $v$. It is easy to check that $F(l) \subset N_R(l')$ and $l'$ is the unique half-line from $p$ with this property. $\square$

The previous lemma gives a recipe to transform half-lines $l$ into half-lines $l'$. Since $\partial\mathbb{H}^n$ is an equivalence relation of half-lines, we define the extension $F\colon \partial\mathbb{H}^n \to \partial\mathbb{H}^n$ by sending $l$ to $l'$.

LEMMA 2.11. *The boundary extension $F\colon \partial\mathbb{H}^n \to \partial\mathbb{H}^n$ is well-defined and injective.*

PROOF. Let $l_1, l_2$ be two half-lines at bounded distance $d(l_1(t), l_2(t)) < M$ for all $t$. If $d(l'_1(t), l'_2(t)) \to \infty$ we get $d\big(F(l_1(t)), F(l_2(t))\big) \to +\infty$, a contradiction since $F$ is Lipschitz. Therefore $l'_1, l'_2$ are at bounded distance.

Injectivity is proved analogously: if $l_1$ and $l_2$ are divergent then $l'_1$ and $l'_2$ also are because $F$ is a pseudo-isometry. $\square$

It remains to prove that the extension $F\colon \overline{\mathbb{H}^n} \to \overline{\mathbb{H}^n}$ is continuous. We start by extending Lemma 2.10 from half-lines to lines.

LEMMA 2.12. *Let $F\colon \mathbb{H}^n \to \mathbb{H}^n$ be a pseudo-isometry. There is a $R > 0$ such that for any line $l$ there is a unique line $l'$ with $F(l) \subset N_R(l')$.*

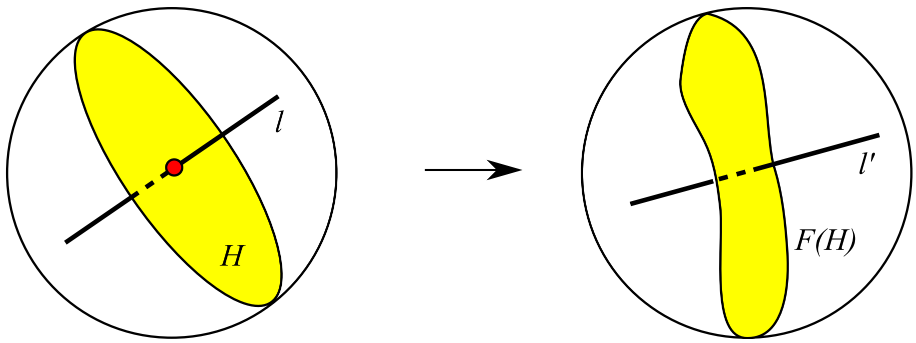



FIGURE 4. Let $l$ and $H$ be a line and an orthogonal hyperplane. The orthogonal projection of $H$ onto $l$ is obviously a point $l \cap H$; the pseudo-isometry $F$ mildly distorts this picture: the image $F(H)$ projects to a bounded segment in $l'$.

PROOF. Parametrize $l$ as $l\colon (-\infty, +\infty) \to \mathbb{H}^n$ with unit speed. By cutting $l$ into two half-lines we know that $F(l(t))$ is a curve that tends to two distinct points $x_\pm \in \partial\mathbb{H}^n$ as $t \to \pm\infty$. Let $l'$ be the line with endpoints $x_\pm$. For any $t > 0$ we have

$$F\big(l([-t,t])\big) \subset N_R\left(\overline{F(l(-t))F(l(t))}\right)$$

and by sending $t \to +\infty$ we deduce that $F(l) \subset N_R(l')$. $\square$

The next lemma says that a pseudo-isometry does not distort much lines and orthogonal hyperplanes.

LEMMA 2.13. *Let $F\colon \mathbb{H}^n \to \mathbb{H}^n$ be a pseudo-isometry. There is a $R > 0$ such that for any line $l$ and hyperplane $H$ orthogonal to $l$, the image $F(H)$ projects orthogonally to $l'$ onto a bounded segment length smaller than $R$.*

PROOF. See Fig. 4. Consider a generic line $s \subset H$ passing through $p = l \cap H$. By the previous lemmas $F(s) \subset N_R(s')$ with $s' \neq l'$, and the orthogonal projection on a line $l$ sends any other line $s'$ onto a segment, bounded by the images of the endpoints of $s'$.

Consider as in Fig. 5 the line $s$, with one endpoint $s^\infty$ and the corresponding endpoint $F(s^\infty)$ of $s'$. The figure shows that the projection $f$ of $F(s^\infty)$ to $l'$ is at bounded distance from a point $q$ which does not depend on $s$. $\square$

LEMMA 2.14. *The extension $F\colon \overline{\mathbb{H}^n} \to \overline{\mathbb{H}^n}$ is continuous.*

PROOF. Consider $x \in \partial\mathbb{H}^n$ and its immage $F(x) \in \partial\mathbb{H}^n$. Let $l$ be a half-line pointing to $x$: hence $l'$ points to $F(x)$. The half-spaces orthogonal to $l'$ determine a neighborhood system for $F(x)$: consider one such half-space $S$.

Let $R > 0$ be as in the previous lemmas. The image $F(l)$ is $R$-close to $l'$, hence for sufficiently big $t$ the point $F(l(t))$ lies in $S$ at distance $> R$ from $\partial S$. By the previous lemma the image $F(H(t))$ of the hyperplane

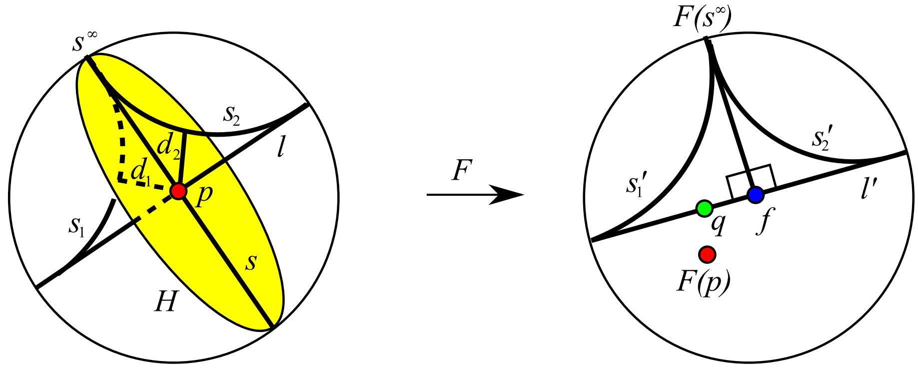



FIGURE 5. The lines $s_1$ and $s_2$ have fixed distance $d_1 = d_2 = \cosh^{-1}\sqrt{2}$ from $p$. The lines $l'$, $s_1'$, and $s_2'$ approximate up to an error $R$ the images of $l$, $s_1$, and $s_2$. The projection $q$ of $F(p)$ on $l'$ is hence $R$-close to $F(p)$, which is in turn $(C_1 d)$-close to the lines $s_i'$. Therefore $q$ is $(C_1 d + 2R)$-close to both $s_1'$ and $s_2'$. This easily implies that $f$ is $(C_1 d + 2R)$-close to $q$.

$H(t)$ orthogonal to $l(t)$ is also contained in $S$. Hence the entire half-space bounded by one such $H(t)$ goes inside $S$ through $F$. This shows that $F$ is continuous at every point $x \in \partial\mathbb{H}^n$. □

With some effort, we have proved that every pseudo-isometry of $\mathbb{H}^n$ extends continuously to the boundary. This has an immediate corollary.

COROLLARY 2.15. *Let $f\colon M \to N$ be a smooth homotopy equivalence between closed hyperbolic $n$-manifolds. Any lift extends to a continuous map $\widetilde{f}\colon \overline{\mathbb{H}^n} \to \overline{\mathbb{H}^n}$ whose restriction $\widetilde{f}|_{\partial\mathbb{H}^n}\colon \partial\mathbb{H}^n \to \partial\mathbb{H}^n$ is a homeomorphism.*

PROOF. Pick a smooth homotopic inverse $g$. Both $f$ and $g$ lift to pseudo-isometries and extend to their boundaries. In the proof of Proposition 2.5 we have seen that $\widetilde{g}\circ\widetilde{f}$ has finite maximum dispacement and hence its extension to $\partial\mathbb{H}^n$ is the identity. Therefore $\widetilde{g}|_{\partial\mathbb{H}^n}$ and $\widetilde{f}|_{\partial\mathbb{H}^n}$ are homeomorphisms. □

**2.4. Conclusion of the proof of Mostow theorem.** To prove Mostow rigidity we still need some lemma.

LEMMA 2.16. *Let $f\colon M \to N$ be a smooth homotopic equivalence of closed hyperbolic $n$-manifolds. The extension $\widetilde{f}\colon \partial\mathbb{H}^n \to \partial\mathbb{H}^n$ of a lift sends the vertices of a regular ideal simplex to the vertices of a regular ideal simplex.*

PROOF. Let $w_0, \ldots, w_n$ be vertices of a regular ideal simplex and suppose by contradiction that their images $\widetilde{f}(w_0), \ldots, \widetilde{f}(w_n)$ span a non-regular ideal simplex. By Theorem 1.10 this simplex has volume smaller than $v_n - 2\delta$ for some $\delta > 0$. By continuity there are neighborhoods $U_i$ of $v_i$ in $\overline{\mathbb{H}^n}$ for $i = 0, \ldots, n$ such that the volume of the simplex with vertices $\widetilde{f}(u_0), \ldots, \widetilde{f}(u_n)$ is smaller than $v_n - \delta$ for any choice of $u_i \in U_i$.

Motivare continuità.

In Section 1.4 we have defined a cycle

$$\alpha(t) = \sum_{\sigma\in\Sigma} \lambda_\sigma(t)\cdot\sigma$$

where $t$ depends on $\varepsilon$. We say that a singular simplex $\sigma\in\Sigma$ is *bad* if its $i$-th vertex is contained in $U_i$ for all $i$. Let $\Sigma^{\rm bad}\subset\Sigma$ be the subset of all bad singular simplexes and define

$$\alpha(t)^{\rm bad} = \sum_{\sigma\in\Sigma^{\rm bad}} \lambda_\sigma(t)\cdot\sigma.$$

We want to estimate $|\alpha(t)|$ and $|\alpha^{\rm bad}|$. We prove that

$$|\alpha(t)| = \sum_{s\in S} |\lambda_\sigma(t)|$$

is a real number independent of $t$: let $S_0\subset S(t)$ be the set of all $t$-simplexes having the first vertex in $D(x)$. It follows from the definitions that $|\alpha(t)|$ equals the measure of $S_0$. Moreover the set $S_0$ is in natural correspondence with the set of all isometries that send $x$ to some point in $D(x)$: its volume does not depend on $t$.

Dire che il volume è pari al volume di $D(x)$ e quindi $M$?

To estimate $|\alpha(t)^{\rm bad}|$ we fix $g_0\in\Gamma$ so that $D(g_0x)\subset U_0$. Let $S^{\rm bad}\subset S(t)$ be the set of all bad $t$-simplexes with first vertex in $D(g_0x)$. If $t$ is sufficiently big, the volume of $S^{\rm bad}$ is bigger than a constant independent of $t$.

Motivare questo.

We have proved that $|\alpha(t)^{\rm bad}|/|\alpha(t)| > C$ for some $C>0$ independent of $t$. On the renormalization $\bar\alpha(t) = \alpha(t)/_{k_t}$ we get the same ratio $|\bar\alpha(t)^{\rm bad}|/|\bar\alpha(t)| > C$. The map $f\colon M\to N$ has degree one and hence sends $\bar\alpha(t)$ to a class

$$f_*(\bar\alpha(t)) = \frac{1}{k_t}\sum_{\sigma\in\Sigma}\lambda_\sigma(t)\cdot(f\circ\sigma)^{\rm st}$$

representing $[N]$. Since a $C$-portion of $\bar\alpha(t)$ is bad, a $C$-portion of simplexes in $f_*(\bar\alpha(t))$ has volume smaller than $v_n-\delta$ and hence

$$\mathrm{Vol}(N) = \int_{f_*(\bar\alpha(t))}\omega < |\bar\alpha(t)|((1-C)v_n + C(v_n-\delta)) = |\bar\alpha(t)|(v_n-\delta C).$$

Since this holds for all $t$ and $|\bar\alpha(t)|\to\|M\|$ we get

$$\mathrm{Vol}(N) < \|M\|(v_n-\delta C) = \mathrm{Vol}(M) - \delta C\cdot\|M\|.$$

Corollary 1.20 gives $\mathrm{Vol}(M)=\mathrm{Vol}(N)$: a contradiction. $\square$

Now we use for the first time the hypothesis $n\geqslant 3$.

Proposition 2.17. *Let $\Delta\subset\mathbb{H}^n$ be a regular ideal simplex and $F$ a facet of $\Delta$. If $n\geqslant 3$ the only regular ideal simplexes in $\mathbb{H}^n$ having $F$ as facet are $\Delta$ and $\Delta'$, obtained mirroring $\Delta$ along $F$.*

PROOF. Every regular simplex of dimension $\geqslant 2$ has a *barycenter* defined by intersecting its *axis*, the unique lines exiting from a vertex and orthogonal to the opposite facet. Take the line orthogonal to the barycenter of $F$: the last vertex of $\Delta'$ must be the endpoint of this line. □

Given an ideal $n$-simplex $\Delta \subset \mathbb{H}^n$, we define $R(\Delta)$ as the set of all $n$-simplexes obtained iteratively from $\Delta$ by mirroring along all the facets. If $\Delta$ is the regular 2- or 3-simplex we obtain two tessellations, see Chapter 3. If $\Delta$ is a regular $n$-simplex with $n \geqslant 4$, its dihedral angle is equal to the dihedral angle of a regular euclidean $(n-1)$-simplex, which does not divide $2\pi$: therefore simplexes overlap a lot and we do not obtain a tessellation. In any case we have the following.

EXERCISE 2.18. Let $\Delta$ be a regular ideal $n$-simplex in $\mathbb{H}^n$. The vertices of all elements in $R(\Delta)$ form a dense subspace of $\partial\mathbb{H}^n$.

Da dimostrare?

We turn back to Mostow rigidity.

PROPOSITION 2.19. *Let $f\colon M \to N$ be a smooth homotopic equivalence between closed hyperbolic orientable manifolds of dimension $n \geqslant 3$. The restriction $\widetilde{f}|_{\partial H^n}\colon \partial\mathbb{H}^n \to \partial\mathbb{H}^n$ is the trace of an isometry $\psi\colon \mathbb{H}^n \to \mathbb{H}^n$.*

PROOF. Let $v_0, \ldots, v_n \in \partial\mathbb{H}^n$ vertices of a regular ideal simplex $\Delta$. The lift $\widetilde{f}$ sends them to the vertices of a regular ideal simplex, and let $\psi$ the unique isometry of $\mathbb{H}^n$ such that $\psi(v_i) = \widetilde{f}(v_i)$ for all $i$.

Iteratively, Proposition 2.17 shows that the two maps coincide on all vertices of $R(\Delta)$, which form a dense set. Therefore $\widetilde{f} = \psi$ on $\partial\mathbb{H}^n$. □

We can finally prove Mostow rigidity theorem.

THEOREM 2.20. *Let $f\colon M \to N$ be a homotopic equivalence between closed orientable hyperbolic manifolds of dimension $n \geqslant 3$. The map $f$ is homotopically equivalent to an isometry.*

PROOF. Set $M = \mathbb{H}^n/_\Gamma$ and $N = \mathbb{H}^n/_\Lambda$, and pick a lift $\widetilde{f}$. We have

$$\widetilde{f} \circ g = f_*(g) \circ \widetilde{f} \quad \forall g \in \Gamma \tag{2}$$

for an isomorphism $f_*\colon \Gamma \to \Lambda$. We may suppose $f$ smooth. The boundary extension of $\widetilde{f}$ is the trace of an isometry $\psi\colon \mathbb{H}^n \to \mathbb{H}^n$. Hence

$$\psi \circ g = f_*(g) \circ \psi \quad \forall g \in \Gamma \tag{3}$$

holds at $\partial\mathbb{H}^n$. All the elements in (3) are isometries, and isometries are determined by their boundary traces: hence (3) holds also in $\mathbb{H}^n$. Therefore $\psi$ descends to an isometry

$$\psi\colon M \to N.$$

A homotopy between $f$ and $\psi$ may be constructed from a convex combination of $\widetilde{f}$ and $\psi$ in $\mathbb{H}^n$. □

**2.5. Consequences of Mostow rigidity.** The most important consequence is that the entire geometry of a closed hyperbolic $n$-manifold with $n \geqslant 3$ is a topological invariant. Numerical quantities like the volume of the manifold, its geodesic spectrum, etc. depend only on the topology of the manifold. We single out another application.

THEOREM 2.21. *Let $M$ be a closed hyperbolic manifold of dimension $n \geqslant 3$. The map*

$$\mathrm{Isom}(M) \to \mathrm{Out}(\pi_1(M))$$

*is an isomorphism.*

PROOF. We already know that it is injective by Proposition 2.22 from Chapter 3. We prove that it is surjective: every automorphism of $\pi_1(M)$ is represented by a homotopy equivalence since $M$ is aspherical (see Corollary 4.4 from Chapter 1), which is in turn homotopic to an isometry by Mostow rigidity. $\square$

We note that this is false in dimension $n = 2$, where $\mathrm{Isom}(S)$ is finite and $\mathrm{Out}(\pi_1(S))$ is infinite.

CHAPTER 9

# Surface diffeomorphisms

We introduce in this chapter some analogies between $\mathbb{H}^n$ and $\mathrm{Teich}(S_g)$. We have already seen that $\mathbb{H}^n$ compactifies to a closed disc $\overline{\mathbb{H}^n}$, that $\mathrm{Isom}(\mathbb{H}^n)$ acts on this closed disc, and that an isometry is *elliptic*, *parabolic*, or *hyperbolic* according to where are its fixed points.

Analogously, we will construct in this chapter a natural compactification of the open ball $\mathrm{Teich}(S_g)$ to a closed disc. The mapping class group $\mathrm{MCG}(S_g)$ acts on this closed disc, and there will be a trichotomy for the elements of $\mathrm{MCG}(S_g)$ which depends on its fixed points.

## 1. Geodesic currents

**1.1. Projective immersion.** Recall from Chapter 4 that $\mathscr{S} = \mathscr{S}(S_g)$ is the set of all simple closed curves in the closed surface $S_g$, seen up to isotopy and changing of orientation (these curves are hence *unoriented*). When $g \geqslant 2$ the length functions provide an injective map

$$i \colon \mathrm{Teich}(S_g) \hookrightarrow \mathbb{R}^{\mathscr{S}}.$$

We have identified $\mathrm{Teich}(S_g)$ with its image and given it the induced topology. With that topology $\mathrm{Teich}(S_g)$ is homeomorphic to an open ball of dimension $6g-6$. We now want to compactify $\mathrm{Teich}(S_g)$: a first tentative could be to take its closure in $\mathbb{R}^{\mathscr{S}}$, but it does not work.

Proposition 1.1. *The subspace* $\mathrm{Teich}(S_g)$ *is closed in* $\mathbb{R}^{\mathscr{S}}$.

Proof. The inclusion map is proper, hence closed (see Chapter 4). $\square$

We now consider the projective space $\mathbb{P}(\mathbb{R}^{\mathscr{S}})$ with the projection

$$\pi \colon \mathbb{R}^{\mathscr{S}} \setminus \{0\} \longrightarrow \mathbb{P}(\mathbb{R}^{\mathscr{S}}).$$

Proposition 1.2. *The composition*

$$\pi \circ i \colon \mathrm{Teich}(S_g) \longrightarrow \mathbb{P}(\mathbb{R}^{\mathscr{S}})$$

*is injective.*

Proof. Suppose that there are two distinct points $h, h' \in \mathrm{Teich}(S_g)$ and a constant $k > 1$ such that $\ell^\gamma(h) = k \cdot \ell^\gamma(h')$ for all $\gamma \in \mathscr{S}$.

Let $\gamma_1, \gamma_2 \in \mathscr{S}$ be two curves with $i(\gamma_1, \gamma_2) = 1$. We take $x_0 = \gamma_1 \cap \gamma_2$ as a basepoint for $\pi_1(S_g, x_0)$. The elements $\gamma_2 * \gamma_1$ and $\gamma_2 * \gamma_1^{-1}$ define two more non-trivial simple closed curves in $S_g$. The formula

$$\mathrm{tr}(A) \cdot \mathrm{tr}(B) = \mathrm{tr}(AB) + \mathrm{tr}(A^{-1}B)$$

holds for any $A, B \in \mathrm{SL}_2(\mathbb{R})$. Proposition 2.11 from Chapter 5 implies that

$$2\cosh\left(\frac{l(\gamma_1)}{2}\right) \cdot \cosh\left(\frac{l(\gamma_2)}{2}\right) = \cosh\left(\frac{l(\gamma_2 * \gamma_1)}{2}\right) + \cosh\left(\frac{l(\gamma_2 * \gamma_1^{-1})}{2}\right).$$

We have obtained a relation between the lengths of $\gamma_1, \gamma_2, \gamma_2 * \gamma_1$, and $\gamma_2 * \gamma_1^{-1}$ that holds for any hyperbolic metric on $S_g$. It may be written as:

$$\cosh\left(\frac{l(\gamma_1) + l(\gamma_2)}{2}\right) + \cosh\left(\frac{l(\gamma_1) - l(\gamma_2)}{2}\right) =$$
$$\cosh\left(\frac{l(\gamma_2 * \gamma_1)}{2}\right) + \cosh\left(\frac{l(\gamma_2 * \gamma_1^{-1})}{2}\right).$$

By contradiction every $h'$-length is $k$ times a $h$-length: this equation is hence valid after multiplying every argument by $k$. It is easy to check that

$$\cosh a + \cosh b = \cosh c + \cosh d, \quad \cosh ka + \cosh kb = \cosh kc + \cosh kd$$

if and only if $\{a, b\} = \{c, d\}$. This leads to a contradiction: the number $l(\gamma_1) + l(\gamma_2)$ is strictly bigger than $l(\gamma_2 * \gamma_1)$ or $l(\gamma_2 * \gamma_1^{-1})$, since $\gamma_2 * \gamma_1$ and $\gamma_2 * \gamma_1^{-1}$ have a non-geodesic representative of length $l(\gamma_1) + l(\gamma_2)$. $\square$

Dove diciamo che le geodetiche chiuse sono le curve più corte?

As we will see, the image of $\mathrm{Teich}(S_g)$ in $\mathbb{P}(\mathbb{R}^{\mathscr{S}})$ is not close.

**1.2. Thurston compactification.** We now embed $\mathscr{S}$ in $\mathbb{P}(\mathbb{R}^{\mathscr{S}})$. A simple closed curve $\gamma \in \mathscr{S}$ defines a functional $i(\gamma) \in \mathbb{R}^{\mathscr{S}}$ as follows:

$$i(\gamma)(\eta) = i(\gamma, \eta).$$

We have constructed a map $i \colon \mathscr{S} \to \mathbb{R}^{\mathscr{S}}$.

Proposition 1.3. *The composition*

$$\pi \circ i \colon \mathscr{S} \longrightarrow \mathbb{P}(\mathbb{R}^{\mathscr{S}})$$

*is injective.*

Proof. Let $\gamma_1, \gamma_2 \in \mathscr{S}$ be distinct. There is always a curve $\eta \in \mathscr{S}$ with $i(\gamma_1, \eta) \neq 0$ and $i(\gamma_2, \eta) = 0$. (If $i(\gamma_1, \gamma_2) > 0$, simply take $\eta = \gamma_2$.) $\square$

We see both $\mathrm{Teich}(S_g)$ and $\mathscr{S}$ as subsets of $\mathbb{R}^{\mathscr{S}}$.

Proposition 1.4. *The sets* $\mathrm{Teich}(S_g)$ *and* $\mathscr{S}$ *are disjoint in* $\mathbb{P}(\mathbb{R}^{\mathscr{S}})$.

Proof. For each $\gamma \in \mathscr{S}$ we have $i(\gamma, \gamma) = 0$, while every curve has positive length on any hyperbolic metric. $\square$

We can state Thurston's compactification theorem. Let $g \geqslant 2$.

Theorem 1.5 (Thurston compactification). *The closure* $\overline{\mathrm{Teich}(S_g)}$ *of* $\mathrm{Teich}(S_g)$ *in* $\mathbb{P}(\mathbb{R}^{\mathscr{S}})$ *is homeomorphic to* $D^{6g-6}$*, whose interior is* $\mathrm{Teich}(S_g)$ *and whose boundary contains* $\mathscr{S}$ *as a dense subset.*

In particular, the closure of $\mathscr{S}$ is a sphere $S^{6g-7}$. To prove this theorem we will introduce some geometric notions.

**1.3. The torus case.** On the torus $T$ we have

$$\operatorname{Teich}(T) = H^2, \qquad \mathscr{S} = \mathbb{Q} \cup \{+\infty\}.$$

The latter equality holds because a unoriented simple closed curve is determined by a pair $(p, q)$ of coprime integers, unique up to switching both sides, and hence by the number $\frac{p}{q} \in \mathbb{Q} \cup \{\infty\}$. We can also see both $\operatorname{Teich}(T)$ and $\mathscr{S}$ inside $\mathbb{R}^{\mathscr{S}}$ and Thurston compactification holds:

PROPOSITION 1.6. *The closure of* $\operatorname{Teich}(T)$ *in* $\mathbb{P}(\mathbb{R}^{\mathscr{S}})$ *is homeomorphic to* $D^2$*, whose interior is* $\operatorname{Teich}(T)$ *and whose boundary contains* $\mathscr{S}$ *as a dense subset.*

PROOF. Exercise 2.14 and Proposition 1.17 from Chapter 5 give

$$i\left(\frac{p}{q}, \frac{r}{s}\right) = \left|\det \begin{pmatrix} p & r \\ q & s \end{pmatrix}\right| = |ps - qr| = s \cdot \left|p - q\frac{r}{s}\right|$$
$$\ell^{\frac{p}{q}}(z) = \frac{|p + qz|}{\sqrt{\Im z}}$$

Consider the closure $\overline{H^2} = H^2 \cup \mathbb{R} \cup \{\infty\}$ of $H^2$ and define for all $z \in \overline{H^2}$ the functional

$$f_z \colon \frac{p}{q} \longmapsto |p + qz| \quad \text{if } z \neq \infty,$$
$$f_\infty \colon \frac{p}{q} \longmapsto |q|.$$

We have constructed a continuous immersion

$$f \colon \overline{H^2} \longrightarrow \mathbb{P}(\mathbb{R}^{\mathscr{S}})$$
$$z \longmapsto f_z.$$

The map is closed because it sends a compact into a Hausdorff space, hence it is a homeomorphism onto its image. By the formulas above, a metric $z \in H^2$ goes to $f_z$ while a curve $\frac{r}{s} \in \mathscr{S}$ goes to $f_{-\frac{r}{s}}$. $\square$

**1.4. Geodesics.** We have seen that on the torus length functions $\ell^\gamma(z)$ and geometric intersection $i(\gamma, \eta)$ may be collected in a single family of functionals. We will do the same also for $g \geqslant 2$.

Let $M$ a complete hyperbolic manifold. We indicate by $\mathscr{G}(M)$ the set of all geodesics $\mathbb{R} \to M$ run at unit speed, seen up to reparametrization $t \mapsto \pm t + k$: in particular the geodesics are not oriented.

We are particularly interested in the set $\mathscr{G} = \mathscr{G}(\mathbb{H}^2)$ of lines in $\mathbb{H}^2$. A line is determined by its extremes, hence there is a bijection

$$\mathscr{G} \longleftrightarrow \left(\partial\mathbb{H}^2 \times \partial\mathbb{H}^2 \setminus \Delta\right)/_\sim$$

where $\Delta = \{(a, a) \mid a \in \partial\mathbb{H}^2\}$ is the diagonal and $(a, b) \sim (b, a)$. We assign to $\mathscr{G}$ the topology of $\left(\partial\mathbb{H}^2 \times \partial\mathbb{H}^2 \setminus \Delta\right)/_\sim$. With the disc model $\partial\mathbb{H}^2 = S^1$.

EXERCISE 1.7. The space $\mathscr{G}$ is homeomorphic to an open Möbius strip.

The isometries of $\mathbb{H}^2$ act naturally on $\mathscr{G}$.

PROPOSITION 1.8. *If $S = \mathbb{H}^2/_\Gamma$ is a complete hyperbolic surface we get a natural bijection*

$$\mathscr{G}(S) \longleftrightarrow \mathscr{G}/_\Gamma$$

PROOF. Every line in $\mathbb{H}^2$ induces a geodesic in $\mathbb{H}^2/_\Gamma$ by composing with the covering $\pi\colon \mathbb{H}^2 \to \mathbb{H}^2/_\Gamma$, and two lines induce the same geodesic if and only if they are connected by the action of an element in $\Gamma$. $\square$

We see a geodesic $\gamma \in \mathscr{G}(S)$ as a $\Gamma$-orbit of lines in $\mathbb{H}^2$. Note that $\gamma$ has compact support in $S$ if and only if it wraps (infinitely many times) a closed geodesic.

**1.5. Geodesic currents.** We introduce this definition.

DEFINITION 1.9. Let $S = \mathbb{H}^2/_\Gamma$ be a complete hyperbolic surface. A *geodesic current* on $S$ is a locally finite, non-trivial, $\Gamma$-invariant Borel measure $\mu$ on $\mathscr{G}$.

We denote by $\mathscr{C} = \mathscr{C}(S)$ the set of all geodesic currents in $S$. It is a subset of the space $\mathscr{M}(\mathscr{G})$ of all Borel measures of $\mathscr{G}$, closed with respect to sum and product with a positive scalar, and inherits a topology.

EXAMPLE 1.10 (Simple closed curves). A simple closed geodesic $\gamma \in \mathscr{G}$ lifts to a $\Gamma$-orbit of disjoint lines in $\mathbb{H}^2$, which is in turn a discrete set in $\mathscr{G}$. The Dirac measure on this discrete set is locally finite and $\Gamma$-invariant, hence a geodesic current.

A simple closed curve $\gamma$ determines a geodesic current: we have constructed a natural map

$$\mathscr{S} \longrightarrow \mathscr{C}.$$

PROPOSITION 1.11. *The map is injective.*

PROOF. Distinct curves have distinct lifts and give Dirac measures with distinct (actually disjoint) supports. $\square$

We now define a current which is supported on the whole $\mathscr{G}$.

**1.6. The Liouville measure.** Let $\gamma\colon \mathbb{R} \to \mathbb{H}^2$ be a geodesic and $U_\gamma \subset \mathscr{G}$ the open set consisting of all lines intersecting $\gamma$ in a point. We can parametrize $U_\gamma$ via the homeomorphism

$$\mathbb{R} \times (0, \pi) \to U_\gamma$$

that sends $(t, \theta)$ to the line that intersects $\gamma$ in the point $\gamma(t)$ at an angle $\theta$. We define a volume 2-form on $U_\gamma$ as follows:

$$\mu_\gamma = \frac{1}{2} \sin\theta \, dt \wedge d\theta.$$

PROPOSITION 1.12. *The charts $U_\gamma$ form a differentiable atlas for $\mathscr{G}$. The 2-forms $\mu_\gamma$ match up to sign and hence give a measure $\mu$ on $\mathscr{G}$.*

PROOF. Every line intersects some other line, hence the charts cover $\mathscr{G}$. We consider a line $r \in U_\gamma \cap U_{\gamma'}$. The charts $U_\gamma$ and $U_{\gamma'}$ have parametrizations $(t,\theta)$ and $(t',\theta')$ and 2-forms

$$\mu = \frac{1}{2}\sin\theta dt \wedge d\theta, \quad \mu' = \frac{1}{2}\sin\theta' dt' \wedge d\theta'.$$

Consider the jacobian

$$J = \begin{pmatrix} \dfrac{\partial\theta'}{\partial\theta} & \dfrac{\partial t'}{\partial\theta} \\ \dfrac{\partial\theta'}{\partial t} & \dfrac{\partial t'}{\partial t} \end{pmatrix}$$

and recall that

$$dt' \wedge d\theta' = \det J \cdot dt \wedge d\theta.$$

We need to show that

$$\det J = \frac{\sin\theta}{\sin\theta'}.$$

Consider first the case $\gamma$, $\gamma'$, and $r$ intersect in the same point $O$. We find

$$\frac{\partial\theta'}{\partial\theta} = 1, \qquad \frac{\partial t'}{\partial\theta} = 0, \qquad \frac{\partial t'}{\partial t} = \frac{\sin\theta'}{\sin\theta}$$

that implies $\det J = \frac{\sin\theta}{\sin\theta'}$. Consider the case $\gamma$ and $\gamma'$ intersect $r$ in two distinct points $P$ and $P'$ at some distance $l > 0$. By the previous case we may suppose that $\gamma$, $\gamma'$ are orthogonal to $r$, hence $\sin\theta = \sin\theta' = 1$. We get

$$\frac{\partial\theta'}{\partial\theta} = \cosh l, \quad \frac{\partial t'}{\partial\theta} = \sinh l, \quad \frac{\partial\theta'}{\partial t} = \sinh l, \quad \frac{\partial t'}{\partial t} = \cosh l$$

that implies $\det J = \cosh^2 l - \sinh^2 l = 1$. □

Dimostrare derivate parziali con disegni o altro. B

The measure $\mu$ on $\mathscr{G}$ is the *Liouville measure*: it is not induced by a global volume 2-form on $\mathscr{G}$ because $\mathscr{G}$ is non-orientable, see Esercise 1.7. The Liouville measure is clearly invariant by the action of $\mathrm{Isom}(\mathbb{H}^2)$. The factor $\frac{1}{2}$ in the definition was chosen to get the following result.

PROPOSITION 1.13. *Let $s \subset \mathbb{H}^2$ be a segment of length $L$. The lines in $\mathbb{H}^2$ intersecting $s$ form a set of measure $L$.*

PROOF. The set has measure

$$\int_0^\pi \int_0^L \frac{1}{2}\sin\theta\, dt\, d\theta = L\int_0^\pi \frac{1}{2}\sin\theta\, d\theta = L.$$

□

**1.7. The Liouville currents.** Let $S = \mathbb{H}^2/_\Gamma$ be a closed hyperbolic surface. The Liouville measure is $\mathrm{Isom}(\mathbb{H}^2)$-invariant: in particular it is $\Gamma$-invariant and hence defines a current $\mu \in \mathscr{C}(S)$, called the *Liouville current.*

In what follows we always suppose $g \geqslant 2$.

PROPOSITION 1.14. *The space $\mathscr{C}(S_g)$ of currents does not depend (up to canonical isomorphisms) on the hyperbolic metric on $S_g$.*

PROOF. Let $h, h'$ be two hyperbolic structures on $S_g$, giving two different coverings $\pi, \pi' \colon \mathbb{H}^2 \to S_g$. The identity map $S_g \to S_g$ lifts to a map between these coverings that extend continuously to their boundaries by Corollary 2.15 from Chapter 8. This induces an isomorphism between the corresponding current spaces. $\square$

Now that $\mathscr{C} = \mathscr{C}(S_g)$ is metric-independent, we note that every metric $h \in \mathrm{Teich}(S)$ induces a Liouville current $\mu_h$. We get a map

$$\mu \colon \mathrm{Teich}(S) \longrightarrow \mathscr{C}.$$

We will see later that $i$ is injective. We have mapped $\mathscr{S}$ and $\mathrm{Teich}(S)$ inside $\mathscr{C}$: we now introduce a bilinear form on $\mathscr{C}$ that extends both the length and the geometric intersection for closed geodesics.

**1.8. Intersection form.** Let $S = \mathbb{H}/_\Gamma$ be a hyperbolic surface. We denote by $\mathscr{I} \subset \mathscr{G} \times \mathscr{G}$ the open subset consisting of all pairs $(\gamma, \gamma')$ of incident distinct lines. We give $\mathscr{I}$ the topology induced by $\mathscr{G} \times \mathscr{G}$: hence $\mathscr{I}$ is a topological 4-manifold. The group $\Gamma$ acts on $\mathscr{I}$ diagonally.

EXERCISE 1.15. The map $\mathscr{I} \to \mathscr{I}/_\Gamma$ is a topological covering.

Therefore $\mathscr{I}/_\Gamma$ is a topological 4-manifold. Note that both $\mathscr{I}$ and $\mathscr{I}/_\Gamma$ are non-compact: the pairs $(\gamma, \gamma')$ of distinct lines intersecting in a fixed points for a non-compact set.

Two currents $\alpha, \beta \in \mathscr{C}$ induce a product measure $\alpha \times \beta$ on $\mathscr{G} \times \mathscr{G}$ and hence on $\mathscr{I}$. Since $\alpha \times \beta$ is $\Gamma$-invariant, it descends to a measure on $\mathscr{I}/_\Gamma$ which we still indicate by $\alpha \times \beta$, defined as follows: the measure on a well-covered connected open set $U \subset \mathscr{I}/_\Gamma$ is the measure of any connected component of its counterimage.

DEFINITION 1.16. The *intersection* $i(\alpha, \beta)$ of two geodesic currents is the total volume of $\mathscr{I}/_\Gamma$ in the measure $\alpha \times \beta$.

It is not obvious that $i(\alpha, \beta)$ is finite since $\mathscr{I}/_\Gamma$ is not compact: we will prove this later. We start by recognizing $i$ in some important cases. Recall that we consider $\mathscr{S}$ inside $\mathscr{C}$.

PROPOSITION 1.17. *If $\alpha, \beta \in \mathscr{S}$, $i(\alpha, \beta)$ is their geometric intersection.*

PROOF. Fix an auxiliary hyperbolic metric for $S_g$ and represent $\alpha$ and $\beta$ as geodesics. The measure $\alpha \times \beta$ is the Dirac measure with support the pairs $(l, l')$ of incident lines in $\mathbb{H}^2$ that cover respectively $\alpha$ and $\beta$. The $\Gamma$-orbits of these pairs are in natural bijection with the points in $\alpha \cap \beta$. Hence the volume of $\mathscr{I}_\Gamma$ is the cardinality of $\alpha \cap \beta$. $\square$

Recall the Liouville map $\mu\colon \mathrm{Teich}(S_g) \to \mathscr{C}$.

PROPOSITION 1.18. *If $\alpha \in \mathscr{S}$ we have $i(\mu_h, \alpha) = \ell^\alpha(h)$.*

PROOF. Give $S$ the metric $h$. The measure $\alpha \times \beta$ has its support on the incident pairs $(l, l')$ where $l$ is arbitrary and $l'$ is a lift of $\beta$.

A segment $s' \subset l'$ of length $L = \ell^\alpha(h)$ is a fundamental domain for the action of $\Gamma$ on the lifts of $l'$. Therefore $i(\mu_h, \alpha)$ is the volume of the pairs $(l, l')$ where $l$ is arbitrary and intersects $s'$. By Proposition 1.13 these pairs have volume $L$. $\square$

COROLLARY 1.19. *The Liouville map $\mu\colon \mathrm{Teich}(S_g) \to \mathscr{C}$ is injective.*

PROOF. We know that $\mathrm{Teich}(S_g)$ embeds in $\mathbb{R}^{\mathscr{S}}$: if $h \neq h'$ there is a curve $\gamma \in \mathscr{S}$ with $\ell^\gamma(h) \neq \ell^\gamma(h')$, hence $i(\mu_h, \gamma) \neq i(\mu'_h, \gamma)$. $\square$

We will consider both $\mathrm{Teich}(S_g)$ and $\mathscr{S}$ as subsets of $\mathscr{C}$. We know the geometric meaning of $i$ on two curves, and on a curve and a metric. What is $i$ on two metrics? We know the answer when they coincide.

PROPOSITION 1.20. *If $\alpha \in \mathrm{Teich}(S_g)$ we have $i(\mu, \mu) = -\pi^2\chi(S)$.*

Da dimostrare

**1.9. Continuity of the intersection function.** The following fact is not obvious, since $\mathscr{G}$ is not compact.

THEOREM 1.21. *The form $i\colon \mathscr{C} \times \mathscr{C} \to \mathbb{R}$ is continuous.*

Prove also that it is finite-valued

We say that a geodesic current *fills* the surface $S$ if every line in $\mathbb{H}^2$ intersects transversely at least one line in the support of $\alpha$. A Liouville measure fills $S$ since its support is the whole of $\mathscr{G}$. We say that $k$ simple closed curves $\gamma_1, \ldots, \gamma_k$ *fill* $S$ if the current $\gamma_1 + \ldots + \gamma_k$ does.

PROPOSITION 1.22. *Let $\gamma_1, \ldots, \gamma_k$ be simple closed geodesics with respect to some metric. If $S \setminus (\gamma_1 \cup \cdots \cup \gamma_k)$ consists of polygons, the curves fill $S$.*

PROOF. Every geodesic in $S$ intersects these curves. $\square$

EXERCISE 1.23. Let $\alpha$ and $\beta$ be currents. We have $i(\alpha, \beta) > 0$ if and only if there are two lines in the supports of $\alpha$ and $\beta$ that intersect transversely.

COROLLARY 1.24. *If $\alpha$ fills then $i(\alpha, \beta) > 0$ for any $\beta \in \mathscr{C}$.*

**1.10. A compactness criterion.** The following compactness criterion is simple and useful.

PROPOSITION 1.25. *If $\alpha \in \mathscr{C}$ fills $S$, the set of all $\beta \in \mathscr{C}$ with $i(\alpha, \beta) \leqslant M$ is compact for all $M > 0$.*

Proof. Let $C \subset \mathscr{C}$ be the set of all $\beta$ with $i(\alpha, \beta) \leqslant M$. It is closed because $i$ is continuous, we show that is is also compact.

Let $l$ be a line in $\mathbb{H}^2$. By hypothesis there is a $l'$ in the support of $\alpha$ which intersects $l$ transversely. Let $U_l$, $U_{l'}$ be two neighborhoods of $l, l'$ in $\mathscr{G}$ sufficiently small so that the following hold:

- every line in $U_l$ intersects a line in $U_{l'}$, hence $U_l \times U_{l'} \subset \mathscr{I}$,
- the product $U_{l'} \times U_l$ is mapped injectively in $\mathscr{I}/_\Gamma$.

If $\beta \in C$ we have

$$\alpha(U_{l'})\beta(U_l) = (\alpha \times \beta)(U_{l'} \times U_l) \leqslant (\alpha \times \beta)\left(\mathscr{I}/_\Gamma\right) = i(\alpha, \beta) < M.$$

Therefore every line $l$ has an open neighborhood $U_l$ such that

$$\beta(U_l) < K_l \quad \forall \beta \in \mathscr{C}$$

for some constant $K_l = M/\alpha(U_{l'})$ depending only on $l$. We may cover $\mathscr{G}$ with countably many such neighborhoods.

Let $\beta_i$ be a sequence in $C$. On each $U_l$ the sequence $\beta_i(U_l)$ is bounded, hence on a subsequence $\beta_i(U_l) \to \beta_\infty(U_l)$ for all $l$. □

Finire usando qualche teo di compattezza o $U_l$ sist di intorni.

Corollary 1.26. *The immersion $\mu\colon \mathrm{Teich}(S) \hookrightarrow \mathscr{C}$ is proper and a homeomorphism onto its image.*

Proof. The immersion is proper: if $h_i \in \mathrm{Teich}(S)$ is a divergence sequence, we know that on a subsequence there is a simple closed curve $\gamma$ such that $\ell^\gamma(h_i) = i(h_i, \gamma) \to \infty$. Since $i$ is continuous $\mu(h_i) \in \mathscr{C}$ diverges.

Controllare bene qui

To show that $\mu$ is a homeomorphism onto its image it is easier to consider the inverse $\mu^{-1}\colon \mu(\mathrm{Teich}(S)) \to \mathrm{Teich}(S)$. The map $\mu^{-1}$ is continuous because $i$ is and $\mathrm{Teich}(S)$ has the weakest topology where the length functions are continuous. We show that it is proper. Let $\gamma_1, \ldots, \gamma_k$ be simple closed curves that fill $S$. If $\mu(h_i)$ is a diverging sequence, by Proposition 1.25 we have $i(\mu(h_i), \sum_t \gamma_t) \to \infty$ and hence $i(\mu(h_i), \gamma_t) \to \infty$ for some $t$. Therefore $h_i$ is divergent also in $\mathrm{Teich}(S)$. □

The compactness criterion implies the following.

Corollary 1.27. *Let $\gamma_1, \ldots, \gamma_k$ be simple closed geodesics that fill $S$. The points $h \in$ Teich$(S)$ with $\ell^{\gamma_i}(h) \leqslant M$ form a compact subset of* Teich$(S)$.

Exercise 1.28. Use the compactness criterion to re-prove that on a hyperbolic $S_g$ there are only finitely many closed geodesics of bounded length.

**1.11. Projective currents.** We compose the immersions of $\mathrm{Teich}(S)$ and $\mathscr{S}$ in $\mathscr{C}$ with the projection

$$\pi\colon \mathscr{C} \longrightarrow \mathbb{P}\mathscr{C}$$

where $\mathbb{P}\mathscr{C} = \mathscr{C}/_\sim$ with $\alpha \sim \lambda\alpha$ for all $\lambda > 0$.

Proposition 1.29. *The space $\mathbb{P}\mathscr{C}$ is compact.*

PROOF. Pick $h \in \mathrm{Teich}(S)$. By the compactness criterion the set $C = \{\alpha \in \mathscr{C} \mid i(\alpha, h) = 1\}$ is compact. By Corollary 1.24 we have $i(\alpha, h) > 0$ for all $\alpha$: hence $\pi\colon C \to \mathbb{P}\mathscr{C}$ is surjective and $\mathbb{P}\mathscr{C}$ is compact. $\square$

PROPOSITION 1.30. *The composition $\pi \circ i\colon \mathscr{S} \to \mathbb{P}\mathscr{C}$ is injective.*

PROOF. Let $\gamma_1, \gamma_2 \in \mathscr{S}$ be distinct. There is always a curve $\eta \in \mathscr{S}$ with $i(\gamma_1, \eta) \neq 0$ and $i(\gamma_2, \eta) = 0$. $\square$

The image of $\mathrm{Teich}(S)$ in $\mathscr{C}$ is closed because the immersion $\mu$ is proper. The image in $\mathbb{P}\mathscr{C}$ is not closed, and its closure is a compact set since $\mathbb{P}\mathscr{C}$ is compact.

PROPOSITION 1.31. *The composition $\pi \circ \mu\colon \mathrm{Teich}(S) \to \mathbb{P}\mathscr{C}$ is injective and a homeomorphism onto its image. The boundary of $\pi(\mu(\mathrm{Teich}(S))$ consists of projective currents $[\alpha]$ with $i(\alpha, \alpha) = 0$.*

PROOF. Consider $\mathrm{Teich}(S)$ already properly embedded in $\mathscr{C}$. Since $i(h, h) = -\pi^2\chi(S)$ is constant on $\mathrm{Teich}(S)$, the composition is injective. The map $\mathrm{Teich}(S) \to \pi(\mathrm{Teich}(S))$ is continuous and proper: consider a diverging sequence $h_i \in \mathrm{Teich}(S)$. By compactness the sequence $[h_i] \in \mathbb{P}\mathscr{C}$ converges on a subsequence to some $[\alpha] \in \mathbb{P}\mathscr{C}$. For each $i$ there is a $\lambda_i > 0$ such that $\lambda_i h_i \to \alpha$ in $\mathscr{C}$. Since $h_i$ diverges in $\mathrm{Teich}(S)$ and hence in $\mathscr{C}$ we get $\lambda_i \to 0$. Since $i$ is continuous we get

$$i(\alpha, \alpha) = \lim_{i\to\infty} i(\lambda_i m_i, \lambda_i m_i) = -\pi^2\chi(S) \lim_{i\to\infty} \lambda_i^2 = 0.$$

In particular $[\alpha] \notin \pi(\mathrm{Teich}(S))$: this implies that $\pi\colon \mathrm{Teich}(S) \to \pi(\mathrm{Teich}(S)$ is proper and hence a homeomorphism onto its image. Moreover the boundary of $\pi(\mathrm{Teich}(S))$ consists of some elements $\alpha$ with $i(\alpha, \alpha) = 0$. $\square$

We consider both $\mathscr{S}$ and $\mathrm{Teich}(S)$ embedded in $\mathbb{P}\mathscr{C}$. To identify $\partial\mathrm{Teich}(S)$ we now analyze the geodesic currents $\alpha$ with $i(\alpha, \alpha) = 0$. These geodesc currents are geometric objects that contain and generalize $\mathscr{S}$.

## 2. Laminations

**2.1. Measured geodesic laminations.** Let $S = \mathbb{H}^2/_\Gamma$ be a hyperbolic surface. A *geodesic lamination* $\lambda$ is a non-empty set of disjoint simple complete geodesics in $S$, whose union is a closed subset of $S$. Each geodesic may be closed or open and is called a *leaf*; their union is the *support* of $\lambda$. We will often confuse $\lambda$ with its support for simplicity.

The following examples are fundamental:

- a finite set of disjoint closed geodeiscs in $S$,
- a set of disjoint lines in $\mathbb{H}^2$ whose union is closed.

The lamination in $\mathbb{H}^2$ may be particularly complicated, see Fig. 1.

EXERCISE 2.1. A set $\lambda$ of disjoint lines in $\mathbb{H}^2$ form a closed set if and only if $\lambda$ is closed as a subset of $\mathscr{G}$.

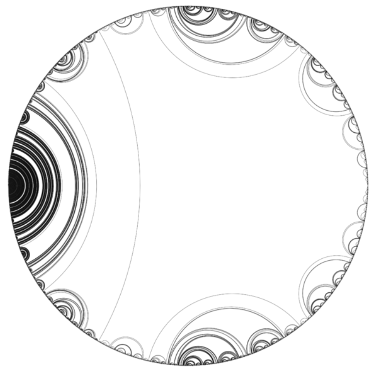


FIGURE 1. A geodesic lamination in $\mathbb{H}^2$.

If a set of disjoint lines in $\mathbb{H}^2$ is not closed, it suffices to take its closure to get a lamination. Those in $\mathbb{H}^2$ are fundamental, since a lamination in $S = \mathbb{H}^2/_\Gamma$ lifts to a $\Gamma$-invariant lamination in $\mathbb{H}^2$. The laminations in $S$ are hence in natural bijection with the $\Gamma$-invariant laminations in $\mathbb{H}^2$.

Let $\lambda \subset S$ be a geodesic lamination. A *transverse arc* to $\lambda$ is the support of a simple regular curve $\alpha\colon [a,b] \to S$ transverse to each leaf of $\lambda$, whose endpoints $\alpha(a)$ and $\alpha(b)$ are not contained in $\lambda$.

DEFINITION 2.2. A *transverse measure* for a lamination $\lambda \subset S$ is a locally finite measure $\mu_\alpha$ on each transverse arc such that:

(1) if $\alpha' \subset \alpha$ is a sub-arc of $\alpha$, the measure $\mu_{\alpha'}$ is the restriction of $\mu_\alpha$;
(2) the support of $\mu_\alpha$ is $\alpha \cap \lambda$;
(3) if there is an isotopy $\alpha_t\colon [a,b] \to S$ between two arcs $\alpha_0$ and $\alpha_1$ such that each level $\alpha_t$ is a transverse arc, then $\mu_{\alpha_0} = (\alpha_0^{-1} \circ \alpha_1)^{-1}(\mu_{\alpha_1})$.

In particular every arc $\alpha$ transverse to $\lambda$ has a finite *length*, defined as the total measure of the arc. The arc has length zero if and only if $\alpha \cap \lambda = \emptyset$. A *measured geodesic lamination* is a geodesic lamination with a transverse measure.

EXAMPLE 2.3. A lamination $\lambda$ formed by a finite set of disjoint closed geodesics $\gamma_1, \ldots, \gamma_k$ has a natural transverse measure: for any transverse arc $\alpha$, the measure $\mu_\alpha$ on $\alpha$ is just the Dirac measure supported in $\alpha \cap \lambda$.

More generaly, we may assign a positive weight $a_i > 0$ at each $\gamma_i$ and define $\mu_\alpha$ by giving a weight $a_i$ at each intersection $\alpha \cap \gamma_i$. When weights vary we get distinct measured laminations with the same support.

**2.2. Currents and measured geodesic laminations.** Let $S = \mathbb{H}^2/_\Gamma$ be a hyperbolic surface and $\beta$ be a geodesic current with $i(\beta, \beta) = 0$. By Exercise 1.23 the support of $\beta$ is a closed $\Gamma$-invariant subset of $\mathscr{G}$ formed by disjoint lines, which projects to a lamination $\lambda$ in $S$.

The lamination $\lambda$ has a natural transverse measure induced by $\beta$, defined as follows. Let $\alpha$ be an arc transverse to $\lambda$; let us lift it to an arc $\tilde{\alpha}$ in $\mathbb{H}^2$.

Up to cutting $\tilde{\alpha}$ in finitely many arcs we may suppose that it intersects each leaf of $\tilde{\lambda}$ in at most one point. We define the measure of a Borel set $U \subset \tilde{\alpha}$ as the $\beta$-measure of the lines in $\beta$ that it intersects.

EXERCISE 2.4. This transverse measure satisfies the axioms and give $\lambda$ the structure of a measured geodesic lamination.

PROPOSITION 2.5. *We have just defined a bijection*

$$\big\{\text{currents } \beta \textit{ with } i(\beta,\beta)=0\big\} \longleftrightarrow \big\{\text{measured geodesic laminations on } S\big\}$$

PROOF. Given a measured geodesic lamination $\lambda$ in $S$, we construct a geodesic current $\beta$. Consider the lift $\tilde{\lambda}$ in $\mathbb{H}^2$. Define the support of $\beta$ as the leaves of $\tilde{\lambda}$. For every leaf $r$ we choose a transverse arc $\alpha$ to $r$ that intersects every leaf in $\tilde{\lambda}$ in at most one point. The leaves intersected by $\alpha$ form an open neighborhood $U_r$ of $r$ in $\tilde{\lambda} \subset \mathscr{G}$. We define the measure $\beta_r$ on $U_r$ by transporting the transverse measure on $\alpha$.

The open sets $U_r$ cover $\tilde{\lambda} \subset \mathscr{G}$ and $\beta_r|_{U_r \cap U_{r'}} = \beta_{r'}|_{U_r \cap U_{r'}}$ for all $r, r'$. We may extract from this a locally finite covering and apply Proposition 3.3 from Chapter 1 to get a measure $\beta$ on $\tilde{\lambda}$ that extends each $\beta_r$.

The measure $\beta$ is $\Gamma$-invariant by construction, hence gives a geodesic current with $i(\mu,\mu) = 0$ since its support consists of disjoint lines. $\square$

We denote by $\mathscr{ML} \subset \mathscr{C}$ the set of all measured geodesic laminations on $S$, thus identified with the currents having zero self-intersection. With this identification the set $\mathscr{ML}$ does not depend on the hyperbolic metric chosen.

Recall that a multicurve is a finite collection of disjoint and homotopically non-trivial simple closed curves. A multicurve determines a measured geodesic lamination: consider $n$ parallel components as a single one with weight $n$ and use Example 2.3. We obtain the inclusions

$$\mathscr{S} \subset \mathscr{M} \subset \mathscr{ML} \subset \mathscr{C}.$$

**2.3. Euclidean singular foliations.** Geodesic laminations are not combinatorial in nature and hence difficult to construct: we introduce a tool that is useful for this task.

Fare GB prima per superfici con bordo angoloso e mettere questo esempio

Let the *Fontana plane* $\mathbb{F}$ be the euclidean plane $\mathbb{C}$ cut along the segment $[-1,1]$, *i.e.* obtained from $\mathbb{C}$ by substituting $(-1,1)$ with two copies of it, one attached to the upper half-plane and one to the lower. The Fontana plane is a flat surface with boundary consisting of two geodesic lines (the two copies of $(-1,1)$) and two cone points of angle $2\pi$.

We fix some terminology: the two cone points are called *singular*, the points in the two copies of $(-1,1)$ are the *boundary* points, and all the other points in $\mathbb{F}$ are *interior* points. We see $\mathbb{F}$ foliated by horizontal geodesic lines: three lines emanate from each singular point.

DEFINITION 2.6. A *euclidean singular foliation* on a closed surface $S$ is a non-empty closed set $\phi \subset S$ equipped with an atlas of charts onto open sets of $\mathbb{F}$, whose transition functions are isometries that preserve the

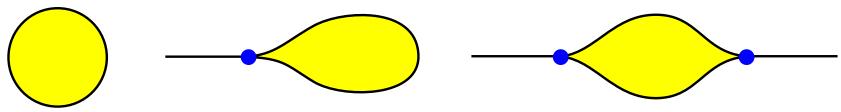


FIGURE 2. We require that each complementary region of a euclidean singular foliation is not a disc, a monogon, or a bigon.

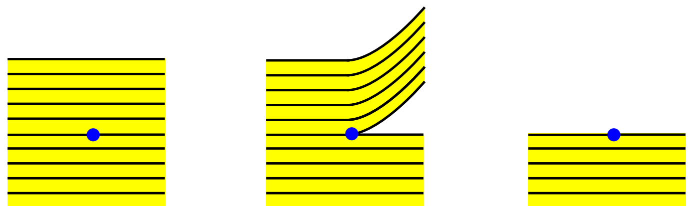


FIGURE 3. An interior, singular, and boundary point in a euclidean singular foliation.

horizontal foliation of $\mathbb{F}$. We require that no component of $S \setminus \mathrm{int}(\phi)$ is a disc, a monogon, or a bigon as in Fig. 2.

A point of $\phi$ is *singular*, *boundary*, or *inner* if it is mapped to a singular, boundary, or inner point of $\mathbb{F}$, see Fig. 3. The subsurface $\phi$ has a structure of flat surface with boundary consisting of geodesic lines (the regular boundary points) and cone points of angle $2\pi$ (the singular points). The *complementary regions* are the connected components of $S \setminus \mathrm{int}(\phi)$: each is a surface with poligonal boundary, with vertices at the singular points. We require that no complementary region is a disc, a monogon, or a bigon.

The singular foliation of $\mathbb{F}$ descends to a singular foliation of $\phi$ into geodesics, called *leaves*. A leaf may be open or close. A leaf is *singular* if it terminates at a singular point. Each singular point is adjacent to three singular leaves (counted with multiplicity). By compactness $\phi$ contains finitely many singular points and leaves: the following exercise shows that most leaves are non-singular.

A euclidean singular foliation will be called simply a *foliation*.

EXERCISE 2.7. The foliation $\phi$ contains uncountably many leaves.

EXAMPLE 2.8. Pick a square in $\mathbb{F}$ disjoint from the boundary as in Fig. and identify its opposite sides via an isometry. We get a torus equipped with a euclidean singular foliation $\phi$.

In this example the foliation has no singular and boundary points and covers the torus, hence $\chi(\phi) = 0$. More generally:

PROPOSITION 2.9. *There are $-2\chi(\phi)$ singular points in $\phi$.*

Fare GB con angoli e bordo.

PROOF. Gauss-Bonnet formula says

$$2\pi\chi(\phi) = \sum \alpha_i$$

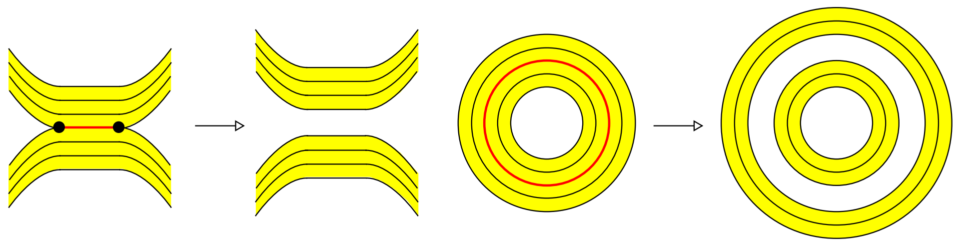


FIGURE 4. We may cut a foliation along an interior compact leaf.

where $\alpha_i$ are the exterior angles of the boundary cone points. The exterior angle of a singular point is $\pi - 2\pi = -\pi$. $\square$

COROLLARY 2.10. *A foliation $\phi$ in $S_g$ with $g \geqslant 2$ is a proper subset.*

PROOF. If $\phi$ has singular points then it is a proper subset. If it has no singular points then $\chi(\phi) = 0$ and $\phi$ is again a proper subset (a union of disjoint annuli). $\square$

COROLLARY 2.11. *There are no foliations in $S^2$.*

COROLLARY 2.12. *A foliation $\phi$ in $S$ contains at most $-6\chi(S)$ singular points. It contains $-6\chi(S)$ singular points if and only if its complementary regions consist of triangles and annuli.*

PROOF. We have $-6\chi(S) = -6\chi(\phi) - 6\chi(C) = 3n - 6\chi(C)$ where $n$ is the number of singular points and $C = C_1 \sqcup \cdots \sqcup C_k$ are the complementary regions. Each $C_i$ is incident to some $n_i$ singular points with $n = n_1 + \ldots n_k$. By hypothesis we have $3\chi(C_i) - n_i \leqslant 0$ and the equality holds if and only if $C_i$ is an annulus or a triangle. Therefore $3\chi(C) \leqslant n$ and we get

$$-6\chi(S_g) = 3n - 6\chi(C) \geqslant 3n - 2n = n.$$

The equality holds if and only if the $C_i$ are annuli and triangles. $\square$

**2.4. Transversal arcs and multicurves.** Let $S$ be equipped with a foliation $\phi$. An arc, curve, or multicurve $\alpha$ is *transverse* to $\phi$ if it is transverse to each leaf. A transverse arc $\alpha$ intersects $\phi$ in finitely many disjoint closed segments, and the endpoints of $\alpha$ may or may not be contained in $\phi$.

A *boundary leaf* of $\phi$ is a leaf contained in the boundary of $\phi$: it may be a segment connecting two singular points or a circle. The foliation $\phi$ sometimes contains also some interior (*i.e.* non-boundary) compact leaf $l$: a segment as in Fig. or a circle. In that case we may *cut* $\phi$ along $l$ as shown in Fig. and obtain a new foliation $\phi'$.

DEFINITION 2.13. A foliation $\phi$ is *reduced* if every compact interior leaf is a circle parallel to a boundary leaf.

Every foliation may be transformed to a reduced foliation by cutting along finitely many compact leaves: simply cut along every interior leaf and

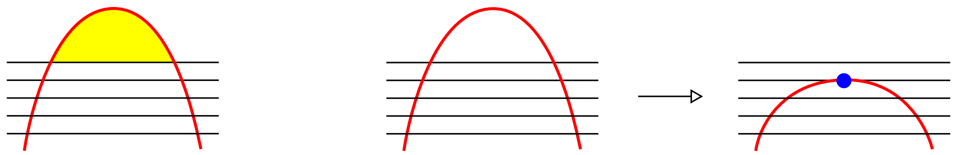


FIGURE 5. A bigon between a multicurve $\mu$ and a foliation $\phi$ (left) and how to eliminate it (right).

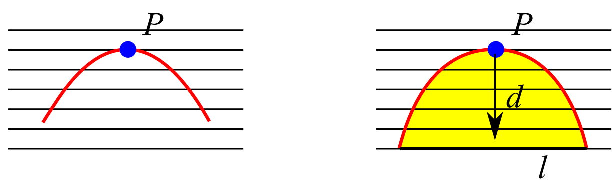



FIGURE 6. If $\mu$ is in general position with respect to $\phi$, it is tangent to some leaf in finitely many points as shown here (left). A perpendicular of small length $d$ points to some leaf $l$ at distance $d$ from $P$, which forms a bigon with $\mu$ (right).

along one circle leaf in each class of parallel closed leaves. It is slightly more convenient to work with reduced foliations.

A multicurve $\mu$ forms a *bigon* with $\phi$ if one component of $\mu$ forms a bigon with a boundary leaf of $\phi$ as in Fig. 5-left.

PROPOSITION 2.14. *Let $\phi$ be a reduced foliation. Every multicurve $\mu$ has a representative transverse to $\phi$ which forms no bigons with $\phi$.*

PROOF. We put $\mu$ in general position with respect to $\phi$, so that $\mu$ is tangent to $\phi$ only in finitely many points as in Fig. 6-left. We eliminate all bigons of $\mu$ using the isotopy shown in Fig. 5-right.

Consider a tangency point $P$. We draw from it a line perpendicular to the leaves as in Fig. 6-right of some length $d > 0$. The line ends at some leaf $l$ at distance $d$ from $P$. For sufficiently small $d$ the leaf $l$ forms a bigon with $\mu$ as in the figure.

Let now $D$ be the supremum of all values of $d$ such that $l$ exists and forms a bigon as in the figure. Some cases may occur: the two configurations of Fig. 7 are excluded because every component of $\mu$ is homotopically non-trivial and $\phi$ is reduced. The remaining possibilities are listed in in Fig. 8 and 9: each can be simplified by an isotopy of $\mu$ that decreases the number of tangency points of $\mu$.

The move in Fig. 9-right does not create any new bigon because no complementary region of $\phi$ is a bigon. If the move in Fig. 9-left creates a new bigon, we eliminate it as in Fig. 5-right: the resulting multicurve $\mu$ has the same number of tangency points as before but one component less in $\mu \setminus \phi$, hence the process ends in finite time and we get a multicurve that forms no bigon and has no tangency points. $\square$

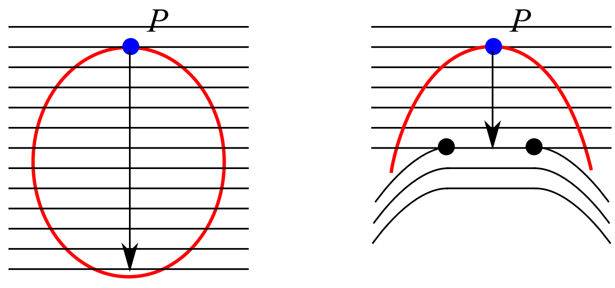



FIGURE 7. These two configurations do not arise because every component of a multicurve is homotopically non-trivial (left) and because the foliation is reduced (right).

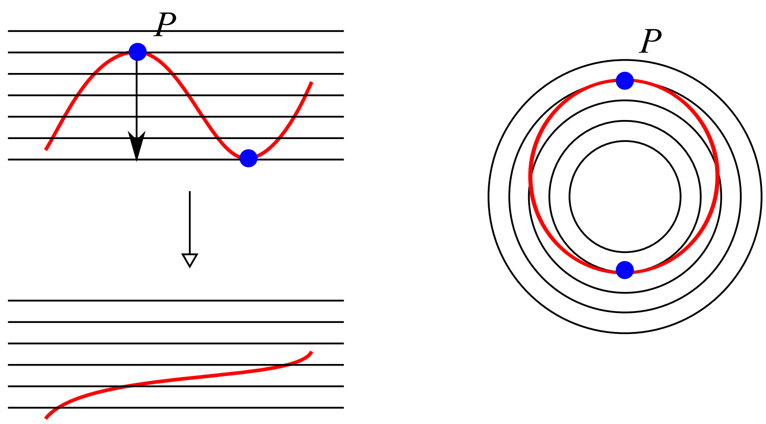



FIGURE 8. Each such configuration can be simplified. In the left picture we reduce the number of tangency points, in the right picture a component of $\mu$ is isotopic to a leaf, hence to a boundary leaf since $\phi$ is reduced: therefore it may be isotoped to be disjoint from $\phi$.

If there are no bigons with $\phi$, there are no bigons with any leaf of $\phi$.

PROPOSITION 2.15. *Let $\phi$ be a foliation in $S$. If a multicurve $\mu$ forms no bigon with $\phi$, then it forms no bigon with any leaf of $\phi$.*

PROOF. Suppose by contradiction that there is a bigon, that is a disc $D$ with $\partial D = \alpha \cup \beta$ with $\alpha \subset \mu$ and $\beta$ contained in some leaf of $\phi$. The arc $\alpha$ intersects $\phi$ transversely in finitely many arcs.

Consider $D$ abstractly. We enlarge $D$ so that the arcs in $\alpha \cap \phi$ are orthogonal to the foliation as in Fig. 10-left. Consider the complementary regions in $D \setminus \phi$ that are adjacent to $\alpha$. By hypothesis there are no bigons as in Fig. 11-left. If there is a triangle as in Fig. 11-right, we eliminate it as shown in the figure.

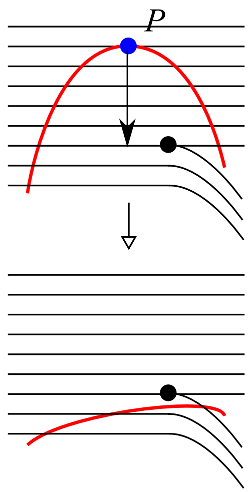



FIGURE 9. Each such configuration can be simplified.

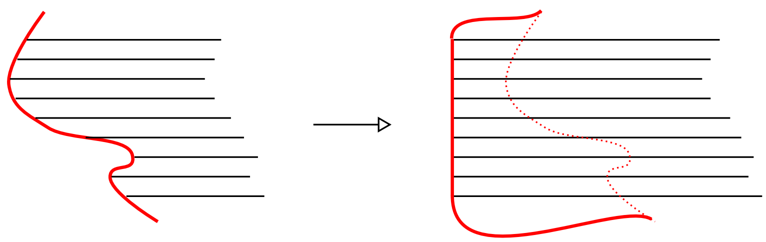


FIGURE 10. We can enlarge $D$ so that $\alpha \subset \partial D$ intersects $\phi$ into lines orthogonal to the leaves.

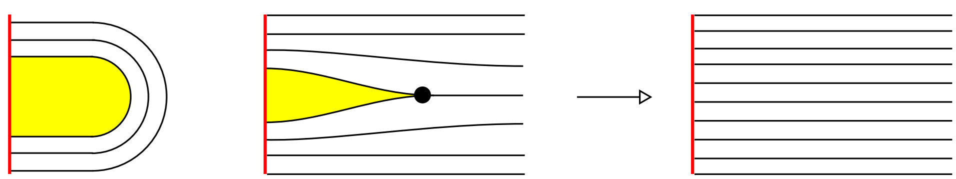


FIGURE 11. There are no complementary bigons as in the left picture. A complementary triangle as in the right picture can be eliminated by gluing together two sides of it. If the sides have different lengths, we enlarge the disc as in Fig. 10 to get sides of the same length.

If we double $D$ and $\phi$ along $\alpha$ we get another disc $D'$ containing a foliation $\phi'$ having $\partial D'$ as a leaf. We double again $D'$ along $\partial D'$ to get a foliation $\phi''$ in the 2-sphere $S^2$. Note that since there are no bigons and triangles in $D$ as in Fig. 10, no complementary region in $D' \setminus \phi'$ and hence in $S^2 \setminus \phi''$ is a disc, a monogon, or a bigon. Hence $\phi''$ is indeed a euclidean singular foliation in $S^2$: however, this is excluded by Corollary 2.11. $\square$

**2.5. Transversal measure.** We equip every transverse simple arc to a foliation $\phi$ with a Borel measure as follows.

Let $\alpha$ be a transverse arc in the Fontana plane $\mathbb{F}$. The support of $\alpha$ has a Borel measure, induced by the linear 1-form $dy$. This measure is preserved by all transition functions, therefore it descends to any simple transverse arc $\alpha$ to a euclidean singular foliation $\phi$ in any surface $S$. We call this measure the *transversal measure* of $\alpha$. Its support $\alpha \cap \phi$ consists of disjoint segments.

In particular, a transverse arc or curve $\alpha$ has a *transversal length* $\ell(\alpha)$ defined as the total measure of $\alpha$.

REMARK 2.16. The transversal length is not bigger than the riemannian length, and the two lengths coincide when $\alpha$ is orthogonal to all leaves.

**2.6. From foliations to laminations.** We transform here foliations into measured geodesic laminations. Let $\phi$ be a foliation on a surface $S_g$ of genus $g \geqslant 2$ and fix an auxiliary hyperbolic metric on $S_g = \mathbb{H}^2/_\Gamma$. The foliation $\phi$ lifts naturally to a foliation $\tilde{\phi}$ on $\mathbb{H}^2$, whose leaves and singular points are the counter-images of leaves and singular points of $\phi$.

LEMMA 2.17. *A non-singular leaf of $\tilde{\phi}$ has two distinct endpoints in $\partial\mathbb{H}^2$.*

PROOF. We may suppose that $\phi$ is reduced, since reducing $\phi$ moves the leaves of a bounded amount and hence does not affect the limit behaviour of its lifts in $\tilde{\phi}$. Choose any geodesic pants decomposition $\mu$. By Proposition 2.14 we may transform $\phi$ via an isotopy so that $\mu$ is transverse and forms no bigon with $\phi$. Its counterimage $\tilde{\mu}$ in $\mathbb{H}^2$ consists of pairwise ultraparellel disjoint lines.

Consider a non-singular leaf $\beta$ of $\phi$ and a lift $\tilde{\beta}$ in $\mathbb{H}^2$. If $\beta$ is contained in a pair-of-pants, it is a simple closed curve parallel to a component of $\mu$, hence homotopic to a geodesic, and $\tilde{\beta}$ is homotopic to a line of $\tilde{\mu}$ with the same endpoints. Otherwise, $\beta$ crosses $\mu$ without forming bigons by Proposition 2.15: therefore $\tilde{\beta}$ crosses $\tilde{\mu}$ without forming bigons, and hence it intersects each line in $\tilde{\beta}$ at most once.

Represent $\overline{\mathbb{H}^2}$ as the closed euclidean disc. The curve $\tilde{\beta}$ intersects infinitely many lines in $\mathbb{H}^2$ as in Fig. 12. Since $\tilde{\mu}$ is a discrete set of lines in $\mathscr{G}$, for any $\varepsilon > 0$ there are only finitely many lines whose endpoints have distance (in the euclidean metric) bigger than $\varepsilon$. The curve $\tilde{\beta}$ intersects infinitely many lines, whose endpoints are nested and infinitely close, hence $\tilde{\beta}$ has two endpoints in $\partial\mathbb{H}^2$. These are distinct since the lines in $\tilde{\mu}$ are ultraparallel. $\square$

The *straightening* of $\tilde{\phi}$ is the set of lines in $\mathbb{H}^2$ obtained by replacing every non-singular leaf of $\tilde{\phi}$ with the line having the same endpoints.

LEMMA 2.18. *The straightening of $\tilde{\phi}$ consists of disjoint lines.*

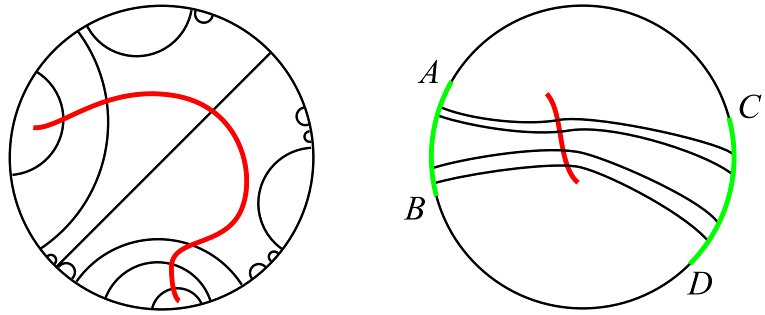



FIGURE 12. The lift $\tilde{\beta}$ of a leaf intersects the ultraparallel lines $\tilde{\mu}$, each at most once: therefore it has disjoint limits as $t \to \pm\infty$ (left). An arc crosses a first and a last leaf of $\tilde{\phi}$, and all other leaves it crosses have endpoints in the segments $AB$ and $CD$ (right).

PROOF. Leaves in $\tilde{\phi}$ are disjoint, hence their endpoints are unlinked. □

The closure of the straightening of $\tilde{\phi}$ is a geodesic lamination $\tilde{\lambda}$ in $\mathbb{H}^2$. The whole construction is $\Gamma$-invariant, hence $\tilde{\lambda}$ also is. Therefore $\tilde{\lambda}$ descends to a geodesic lamination $\lambda$ on $S$, called the *straightening* of $\phi$.

REMARK 2.19. The straightening induces a natural map

$$\{\text{leaves of } \phi\} \longrightarrow \{\text{leaves of } \lambda\}$$

which may be neither injective nor surjective: many leaves may straighten to the same one, and some leaves in $\lambda$ were added in the closure.

We now define a transverse measure on $\lambda$. Informally, the transverse measure on $\phi$ induces one on $\lambda$. More formally, we define a measure on $\mathscr{G}$ with support $\tilde{\lambda}$.

Let $\alpha$ be an arc in $\mathbb{H}^2$ transverse to $\tilde{\phi}$, which intersects every leaf at most once. Let $\alpha^* \subset \alpha$ be the set of points contained in a non-singular leaf of $\tilde{\phi}$. We may orient all leaves intersecting $\alpha^*$ coherently and define a map

$$s_\alpha \colon \alpha^* \to \partial\mathbb{H}^2 \times \partial\mathbb{H}^2$$

that sends $x$ to the oriented pair of endpoints of the leaf containing $x$. By construction each component of $s_\alpha$ is monotone, hence measurable: therefore $s_\alpha$ is a measurable function.

Let $[A, B]$ and $[C, D]$ be as in Fig. 12-(right) the smallest closed segments in $\partial\mathbb{H}^2$ (which might reduce to points) such that $[A, B] \times [C, D]$ contains the image of $s_\alpha$. We have $(A, B) \cap (C, D) = \emptyset$ and every leaf of $\tilde{\phi}$ with endpoints in $(A, B)$ and $(C, D)$ crosses $\alpha$. We equip the open set

$$U_\alpha = (A, B) \times (C, D) \subset \mathscr{G}$$

with the push-forward $\mu_\alpha$ via $s_\alpha$ of the transverse measure on $\alpha$ induced by $\tilde{\phi}$. We do this for all transverse arcs $\alpha$ that cross each leaf at most once.

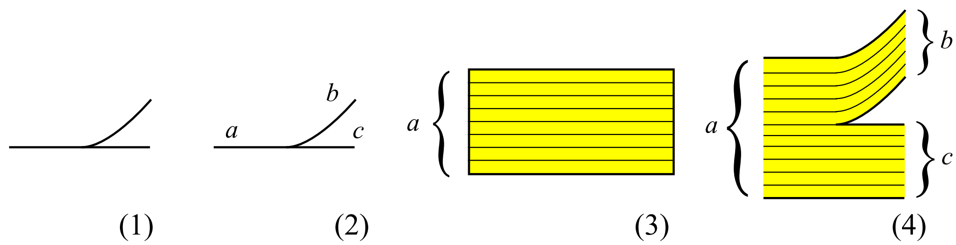



FIGURE 13. A switch (1), the switch condition requires that $a = b+c$ (2), to construct a foliation, replace every branch with weight $a$ with a square of width $a$ (3), every switch creates a singular point (4).

PROPOSITION 2.20. *There is a unique measure $\mu$ on $\mathscr{G}$ with support $\tilde{\lambda}$ which restricts to $\mu_\alpha$ on each $U_\alpha$.*

PROOF. We check that $\mu_\alpha = \mu_{\alpha'}$ whenever $U_\alpha \cap U_{\alpha'} \neq \emptyset$. Let $\phi' \subset \phi$ be the leaves having endpoints in $U_\alpha \cap U_{\alpha'}$. These leaves cross both $\alpha$ and $\alpha'$ and hence induce a bijection $i\colon \alpha \cap \phi' \to \alpha' \cap \phi'$ such that $s_\alpha = s_{\alpha'} \circ i$. The map $i$ preserves the transverse measure, since the close leaves in $\phi'$ stay at the same euclidean distance when they go from $\alpha$ to $\alpha'$. Hence $\mu_\alpha = \mu_{\alpha'}$.

Now we apply Proposition 3.4 from Chapter 1 to get a unique $\mu$ on the union of the open sets $U_\alpha$. Using a pants decomposition as in the proof of Proposition 2.17 one sees that every non-singular leaf is contained in some $U_\alpha$: therefore $\mu$ is supported on $\tilde{\lambda}$. □

The measure $\mu$ is clearly $\Gamma$-invariant (as everything is) and therefore descends to a measured lamination $\lambda$ in $S$.

We can now transform canonically every foliation $\phi$ into a measured geodesic lamination $\lambda$. Let $\phi'$ be obtained by reducing $\phi$. For a multicurve $\mu$, we define $i(\phi, \mu)$ as the transversal length of any isotopic representative for $\mu$ which is transverse to $\phi'$ and forms no bigon with it, which exists by Proposition 2.14. This number is well-defined and useful because of the following.

PROPOSITION 2.21. *We have $i(\phi, \mu) = i(\lambda, \mu)$.*

PROOF. Every component of $\mu$ has no bigons with the leaves of $\phi$ by Proposition 2.15 and hence lifts to a segment in $\mathbb{H}^2$ that intersects every leaf at most once. The measure of the leaves it crosses equals its transverse measure by the definition above. □

**2.7. Train tracks.** Laminations are better constructed via foliations, and foliations are better modeled via train tracks.

A *train track* on a closed surface $S$ is a closed subset $\tau \subset S$ built by taking a finite set of points (called *vertices* or *switches*) and joining them with disjoint arcs called *branches*. We require that every switch looks locally like Fig. 13-(left): there are three branches all with the same tangent line, two from one side and one from the other. As for foliations, we also require

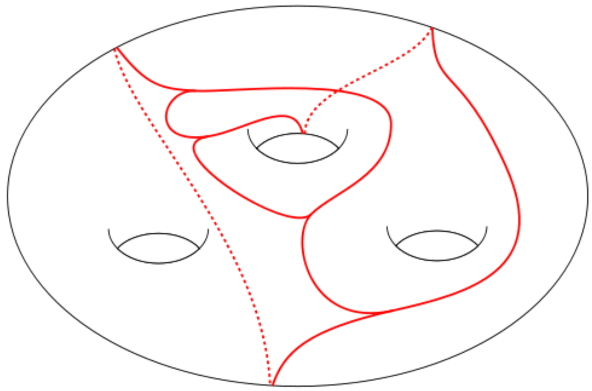


FIGURE 14. A train track on $S_3$.

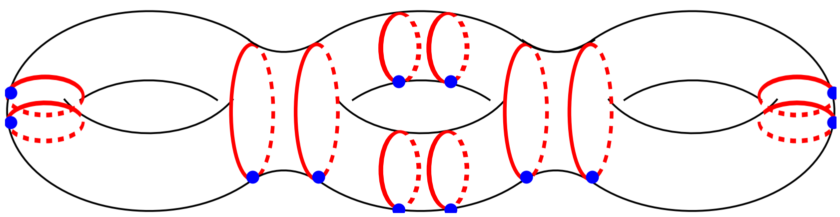


FIGURE 15. Decompose $S$ into pants and annuli and fix a blue point in each curve.

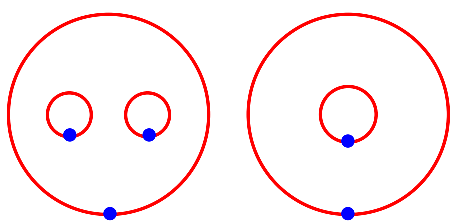


FIGURE 16. Identify each pant and annulus with one of these two fixed models.

that no complementary region in $S \setminus \tau$ is a disc, a monogon, or a bigon. An example is shown in Fig. 14.

A *weight system* on $\tau$ is constructed by assigning a non-negative real number, called *weight*, to each branch of $\tau$, such that the *switch conditions* hold: at every switch as in Fig. 13-(2) we must have $a = b + c$.

A weighted train track determines a foliation as follows: replace every branch with weight $a$ with a euclidean rectangle of width $a$ and arbitrary length as in Fig. 13. Thanks to the switch conditions, these rectangle glue nicely at each switch as in Fig. 13, producing a singular vertex.

**2.8. A parametrization for $\mathscr{ML}$.** Let $S = S_g$ with $g \geqslant 2$ be oriented. We want to parametrize the space $\mathscr{ML}$ of all geodesic laminations on $S$, and to this purpose we fix a frame similar to the one needed in Fenchel-Nielsen coordinates.

Let a *decomposition into pants and annuli* be a multicurve as in Fig. 15, obtained from a pants decomposition by duplicating each curve. A *frame* here consists of a decomposition of $S$ into pants and annuli, with a marked (blue) point in each closed curve as in Fig. 15: here we fix a diffeomorphism (preserving orientation and marked points) between each pants and annulus with one model from Fig. 16.

We fix an arbitrary frame for $S_g$.

Definition 2.22. Assign a triple $a_i, b_i, c_i$ of non-negative numbers to each annulus of the frame, such that one of the following holds:

$$a_i = b_i + c_i, \quad b_i = c_i + a_i, \quad c_i = a_i + b_i.$$

This assignment is called a *coloring.*

We transform a coloring into a weighted train track as follows. Every pair-of-pants is adjacent to three annuli: the $a_i$ colorings of these annuli form a triple $a_i, a_j, a_k$. We insert a portion of train track as in Fig. 17: its shape depends on the position of $[a_i, a_j, a_k]$ in $\mathbb{RP}^2$ and its weights depend linearly on $a_i, a_j, a_k$.

We extend the train track inside an annulus colored with $a_i, b_i, c_i$ as shown in Fig. 18. A coloring determines a weighted train track, hence a foliation, hence a measured geodesic lamination.

Theorem 2.23. *The construction induces a bijection*

$$\big\{\text{colorings}\big\} \longleftrightarrow \mathscr{ML}.$$

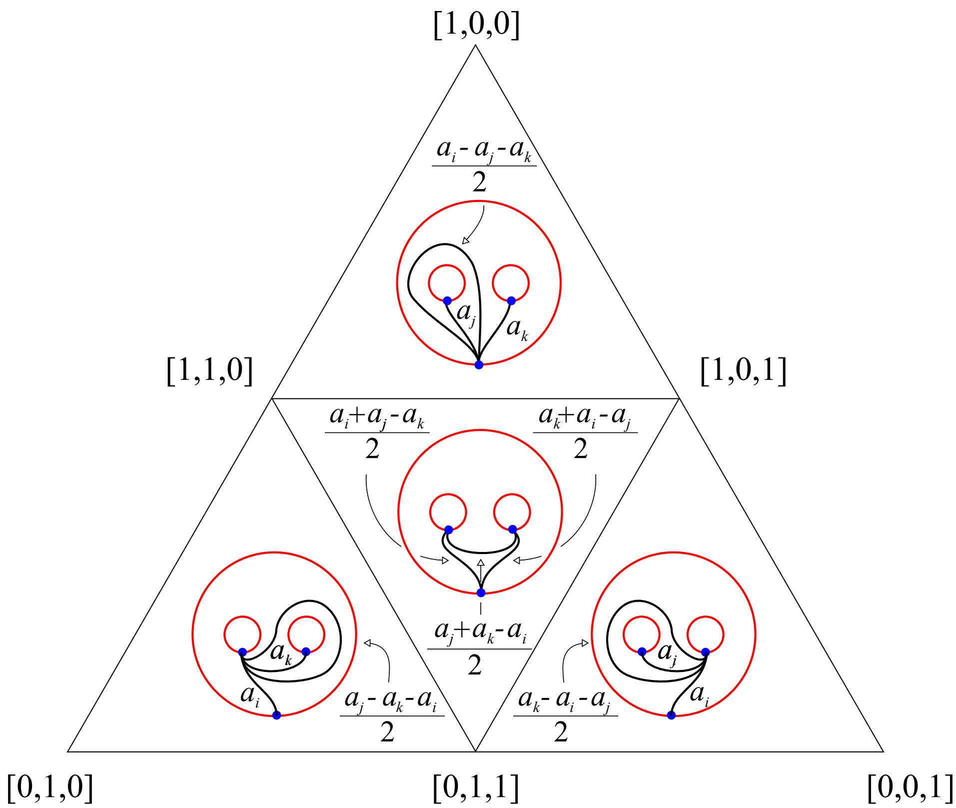



Figure 17. The portion of train track determined by the triple $(a_i, a_j, a_k)$. The triangle in $\mathbb{RP}^2$ with vertices $[1,0,0], [0,1,0], [0,0,1]$ subdivides into four triangles and the shape of the train track depends on the position of $P = [a_i, a_j, a_k]$. When $P$ lies in the frontier of two triangles some branch has weight zero and the different shapes actually coincide after deleting this branch.

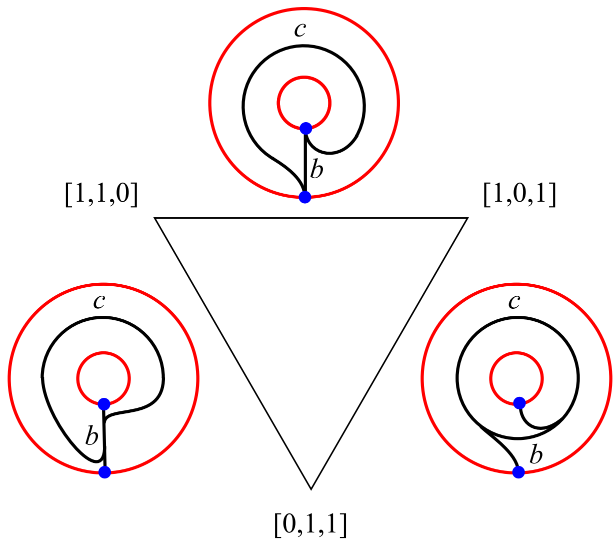



FIGURE 18. The portion of train track determined by the triple $(a_i, b_i, c_i)$. The boundary of the triangle in $\mathbb{RP}^2$ with vertices $[1,0,0], [0,1,0], [0,0,1]$ subdivides into three segments and the shape of the train track depends on the position of $P = [a_i, b_i, c_i]$ in this boundary. When $P$ lies in a vertex some branch has weight zero and the two shapes coincide after deleting this branch.