Geometry III/V

Anna Felikson

Durham University, 2023–2024

I devote this course to solidarity with Ukraine,
to solidarity with Israel,
with all people in hardship,
and with all people standing for peace and freedom!

Contents

Here "(NE)" marks non-examinable sections.

0 Introduction and History

0.1 Introduction

What to expect or 8 reasons to expect difficulties.

Our brain has two halves: one is responsible for multiplication of polynomials and languages, and the other half is responsible for orientation of figures in space and the things important in real life.
Mathematics is geometry when you have to use both halves.

Vladimir Arnold

Geometry is an art of reasoning well from badly drawn diagrams.

Henri Poincaré

1. Structure of the course:

- It will be a zoo of different 2-dimensional geometries - including Euclidean, affine, projective, spherical and Möbius geometries, all of which will appear as some aspects of hyperbolic geometry.

  Why to study all of them?
  - They are beautiful!
  - we will need all of them to study hyperbolic geometry.

  Why to study hyperbolic geometry?
  - Important in topology and physics, for example.

  Example. When one looks at geometric structures on 2-dimensional closed surfaces, one can find out that only the sphere and torus carry spherical and Euclidean geometries on them, and infinitely many other surfaces (all other closed surfaces) are hyperbolic (see Fig. 1).
  (Given the time there will be more on that at the very end of the second term).

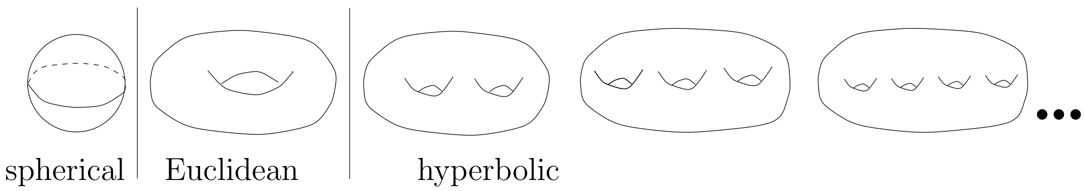


Figure 1: Geometric structures on surfaces.

- There will be just a bit on each geometry, hence the material may seem too easy.
- But it will get too difficult if you will miss something (as we are going to use extensively almost everything...)

2. Two ways of doing geometry: "synthetic" and "analytic"

- "Synthetic" way:
  - List axioms and definitions.
  - Then formally derive theorems.
  Question: is there any object satisfying the axioms?
  - Build a "model": an object satisfying the axioms (and hence, theorems).

- "Analytic" way:
  - Build a model
  - Work in the model to prove theorems (using properties of the model).

The same object may have many different models.

Example 0.1. A group $G_2 = \{e, r\} = \langle r | r^2 = e \rangle$
(Group with 2 elements, e, r, with one generator r and relation $r^2 = e$).
Model 1: Let r be a reflection on $\mathbb{R}^2$ (and e an identity map).
Model 2: $\{1, -1\} \in \mathbb{Z}$ with respect to multiplication.

We will sometimes use different models for the same geometry - to see different aspect of that geometry.

3. "Geometric" way of thinking:

Example 0.2. Claim. *Let ABC be a triangle, let M and N be the midpoints of AB and BC. Then $AC = 2MN$.*

We will prove the claim in two ways: geometrically and in coordinates. Geometric proof will be based on Theorem0.3.

Notation: given lines l and m, we write $l||m$ when l is parallel to m.

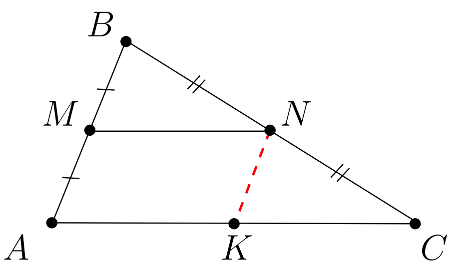
Theorem 0.3. *If ABC is a triangle, $M \in AB$, $N \in BC$, then $MN||AC \Leftrightarrow \frac{BA}{BM} = \frac{BC}{BN}$.*

Proof. (Geometric proof):
1) $MN||AC$ (by Theorem 0.3).
2) Draw $NK||AM$, $K \in AC$ (see Fig. 2, left).
3) Then $AK = KC$ (by Theorem 0.3).
4) The quadrilateral $AMNK$ is a parallelogram
(by definition of a parallelogram - as it has two pairs of parallel sides).
5) Hence, $MN = AK$ (by a property of a parallelogram).
6) $AC = 2AK = 2MN$ (by steps 3 and 5).

□

Proof. (Computation in coordinates):
We can assume that $A = (0, 0)$, $C = (x, 0)$, $B = (z, t)$ (see Fig. 2, right).
Then $M = (\frac{z}{2}, \frac{t}{2})$, $N = (\frac{x+z}{2}, \frac{z}{2})$.
Therefore, $MN^2 = (\frac{x+z}{2} - \frac{z}{2})^2 + (\frac{t}{2} - \frac{t}{2})^2 = (\frac{x}{2})^2$. Hence, $MN = x/2$, while $AC = x$.

□



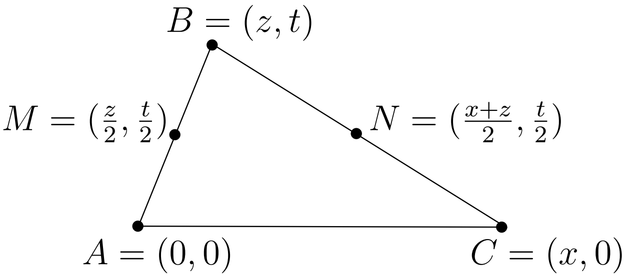


Figure 2: Two proofs of the theorem about midlines.

Note that even in the second proof we used geometry to simplify the computation: we assumed that $A = (0, 0)$, i.e. that all points of the plane are equally good, and that after taking A to the origin we can rotate the whole picture so that C get to the horizontal line.

4. We will use some results from Euclidean geometry without reproving.
- We need some basics.
- The complete way from axioms takes time.
- It is not difficult (was previously taught in schools).
- You can find proofs in books (will give some references).
- Hopefully, by now you have already mastered logical/mathematical thinking (and don't need the course on Euclidean geometry as a model for mastering them).

5. We will use many diagrams:
- They are useful
- but be careful: wrong diagrams may lead to mistakes.

Example 0.4. "Proof" that all triangles are isosceles (with demystification): http://jdh.hamkins.org/all-triangles-are-isosceles/

6. Problem solving in Geometry
- Is not algorithmic (one needs practice!)
- Solution may be easy – but how to find it?
(additional constructions? which model to use? which coordinates to choose? ...
- all needs practice!)

For getting the practice we will have Problem Classes and Assignments:
- weekly sets of assignments;
- some questions will be starred - to submit for marking fortnightly (via Gradescope).
- other questions - to solve!
- There will be hints - use them if you absolutely don't know how to start the question without them (it is much better to attempt the questions with hints than just to read the solutions).

7. “Examples” will be hard to tell from “Theory”:
“Problem”=”one more theorem”
“Proof of a Theorem”=”Example on problem solving”.

8. Group approach to geometry
Klein’s Erlangen Program: In 1872, Felix Klein proposed the following:
each geometry is a set with a transformation group acting on it.
To study geometry is the same as to study the properties preserved by the group.

Example 0.5. Isometries preserve distance;
Affine transformations preserve parallelism;
Projective transformations preserve collinearity;
Möbius transformations preserve property to lie on the same circle or line.

Why to speak about possible difficulties now?
- not with the aim to frighten you but
- to make sure you are aware of them;
- to inform you that they are in the nature of the subject;
- to inform you that I know about the difficulties- and will try my best to help;
- to motivate you to ask questions.

Remark. For seven top reasons to enjoy geometry check Chapter 0 here:
http://www1.maths.leeds.ac.uk/ kisilv/courses/math255.html

0.2 Axiomatic approach to geometry

Ptolemy I: *Is there any shorter way than one of Elements?*
Euclid: *There is no royal road to geometry.*

Proclus

“One must be able to say at all times instead of points, straight lines and planes - tables, chairs and beer mugs.”

David Hilbert

Geometry in Greek: $\gamma\epsilon\omega\mu\epsilon\tau\rho\iota\alpha$, i.e. measure of land (“geo”=land, “metry”=measure).

Brief History:

- Origin: Ancient Egypt $\approx$ 3000 BC
(measuring land, building pyramids, astronomy).
- First records: Mesopotamia, Egypt $\approx$ 2000 BC.
Example: Babylonians did know Pythagorean theorem
at least 1000 years before Pythagoras.
- Greek philosophy brought people to the idea
that geometric statements should be deductively proved.
- Euclid ($\approx$ 300 BC) realised that the chain of proofs cannot be endless:
A holds because of B (Why B holds?)
B holds because of C (Why C?)
C ....

  To break this infinite chain ... $\Rightarrow C \Rightarrow B \Rightarrow A$ we need to
  - Accept some statements as axioms without justification;
  - Agree on the rules of logic.

Euclid's Postulates:

1. For every point A and for every point B not equal to A there exists a unique line that passes through A and B.
2. For every segment AB and for every segment CD there exists a unique point E such that B is between A and E and such that segment CD is congruent to segment BE.
3. For every point O and every point A not equal to O, there exists a circle with centre O and radius OA.
4. All right angles are congruent to each other.
5. (Euclid's Parallel Postulate) For every line l and for every point P that does not lie on l, there exists a unique line m passing through P that is parallel to l.

In "Elements" Euclid derives all known by that time statements of geometry and number theory from these five postulates.

Hilbert's axioms

By XIXth century it is clear that Euclid's axioms are not sufficient: Euclid still used some implicit assumptions.

Example 0.6 (Euclid's Theorem 1)**.** : *There exists an equilateral triangle with a given side AB.*

Euclid's proof:
- Draw a circle C_A centred at A of radius AB (see Fig. 5).
- Draw a circle C_B centred at B of radius AB.
- Take their intersection $C = C_A \cap C_B$ and show that $\triangle ABC$ is equilateral.

What is wrong with the proof: Why do we know that the circles do intersect?

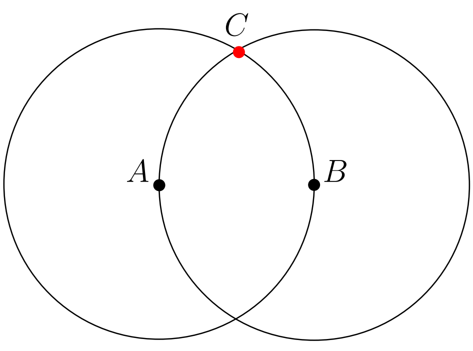


Figure 3: Euclid's proof of existence of equilateral triangle.

This shows that we need to have more axioms. Hilbert has developed such a system of axioms, which contains 5 groups of axioms (roughly corresponding to Euclid's postulates). See handout for the list.
You don't need to memorise - neither Euclid's nor Hilbert's axioms!

Example 0.7. Given a triangle ABC and a line l crossing the segment AB, can we state that l we cross the boundary of ABC again on it's way "out of the triangle"? See Fig. 4.

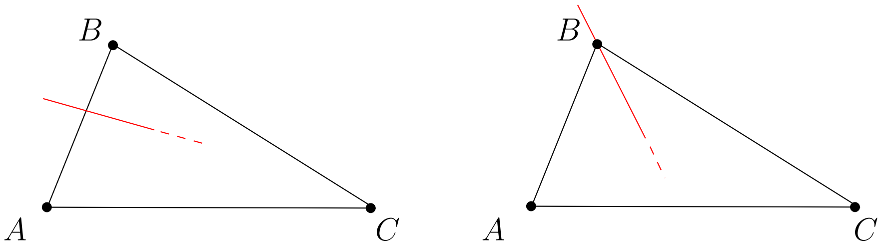


Figure 4: Pasch's Theorem and Crossbar Theorem.

If we want to derive this obvious fact from the axioms, we need to work quite a lot, in particularly, using Betweenness Axiom BA4. It will go as follows.

Definition 0.8. Given a line l and points $A, B \notin l$ we say that A and B are *on the same side* of l if $A = B$ or the segment AB does not intersects l. Otherwise, A and B are *on the opposite sides* of l. We will denote these situations $A, B|*$ and $A|B$ respectively (when it is clear which line is considered).

Axiom 0.9 (BA4, Plane separation)**.** (a) $A, C|*$ and $B, C|*$ imply $A, B|*$.

(b) $A|C$ and $B|C$ imply $A, B|*$.

Remark 0.10. The Axiom BA4 guarantees that the geometry we get is 2-dimensional.

Theorem 0.11 (Pasch's Theorem)**.** *Given a triangle ABC, line l, and points $A, B, C \notin l$. If l intersects AB then l intersects either AC or BC.*

Proof. (1) By Definition 0.8, we have $A|B$.
(2) Since $C \notin l$, BA4(a) implies that either $A, C|*$ or $B, C|*$.
(3) Suppose that $A, C|*$. By BA4(a), this implies that $C|B$ (otherwise we have $A, B|*$ in contradiction to (1)). Therefore, $l \cap BC \neq \emptyset$ (by Definition 0.8).
(4) The case if $B, C|*$ is considered similarly.

□

Remark 0.12. In the case, when l enters the triangle ABC through a vertex C one can show that l intersects AB (this statement is called Crossbar Theorem and its proof is more than twice longer).

Remarks

1. We will not work with axioms (neither in Euclidean geometry no in any other).
2. We appreciate this magnificent building of knowledge and use theorems of Euclidean geometry when we need them.

3. Some basic theorems are listed in the handout out Euclidean geometry (with brief ideas of proofs and references, where available).

4. More detailed treatment of basics can be found in
M. J. Greenberg, *Euclidean and Non-Euclidean Geometries*, San Francisco: W. H. Freeman, 2008.

5. Sometimes one can find many proofs of the same theorem.
For example, see https://www.cut-the-knot.org/pythagoras/ for 122 proofs of Pythagorean theorem.

What to do with the list of Theorems?

1. You don't need to memorise!
(This is just an index for the references later on).

2. **Read, understand** and **illustrate** the statements (to be aware of them).

3. Do HW Question 1.1 (we will collect the data anonymously during Lecture 3!).

Remark 0.13. Axiomatic approach is designed to eliminate geometry from geometry. Now belongs to the history of mathematics. However, some elements of it still could be useful as parts of school education as
- an example of logical arguing;
- a demonstration that even "evident" statements should be justified.

Example: "My opinion is the right one".

Remark 0.14. Hilbert's axiom system is shown to be

1. Consistent (i.e. there exists a model for it).

2. Independent (i.e. when removing any axiom one gets another set of theorems).

3. Complete (for any statement A in this language either holds "A" or its negation "not A").

Remark 0.15 (Hilbert's completeness and Gödel's incompleteness)**.** One may ask why completeness of Hilbert's system of axioms does not contradict to Gödel's Incompleteness Theorem, stating that:

> **Gödel's Incompleteness Theorem.** *Any consistent formal system F within which a certain amount of elementary arithmetic can be carried out is incomplete.*

In other words, Gödel's Incompleteness Theorem states that a theory cannot at the same time: (1) to be consistent, (2) to be complete, (3) to contain elementary arithmetic.

Here, "to contain elementary arithmetic" means that the theory has a universal tool to represent addition *and* multiplication. In particular, geometry allows a sort of addition (given two segments of lengths a and b, we can construct a segment of length $a + b$). However, there is no similar procedure for multiplication.

This shows that there is no contradiction in geometry being consistent and complete. It just does not contain arithmetic (though, we are not providing a proof of that).

0.3 References

- A further discussion of Klein's Erlangen Program can be found in Section 5 of Nigel Hitchin, *Projective Geometry*, Lecture notes. Chapters 1, 2, 3, 4.
  (See also "Other Resources" on DUO if you want to have all chapters in one pdf).

- Elementary exposition of most basic facts of Euclidean geometry can be found in A. D. Gardiner, C. J. Bradley, *Plane Euclidean Geometry*, UKMT, Leeds 2012.
  (The book is available from the library).

- Elementary but detailed exposition of basic facts of Euclidean geometry (and of many other topic of the current module):
  A. Petrunin, *Euclidean plane and its relatives. A minimalist introduction.*

- For the detailed treatment of axiomatic fundations of Euclidean geometry see M. J. Greenberg, *Euclidean and Non-Euclidean Geometries*, San Francisco: W. H. Freeman, 2008.
  (The book is available from the library).

- *Euclid's "Elements"*, complete text with all proofs, with illustration in Geometry Java applet, website by David E. Joyce.

1 Euclidean Geometry

1.1 Isometry group of Euclidean plane, $Isom(\mathbb{E}^2)$.

From now all, we will forget about axiomatic and will use some facts of Euclidean geometry as "preknown".

By Euclidean plane $\mathbb{E}^2$ we will understand $\mathbb{R}^2$ together with a distance function $d(A,B)$ on it satisfying the following axioms M1-M3 of a metric:

Definition 1.1. A distance on a space X is a function
$d : X \times X \to \mathbb{R}$, $(A,B) \mapsto d(A,B)$ for $A, B \in X$ satisfying

M1. $d(A,B) \geq 0 \quad (d(A,B) = 0 \Leftrightarrow A = B)$;

M2. $d(A,B) = d(B,A)$;

M3. $d(A,C) \leq d(A,B) + d(B,C)$ (triangle inequality).

Remark. Triangle inequality appears in the list of Euclidean facts as E25. It was proved using Cauchy-Schwarz inequality in Linear Algebra I, see also Section 1 of
G. Jones, Algebra and Geometry , *Lecture notes,*
which you can find in "Other Resources" on DUO.

We will use the following two models of Euclidean plane:
a Cartesian plane: $\{(x,y) \mid x,y \in \mathbb{R}\}$ with the distance $d(A_1,A_2) = \sqrt{(x_1-x_2)^2 + (y_1-y_2)^2}$;
a Gaussian plane: $\{z \mid z \in \mathbb{C}\}$, with the distance $d(u,v) = |u-v|$.

Definition 1.2. An isometry of Euclidean plane E^2 is a distance-preserving transformation of $\mathbb{E}^2$, i.e. a map $f : \mathbb{E}^2 \to \mathbb{E}^2$ satisfying $d(f(A), f(B)) = d(A,B)$ for every $A, B \in \mathbb{E}^2$.

We will show that isometries of $\mathbb{E}^2$ form a group, but first we recall the definition.

Definition. A set G with operation $\cdot$ is a group if the following for properties hold:

1. (Closedness) $\forall g_1, g_2 \in G$ have $g_1 \cdot g_2 \in G$;
2. (Associativity) $\forall g_1, g_2, g_3 \in G$ have $(g_1 \cdot g_2) \cdot g_3 = g_1 \cdot (g_2 \cdot g_3)$;
3. (Identity) $\exists e \in G$ such that $e \cdot g = g \cdot e = g$ for every $g \in G$;
4. (Inverse) $\forall g \in G\ \exists g^{-1} \in G$ s.t. $g \cdot g^{-1} = g^{-1} \cdot g = e$.

Theorem 1.3. *(a) Every isometry of $\mathbb{E}^2$ is a one-to-one map.*

(b) A composition of any two isometries is an isometry.

(c) Isometries of $\mathbb{E}^2$ form a group (denoted $Isom(\mathbb{E}^2)$) with composition as a group operation.

Proof. (a) Let f be an isometry. By M1, if $f(A) = f(B)$ then $d(f(A), f(B)) = 0$. So, by definition of isometry, $d(A,B) = 0$, which by M1 implies that $A = B$. Hence, f is injective.

Sketch of proof of surjectivity:

- Suppose $X \notin f(\mathbb{E}^2)$. Let $y = f(A)$.
- Consider a circle $C_A(r)$ centred at A of radius $r = d(X,Y)$. Notice that $f(C_A(r)) \subset C_y(r)$.
- Take $B \in C_A(r)$, consider $f(B) \in C_y(r)$.
- There are two points on $C_A(r)$ on any given distance smaller than $2r$ from B. Hence, $C_A(r)$ contains two points on distance $d(f(B),X)$. Therefore, $X \in f(C_A(r))$. The contradiction proves surjectivity, and (a) is done.

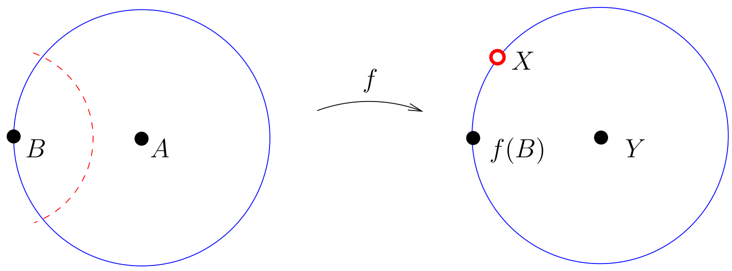


Figure 5: To the proof of surjectivity of isometry.

(b) Given two isometries f and g, we need to check that the composition $g \circ f$ is an isometry. Indeed,

$$d(g(f(A)), g(f(B)) \stackrel{g}{=} d(f(A), f(B)) \stackrel{f}{=} d(A,B),$$

where the first (resp. second) equality holds since g (resp. f) is an isometry.

(c) We need to prove 4 properties (axioms of a group):
1. Closedness is proved in (b).
2. Associativity follows from associativity of composition of maps.
3. Identity $e := id_{\mathbb{E}^2}$ is the map defined by $f(A) = A \ \forall A \in \mathbb{E}^2$. It clearly belongs to the set of isometries.
4. Inverse element g^{-1} does exist as g is one-to-one (and it is an isometry).

□

Example 1.4. Examples of isometries of $\mathbb{E}^2$:

- Translation $T_t : a \mapsto a + t$;
- Rotation $R_{\alpha,A}$ about centre A by angle α.
  On complex plane, $R_{\alpha,0}$ writes as $z \mapsto e^{i\alpha} z$;
- Reflection r_l in a line. Example: if the line l is the real line on $\mathbb{C}$, then $r_l : z \to \bar{z}$.
  For a general formula of reflection: see HW 2.7.
- Glide reflection: given a vector a and a line l parallel to a, consider $t_a \circ r_l = r_l \circ t_a$.

Definition 1.5. Let ABC be a triangle labelled clockwise.
An isometry f is orientation-preserving if the triangle $f(A)f(B)f(C)$ is also labelled clockwise. Otherwise, f is orientation-reversing.

Proposition 1.6 (Correctness of Definition 1.5)**.** *Definition 1.5 does not depend on the choice of the triangle ABC.*

Proof. Suppose that $\triangle ABC$ has the same orientation as $f(ABC)$. Take a point D on the same side of the line AB as C. Then $\triangle ABD$ has the same orientation as $f(ABD)$ (indeed, otherwise the segment $f(CD)$ does intersect the segment $f(AB)$ while AB and CD are disjoint; this would violate that f is a bijection). Hence, given the points A, B, Definition 1.5 does not depend on the choice of C.

Now we change points one by one moving from any triangle to any other as follows: $ABC \to A'BC \to A'B'C \to A'B'C'$. (One should be a bit more careful here if some triples of points are collinear, but then we just insert an extra step and may be change the order. We skip the details here).

□

Example 1.7. Translation and rotation are orientation-preserving,
reflection and glide reflection are orientation-reversing.

Remark 1.8. Composition of two orientation-preserving isometries is orientation-preserving;
composition of an or.-preserving isometry and an or.-reversing one is or.-reversing;
composition of two orientation-reversing isometries is orientation-preserving.

Proposition 1.9. *Orientation-preserving isometries form a subgroup (denoted $Isom^+(\mathbb{E}^2)$) of $Isom(\mathbb{E}^2)$.*

Proof. We need to check the set $Isom^+(\mathbb{E}^2)$ forms a group, i.e. satisfies the four properties of a group:

1. Closedness follows from Remark 1.8;

2,3. Associativity and Identity follow in the same way as in the proof of Theorem 1.3.

4. Inverse element: consider $g \in Isom^+(\mathbb{E}^2)$ and let $g^{-1} \in Isom(\mathbb{E}^2)$ be the inverse in the big group. Suppose that g^{-1} is orientation-reversing. Then by Remark 1.8 $g \circ g^{-1}$ is also orientation-reversing, which contradicts to the assumption that $g \circ g^{-1} = e$ when considered in the whole group $Isom(\mathbb{E}^2)$. The contradiction shows that g^{-1} is orientation-preserving, and hence $Isom^+(\mathbb{E}^2)$ contains the inverse element.

□

Definition. Triangles $\triangle ABC$ and $\triangle A'B'C'$ are <u>congruent</u> (write $\triangle ABC \cong \triangle A'B'C'$) if $AB = A'B'$, $AC = A'C'$, $BC = B'C'$ and $\angle ABC = \angle A'B'C'$, $\angle BAC = \angle B'A'C'$, $\angle ACB = \angle A'C'B'$.

Theorem 1.10. *Let ABC and $A'B'C'$ be two congruent triangles. Then there exists <u>a unique</u> isometry sending A to A', B to B' and C to C'.*

Proof. <u>Existence</u>:

1. Let f_1 be any reflection sending $A \to A'$, $A' \to A$ (if $A \neq A'$, f_1 is unique and given by reflection with respect to perpendicular bisector to AA', see Fig. 6, left; if $A = A'$ we can take $f_1 = id$, identity map).

2. Let f_2 be a reflection s.t. $f_2(A') = A'$, $f_2(f_1(B)) = B'$. This f_2 does exist: it is given by reflection with respect to perpendicular bisector to BB', see Fig. 6, middle (denote the perpendicular bisector by l_2). Notice that $A' \in l_2$.
Exercise: Show that $A' \in l_2$ by using E14.

3. We have $A' = f_2(f_1(A))$, $B' = f_2(f_1(B))$.
If $f_2(f_1(C))$ and C' lie in the same half-plane with respect to $A'B'$, then the congruence $\triangle ABC \cong \triangle A'B'C'$ implies $C' = f_2(f_1(C))$: (indeed, in this case triangles $\triangle A'C'f_2(f_1(C))$ and $\triangle B'C'f_2(f_1(C))$ are isosceles, so the heights of these triangles dropped from the points A' and B' respectively are two different perpendicular bisectors for the segment $C'f_2(f_1(C))$, which contradicts to E9, see Fig. 6, right). So, $f_2 \circ f_1$ maps ABC to $A'B'C'$
If $f_2(f_1(C))$ and C' lie in different half-plane with respect to $A'B'$, apply $f_3 = r_{A'B'}$ (reflection with respect to $A'B'$), then use the above reasoning to see that $f_3 \circ f_2 \circ f_1$ maps ABC and $A'B'C'$.

Uniqueness: Suppose the contrary, i.e. there exist $f, g \in Isom(\mathbb{E}^2)$, $f \neq g$ such that $f : \triangle ABC \to \triangle A'B'C'$ and $g : \triangle ABC \to \triangle A'B'C'$. Then $\varphi := f^{-1} \circ g \neq id$ and $\varphi(\triangle ABC) = \triangle ABC$. Choose $D \in \mathbb{E}^2 : \varphi(D) \neq D$ (it exists as φ is non-trivial!). Then $d(A, D) = d(A, \varphi(D))$, $d(B, D) = d(B, \varphi(D))$, $d(C, D) = d(C, \varphi(D))$, which by E14 means that all three points A, B, C lie on the perpendicular bisector to $D\varphi(D)$. This contradicts to the assumption that ABC is a triangle.

$\square$

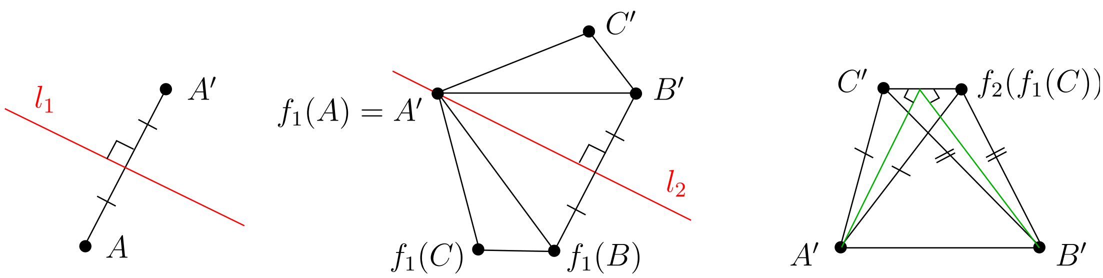


Figure 6: Isometry as a composition of reflections.

Corollary 1.11. *Every isometry of $\mathbb{E}^2$ is a composition of at most 3 reflections. (In particular, the group $Isom(\mathbb{E}^2)$ is generated by reflections).*

Remark 1.12. The way to write an isometry as a composition of reflections is not unique.

Example 1.13. We can write rotation and translation as compositions of two reflections (see (a) and (b) below; a glide deflection can be written as a composition of three reflection (see (c)).

(a) Let $l_1 || l_2$ be two parallel lines on distance d. Then $r_{l_2} \circ r_{l_1}$ is a translation by $2d$ along a line l perpendicular to l_1 and l_2.

(b) Let $0 = l_1 \cap l_2$ be two lines intersecting at O. Let φ be angle between l_1 and l_2. Then $r_{l_2} \circ r_{l_1}$ is a rotation about O through angle 2φ.

(c) Let l be a line, and a a vector parallel to l. To write the glide reflection $t_a \circ r_l$, use (a): consider two lies $l_1 || l_2$ orthogonal to l lying on the distance $a/2$ from each other. Then by (a) $t_a = r_{l_1} \circ r_{l_2}$, so that $t_a \circ r_l = r_{l_1} \circ r_{l_2} \circ r_l$.

Theorem 1.14 (Classification of isometries of $\mathbb{E}^2$). *Every non-trivial isometry of $\mathbb{E}^2$ is of one of the following four types: reflection, rotation, translation, glide reflection.*

Proof. We can see from the proof of Theorem1.10 that every isometry of $\mathbb{E}^2$ is a composition of at most 3 reflections. Consider possible compositions:

0. Composition of 0 reflections is an identity map id.
1. Composition of 1 reflection is the reflection.
2. Composition of 2 reflections is either translation or rotation (see Example 1.13).
3. Composition of 3 reflections: one can prove that is a glide reflection (this is not done in Example 1.13!), for the proof see HW 2.3.

□

Definition 1.15. Let $f\ Isom(\mathbb{E}^2)$. Then the set of fixed points of f is $Fix_f = \{x \in \mathbb{E}^2 \mid f(x) = x\}$.

Example 1.16. Fixed points of id, reflection, rotation, translation and glide reflection are $\mathbb{E}^2$, the line, a point, $\emptyset$, $\emptyset$ respectively.

Remark 1.17. Fixed points together with the property of preserving/reversing the orientation uniquely determine the type of the isometry.

Proposition 1.18. *Let $f, g \in Isom(\mathbb{E}^2)$.*

(a) $Fix_{gfg^{-1}} = gFix_f$;

(b) gfg^{-1} *is an isometry of the same type as* f.

Proof. (a) We need to proof that $g(x) \in Fix_{gfg^{-1}} \Leftrightarrow x \in Fix_f$.
See HW 3.2 for the proof.

(b) Applying (a) we see that fixed points of f and gfg^{-1} are of the same type, also they either both preserve the orientation or both reverse it. Hence, the isometries f and gfg^{-1} are of the same type by Remark 1.17

□

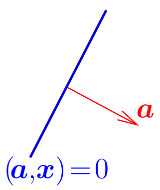
1.2 Isometries and orthogonal transformations

a. Isometries preserving the origin $O = (0, 0)$

- From HW 2.7 we see, that a reflection preserving O is a linear map:

$$\boldsymbol{x} \to A\boldsymbol{x} \quad A \in GL_2(\mathbb{R}).$$

More precisely, if l is a line through O and $\boldsymbol{a}$ a vector normal to l (i.e. the line l is given by equation $(\boldsymbol{a}, \boldsymbol{x}) = 0$, where $(*, *)$ is the dot product), then



$$r_l(\boldsymbol{x}) = \boldsymbol{x} - \frac{(\boldsymbol{a},\boldsymbol{x})}{(\boldsymbol{x},\boldsymbol{x})}\boldsymbol{a}.$$

- Every isometry preserving O is a composition of at most 2 reflections (this follows from the proof of Theorem 1.10, or, alternatively, from the classification of isometries). Hence, it is either an identity map, or a reflection or a rotation.

- So, if $f \in Isom(\mathbb{E}^2)$ and $f(O) = O$, then $f(\boldsymbol{x}) = A\boldsymbol{x}$ for some $A \in GL_2(\mathbb{R})$.

Proposition 1.19. *A linear map $f : \mathbf{x} \to A\mathbf{x}$, $A \in GL(2, \mathbb{R})$ is an isometry if and only if $A \in O(2)$, orthogonal subgroup of $GL(2, \mathbb{R})$ (i.e. iff $A^T A = I = AA^T$, where A^T is A transposed).*

Proof. See HW 3.3. □

b. General case

Let $(b_1, b_2) = f(O)$, denote $\boldsymbol{b} = (b_1, b_2)$. Then $t_{-b} \circ f(O)$ preserves O. So, in view of Proposition 1.19, $t_{-b} \circ f(\boldsymbol{x}) = A\boldsymbol{x}$ for some $A \in O_2(\mathbb{R})$, which implies that

$$f(\boldsymbol{x}) = t_b \circ (A\boldsymbol{x}) = A\boldsymbol{x} + \boldsymbol{b}.$$

Proposition 1.20. *(a) Every isometry f of $\mathbb{E}^2$ may be written as $f(\mathbf{x}) = A\mathbf{x} + \mathbf{t}$.*

(b) The linear part A does not depend on the choice of the origin.

Proof. (a) is already shown. (b) Move the origin to arbitrary other point $\boldsymbol{u} = (u_1, u_2)$ and denote by $\boldsymbol{y} = \boldsymbol{x} - \boldsymbol{u}$ the new coordinates (see Fig. 7). Then

$$f(\boldsymbol{y}) = f(\boldsymbol{x}) - \boldsymbol{u} = A\boldsymbol{x} + \boldsymbol{b} - \boldsymbol{u} = A(\boldsymbol{y} + \boldsymbol{u}) + \boldsymbol{b} - \boldsymbol{u} = A\boldsymbol{y} + (A\boldsymbol{u} + \boldsymbol{b} - \boldsymbol{u}).$$

□

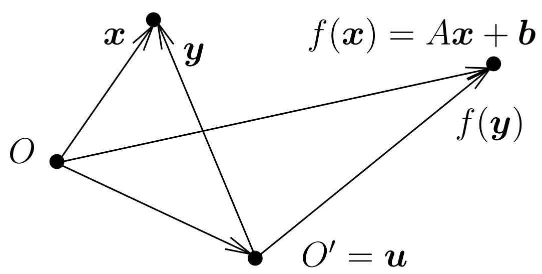


Figure 7: Linear part of isometry: independence of the origin.

Example 1.21. Let $A \in O_2(\mathbb{R})$ then $\det A = \pm 1$.

- Consider the reflection $r_{x=0}$ with respect to the line $x = 0$: $r_{x=0} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$. Clearly, in this case $\det A = -1$.

- Consider a rotation by angle α, $R_{O,\alpha} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}$. In this case $\det A = 1$.

Proposition 1.22. *Let $f(x) = A\mathbf{x} + \mathbf{t}$ be an isometry.*
f is orientation-preserving if $\det A = 1$ *and orientation-reversing if* $\det A = -1$.

Proof. First, notice that translation does not affect the orientation, so. we can assume that f preserve the origin. An origin-preserving isometry is either identity, or reflection, or rotation, and for all of them the statement holds.

□

Remark. Let l be a line through O forming angle α with the horizontal line $x = 0$. Then $r_l = g^{-1} r_{x=0} g$, where $g = R_{O,-\alpha}$ (check this!). So,

$$\det r_l = \det g^{-1} \ \det r_{x=0} \ \det g = -1.$$

Exercise 1.23. (a) Show that any two reflections are conjugate in $Isom(\mathbb{E}^2)$. (i.e. that given any two reflections r_1 and r_2 there exists an isometry $g \in Isom(\mathbb{E}^2)$ such that $r_1 = g^{-1} r_2 g$).

Hint. If l is a line not through the origin, then there exists a translation t such that $l' = t(l)$ is a line through the origin and $r_l = t^{-1} r_{l'} t$.

(b) Not all rotations are conjugate (only rotations by the same angle),
not all translations are conjugate (only the ones by the same distance)
and not all glide reflections are conjugate (only the ones with translational part by the same distance).

Proposition 1.24. *Let $A, C \in l \in \mathbb{E}^2$. Then the line l gives the shortest path from A to C.*

Proof. Idea: approximate the path from A to C by a broken line $AA_1A_2A_3 \dots A_{n-1}A_nC$ and apply triangle inequality $|AC| \le |AB| + |BC|$ repeatedly:

$$|AC| \le |AA_1| + |A_1C| \le |AA_1| + |A_1A_2| + |A_2C| \le \dots \le |AA_1| + \cdot |A_nC|,$$

with at least one inequality being strict if $AA_1A_2A_3 \dots A_{n-1}A_nC \ne AC$.

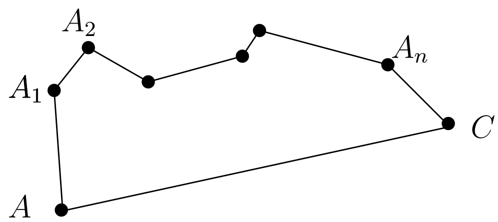


Figure 8: A broken line approximating a path.

Analytically: given a path $\gamma : [0, 1] \to \mathbb{E}^2$ with $\gamma(0) = A = (0, 0)$ and $\gamma(1) = C = (c, 0)$, write

$$l(\gamma\Big|_A^C) = \int_0^1 \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} dt \ge \int_0^1 \sqrt{\left(\frac{dx}{dt}\right)^2} dt$$
$$= \int_0^1 \left|\frac{dx}{dt}\right| dt \ge \int_0^1 \frac{dx}{dt} dt = x(t)\Big|_0^1 = x(1) - x(0) = b - 0 = d(A, B).$$

□

1.2.1 Remarks on groups

There are two ways to define a group G:

- To describe the set of elements of the group G and the group operation.

  Example: Matrix groups are usually defined in this way, i.e. $GL(2, \mathbb{R})$ (nondegenerate real 2×2 matrices), $SL(n, \mathbb{Z})$ ($n \times n$ real matrices with $det = 1$), etc.... The group operation in these groups is a matrix multiplication

- To describe the group G by "generators and relations", where

  - *Generators* are given as a set S of (finitely or infinitely many) elements such that for any $g \in G$ can be written as a finite word $w = s_1 \circ s_2 \circ \cdots \circ s_n$, where either s_i or s_i^{-1} lies in the set S. (Notice, that this n depends on $g \in G$ and is not required to be bounded).

    In other words, G if a minimal group containing all the generators.
  - *Relations:* A word w in the alphabet S, S^{-1} is a relation, if $w = e$ in G.
  - *Defining relations:* is a list of relations $w_1, \ldots, w_n$ such that any relation in w follows from these relations.

  Example 1. $G = \langle r \mid r^2 = e \rangle$ is a group generated by element r satisfying the relation $r^2 = e$. This group contains two elements: e and r (as any longer word in the alphabet r, r^{-1} can be reduced to one of these two.

  Example 2. $G = \langle r_1, r_2 \mid r_1 r_2 = r_2 r_1 = e \rangle$. In this group, every element $g \in G$ can be rewritten as $g = r_1^k r_2^l$, so $G = \mathbb{Z} \oplus \mathbb{Z}$.

Remark. Not every group has a presentation with finitely many generators and finitely many relations. The groups satisfying this property are called *finitely-presented.*

1.3 Discrete groups of isometries acting on $\mathbb{E}^2$

Definition 1.25. A group acts on the set X (denoted $G : X$) if
$\forall g \in G\ \exists f_g$, a bijection $X \to X$, s.t. $f_{gh}(x) = (f_g \circ f_h)(x), \forall x \in X, \forall g, h \in G$.

Example 1.26. Here are some examples of group actions:

(a) Let $G = \langle t_{\boldsymbol{a}} \rangle$ be a group generated by a translation $t_{\boldsymbol{a}}$. Every element of G can be written as $t_{\boldsymbol{a}}^k$ for some $k \in \mathbb{Z}$. Clearly $G : \mathbb{E}^2$ with all elements of G acting as translations $t_{\boldsymbol{a}}^k = t_{k\boldsymbol{a}}$.

(b) $Isom(\mathbb{E}^2)$ acts on the set of all regular pentagons.

(c) $(\mathbb{Z}, +) : \mathbb{E}^2$ in the following way:
Take any vector $\boldsymbol{a}$, then $n \in \mathbb{Z}$ will act on $\mathbb{E}^2$ as the translation $t_{n\boldsymbol{a}}$.

Definition 1.27. An action $G : X$ is transitive if $\forall x_1, x_2 \in X\ \exists g \in G : f_g(x_1) = x_2$.

Example 1.28. (a) The action of $Isom(\mathbb{E}^2)$ on the set of regular pentagons is not transitive (it cannot take a small pentagon to a bigger one).

(b) Theorem 1.10 shows that $Isom(\mathbb{E}^2)$ acts transitively on the set of all triangles congruent to the given one.

(c) $Isom(\mathbb{E}^2)$ acts transitively on points of $\mathbb{E}^2$ (this directly follows from (b)).

(d) The action of $Isom(\mathbb{E}^2)$ on lines is transitive (as for any two lines l_1 and l_2 there is an isometry taking l_1 to l_2.

(e) Theorem 1.10 also implies that $Isom(\mathbb{E}^2)$ acts transitively on flags in $\mathbb{E}^2$, where a flag is a triple (p, r, H^+) such that $p \in \mathbb{E}^2$ is a point, r is a ray from p and H^+ is a half-plane bounded by the line containing r.

(f) $Isom(\mathbb{E}^2)$ does not act transitively on the circles or triangles.

Definition 1.29. Let $G : X$ be an action. An orbit of $x_0 \in X$ under the action $G : X$ is the set $orb(x_0) := \bigcup_{g \in G} gx_0$.

Example 1.30. (a) The group O_2 of isometries preserving the origin O acts on $\mathbb{E}^2$. For this action $orb(O) = O$ (i.e. orbit of the origin is one point) and all other orbits are circles centred at O (see Fig. 9, left).

(b) The group $\mathbb{Z} \times \mathbb{Z}$ acts on $\mathbb{E}^2$ by integer translations (a, b) (where $a \in \mathbb{Z}$ and $b \in \mathbb{Z}$ are the first and the second components respectively). Then the orbit of any point is a shift of the set of all integer points (see Fig. 9, right).

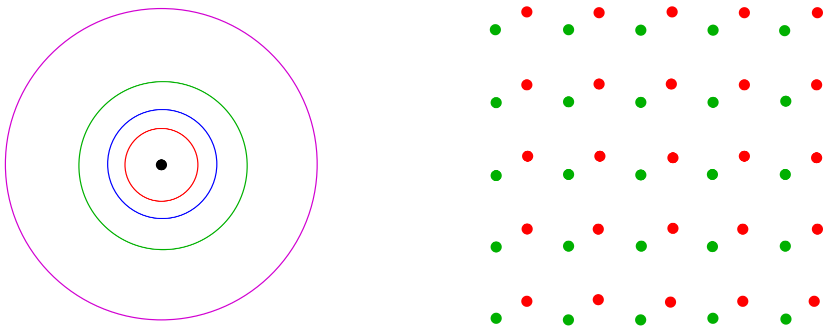


Figure 9: Orbits of O_2 (left) and $\mathbb{Z}^2 = \mathbb{Z} \times \mathbb{Z}$ (right)

Definition 1.31. Let X be a metric space. An action $G : X$ is discrete if none of its orbits possesses accumulation points, i.e. given an orbit $orb(x_0)$, for every $x \in X$ there exists a ball B_x centred at x s.t. the intersection $orb(x_0) \cap B_x$ contains at most finitely many points.

Example 1.32. (a) Consider the action $\mathbb{Z} : \mathbb{E}^1$ defined by $g_n x = 2^n x$ for $n \in \mathbb{Z}$. The action is not discrete as $orb(1) = \{2^n\}$ and the sequence $1/2^n$ converge to $0 \in \mathbb{E}^1$, see Fig. 10, left.

(b) The action $\mathbb{Z} \times Z$ acts on $\mathbb{E}^2$ by translations: let $G = \langle t_1, t_2 \rangle$, where t_1, t_2 are translations in non-collinear directions. This action is discrete as every orbit consists of isolated points, see Fig. 9, right.

(c) (Reflection group). Given an isosceles right-angled triangle, one can generate a group G by reflections in its three sides, $G = \langle r_1, r_2, r_3 \rangle$. Then $G : \mathbb{E}^2$ is a discrete action.

To show that the action is discrete, consider a tiling of $\mathbb{E}^2$ by isosceles right-angled triangles such that any adjacent tiles are reflection images of each other, see Fig. 10, right. Then

- each of the three generators r_1, r_2, r_3 preserves the triangular tiling;
- there are finitely many isometries taking a tile to itself (2 isometries here);
- hence, every tile contains only finitely many points of any given orbit;
- every ball intersects only finitely many tiles;
- which implies that every ball contains finitely many points of each orbit, i.e. the group acts discretely.

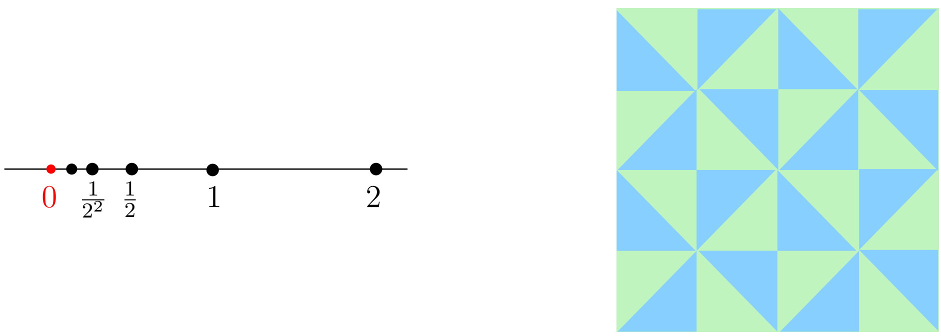


Figure 10: A non-discrete action (left) and a discrete action (right).

Definition 1.33. An open connected set $F \subset X$ is a fundamental domain for an action $G : X$ if the sets $gF, g \in G$ satisfy the following conditions:

1) $X = \bigcup_{g \in G} \overline{gF}$ (where $\overline{U}$ denotes the closure of U in X);

2) $\forall g \in G,\ g \neq e,\ F \cap gF = \emptyset$;

3) There are only finitely many $g \in G$ s.t. $\overline{F} \cap \overline{gF} \neq \emptyset$.

Remark. A set is open if it contains a disc neighbourhood of each point. The closure $\overline{U}$ of U in X is the set of point $\overline{U} = U \cup \{x \in X \mid \forall \varepsilon > 0, B_\varepsilon(x) \cap U \neq \emptyset\}$.

Examples of fundamental domains: any of the triangles in the tiling shown in Fig. 10 is a fundamental domain for the action described in Example 1.32(c).

Definition 1.34. An orbit space X/G for the discrete action $G : X$ is a set of orbits with a distance function

$$d_{X/G} = \min_{\hat{x} \in orb(x),\ \hat{y} \in orb(y)} \{d_x(\hat{x}, \hat{y})\}.$$

Example 1.35. (a) $\mathbb{Z} : \mathbb{E}^1$ acts by translations, then an interval is a fundamental domain. Identifying its endpoints we see that the orbit space $\mathbb{E}^1/Z$ is a circle.

(b) $\mathbb{Z}^2 : \mathbb{E}^2$ (generated by two non-collinear translations), then a parallelogram is a fundamental domain of the action and the orbit space $\mathbb{E}^2/\mathbb{Z}^2$ is a torus.

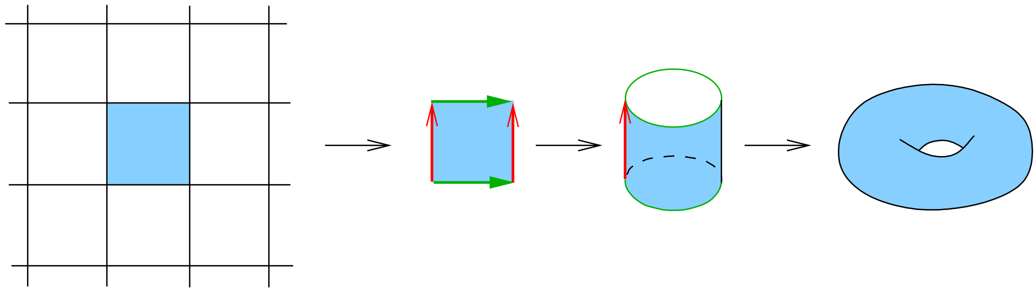


Figure 11: Fundamental domain for $\mathbb{Z}^2 : \mathbb{E}^2$ and a torus as an orbit space.

Remark. One can find some (artistic) tilings of Euclidean plane produced M. C. Escher here, on Escher's official website.

1.4 3-dimensional Euclidean geometry

Model: Cartesian space (x_1, x_2, x_3), $x_i \in \mathbb{R}$, with distance function

$$d(x,y) = (\sum_{i=1}^{3}(x_i - y_i)^2)^{1/2} = \sqrt{\langle x-y, x-y\rangle}.$$

We will not list all the axioms but will mention some essential properties.

Properties:

1. For every plane α there exists a point $A \in \alpha$ and a point $B \notin \alpha$;
2. If two distinct planes α and β have a common point A then they intersect by a line containing A.
3. Given two distinct lines l_1 and l_2 having a common point, there exists a unique plane containing both l_1 and l_2.

Example. Three flies are flying randomly in one room. Find the probability that they are all in one plane at some given moment of time.

Proposition 1.36. *For every triple of non-collinear points there exists a unique plane through these points.*

Proof. Let A, B, C be the three non-collinear points. The lines AB and AC have a common point A. Therefore, there exists a unique plane α containing the lines AB and AC, and hence, containing all three points A, B, C.

□

Definition 1.37. Given a metric space X, a distance between two sets $A, B \in X$ is $d(A,B) := \inf\limits_{a\in A, b\in B}(d(a,b))$.

In particular, the distance between a point A and a plane α is $d(A,\alpha) := \min\limits_{X\in\alpha}(d(A,X))$.

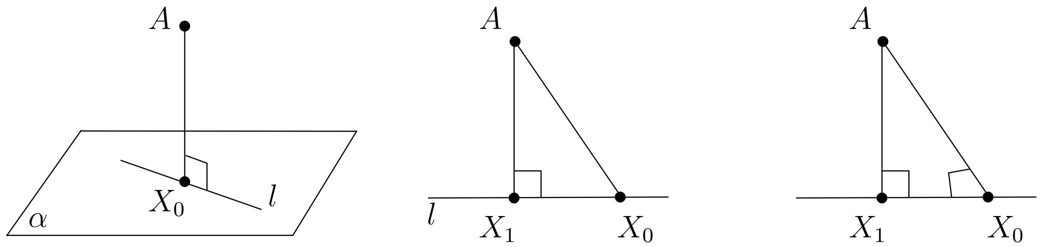


Figure 12: Distance between a point and a plane (see Proposition 1.38).

Proposition 1.38. *Given a plane α, a point $A \notin \alpha$ and a point $X_0 \in \alpha$, $AX_0 = d(A,\alpha)$ if and only if $AX_0 \perp l$ for every $l \in \alpha$, $X_0 \in l$.*

Proof. "$\Rightarrow$": First, we prove that $AX_0 = d(A, \alpha)$ implies that $AX_0 \perp l$ for every $l \in \alpha$, $X_0 \in l$. Suppose that $l \in \alpha$, $X_0 \in l$ and l is not orthogonal to AX_0, see Fig. 12, in the middle. Then there exists $X_1 \in l$ such that $d(X_1, A) < d(X_0, A)$ (indeed, this is the case when X_1 is the point such that $AX_1 \perp l$).

"$\Leftarrow$": Suppose that $AX_0 \perp l$, but $d(A, X_0) \neq d(A, \alpha) = d(A, X_1)$, see see Fig. 12, right. As it is shown above, $AX_1 \perp X_1X_0$. Then there are two distinct lines through A perpendicular to l, in contradiction with E9.

$\square$

Corollary. Given a plane α and a point $A \notin \alpha$, the closest to A point $X_0 \in \alpha$ is unique.

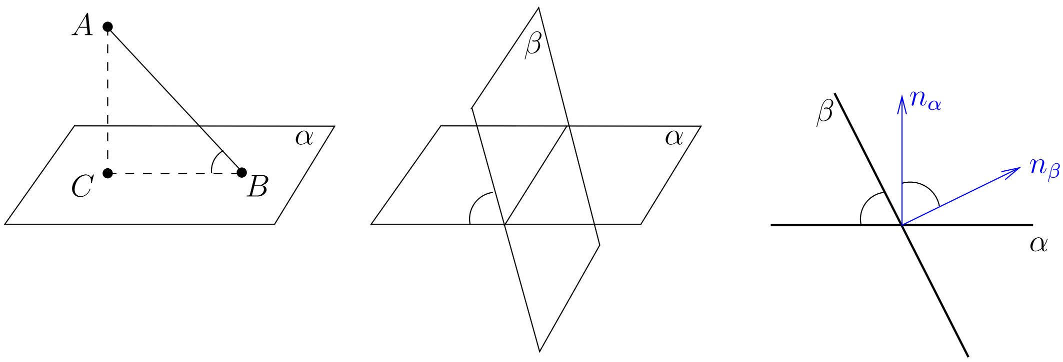


Figure 13: Angle between a line and a plane (left) and between two planes (right).

Definition 1.39. (a) The point $X_0 \in \alpha$ s.t. $d(A, \alpha) = AX_0$ is called an <u>orthogonal projection</u> of A to α. Notation: $X_0 = proj_\alpha(A)$.

(b) Let α be a plane, AB be a line, $B \in \alpha$, and $C = proj_\alpha(A)$. The <u>angle</u> between the line AB and the plane α is $\angle(AB, \alpha) = \angle ABC$, where $C = proj_\alpha(A)$, (see Fig. 13, left).
<u>Equivalently</u>, $\angle(AB, \alpha) = \min\limits_{X \in \alpha}(\angle ABX)$.

Exercise: Check the equivalence. *Hint:* use cosine rule.

Remark. Definition 1.37 implies that if $AC \perp \alpha$ then $AC \perp l$ for all $l \in \alpha$, $C \in l$.

Definition 1.40. The <u>angle</u> $\angle(\alpha, \beta)$ between two intersecting planes α and β is the angle between their normals (see Fig. 13 middle and right).

<u>Equivalently</u>, if $B \in \beta$, $A = proj_\alpha(B)$, $C = proj_l(A)$ where $l = \alpha \cap \beta$,
then $\angle(\alpha, \beta) = \angle BCA$.

Exercise:

1. Check the equivalence.
2. Let γ be a plane through BCA. Check that $\gamma \perp \alpha$, $\gamma \perp \beta$.
3. Let α be a plane, $C \in \alpha$. Let B be a point s.t. $BC \perp \alpha$. Let β be a plane through C, $\beta \perp \alpha$. Then $B \in \beta$.

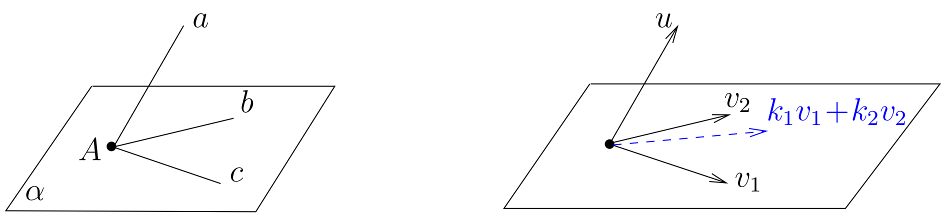


Figure 14: To Proposition 1.41.

Proposition 1.41. *Given two intersecting lines b and c in a plane α, $A = b \cap c$, and a line a, $A \in a$, if $a \perp b$ and $a \perp c$ then $a \perp \alpha$ (i.e. $a \perp l$ for every $l \in \alpha$).*

Proof. Given three vectors $\boldsymbol{u}, \boldsymbol{v}_1, \boldsymbol{v}_2$ in $\mathbb{R}^3$ such that $(\boldsymbol{u}, \boldsymbol{v}_1) = 0$ and $(\boldsymbol{u}, \boldsymbol{v}_2) = 0$ we have $(\boldsymbol{u}, k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_2) = 0$ for any $k_1, k_2 \in \mathbb{R}$.

$\square$

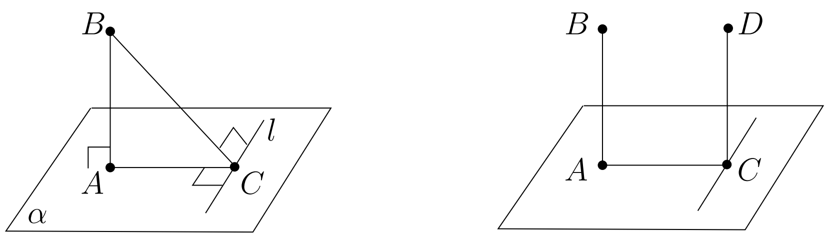


Figure 15: Theorem of three perpendiculars (and it's proof).

Theorem 1.42 (Theorem of three perpendiculars)**.** *Let α be a plane, $l \in \alpha$ be a line and $B \notin \alpha$, $A \in \alpha$ and $C \in l$ be three points. If $BA \perp \alpha$ and $AC \perp l$ then $BC \perp l$.*

Proof. 1. Let CD be a line through C parallel to AB, see Fig. 15. Then $CD \perp \alpha$ (as $AB \perp \alpha$).
2. Then $CD \perp l$ (as $CD \perp l' \quad \forall l' \subset \alpha$. Also, $l \perp AC$ (by assumption).
3. Hence, by Proposition 1.41 $l \perp$ (plane ACD), i.e. $l \perp BC$ (as $BC \subset$ plane ACD).

$\square$

1.5 References

- A nice discussion of the group of isometries of Euclidean plane can be found in G. Jones, *Algebra and Geometry*, Lecture notes (Section 1).
  (The notes are available on ULTRA, see "Other Resources" section).

- Discussion of the geometric constructions and constructibility of various geometric objects can be found in
  G. Jones, *Algebra and Geometry*, Lecture notes (Section 8).
  (The notes are available on ULTRA, see "Other Resources" section).

- More detailed discussion of Euclidean isometries can be found here:
  N. Peyerimhoff, *Geometry III/IV*, Lecture notes (Section 1).

- To read more about the role of reflections for $Isom(E^2)$, look at
  O. Viro, *Defining relations for reflections I*, arXiv:1405.1460v1.

- The following book (Section 1) provides an introduction to group actions:
  T. K. Carne, *Geometry and groups.*
  Also, one can find here a detailed discussion of the group of Euclidean isometries (Sections 2-4) - as well as many other topics.

- Another source concerning groups actions:
  A. B Sossinsky, *Geometries*, Providence, RI : American Mathematical Soc. 2012.
  One can find the book in the library, see also Section 1.3 (pp.9–11) here.
  Section 2.7 (pp.26-27) of the same source introduces group presentations and gives many examples.

- Exposition of 3-dimensional Euclidean Geometry can be found in Chapter 1 of Kiselev's Geometry, *Book II. Stereometry.* (Adopted from Russian by Alexader Givental).
  (The book is not easily reachable at the moment. You can find a reference to Amazon on Giventhal's homepage. I should probably order the book for our library... Please, tell me if you are interested in this book).

- Webpages, etc:

  - Cut-the-knot portal by Alexander Bogomolny.
  - Drawing a Circle with a Framing Square and 2 Nails.
  - One can find some (artistic) tilings of Euclidean plane produced M. C. Escher here, on Escher's official website.

2 Spherical geometry

In this section we will study geometry on the surface of the sphere.

Model of the sphere S^2 in $\mathbb{R}^3$: (sphere of radius $R = 1$ centred at $O = (0,0,0)$)

$$S^2 = \{\ (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1^2 + x_2^2 + x_3^2 = 1\}$$

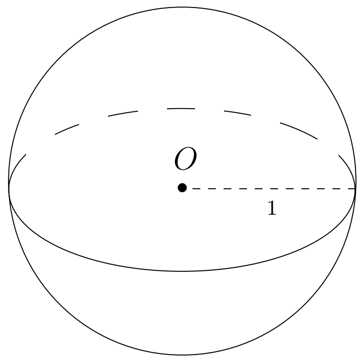


Figure 16: Sphere.

Sometimes we will consider sphere of radius R: $\{\ (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1^2 + x_2^2 + x_3^2 = R\}$.

2.1 Metric on S^2

Definition 2.1.
- Points A and A' of S^2 will be called antipodal if $O \in AA'$.
- A great circle on S^2 is the intersection of S^2 with a plane passing though O, see Fig. 17, left.

Remark 2.2. Given two distinct non-antipodal points $A, B \in S^2$, there exists a unique great circle through A and B (as there is a unique 2-dimensional plane through 3 non-collinear points A, B, O).

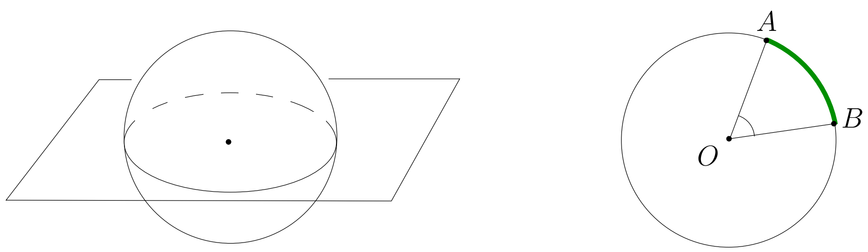


Figure 17: Great circles and distance on the sphere.

Definition 2.3. Given a sphere S^2 of radius R, a distance $d(A, B)$ between the points $A, B \in S^2$ is πR, if A is diametrically opposed to B, and the length of the shorter arc of the great circle through A and B, otherwise.

Equivalently, $d(A, B) := \angle AOB \cdot R$ (with $R = 1$ for the case of unit sphere).
See Fig. 17, right.

Theorem 2.4. *The distance $d(A, B)$ turns S^2 into a metric space, i.e. the following three properties hold:*
M1. $d(A, B) \geq 0 \quad (d(A, B) = 0 \Leftrightarrow A = B)$;
M2. $d(A, B) = d(B, A)$;
M3. $d(A, C) \leq d(A, B) + d(B, C)$ *(triangle inequality).*

Proof. M1 and M2 hold by definition. To prove $M3$ we need to show

$$\angle AOC \leq \angle AOB + \angle BOC.$$

We will do it in the following 8 steps.

1. If B lies on a great circle $\mathcal{C}_{AC}$ through A and C, then M3 holds (may turn into equality). Assume $B \notin \mathcal{C}_{AC}$.

2. Suppose that $\angle AOC > \angle AOB + \angle BOC$, in particular, $\angle AOC > \angle AOB$.

3. Choose B_1 inside AC so that $\angle AOB_1 = \angle AOB$, see Fig. 18.
   Choose $B_2 \in OB$ so that $OB_2 = OB_1$.
   Then $AB_1 = AB_2$ (since $\triangle AB_1O$ is congruent to $\triangle AB_2O$ by SAS).

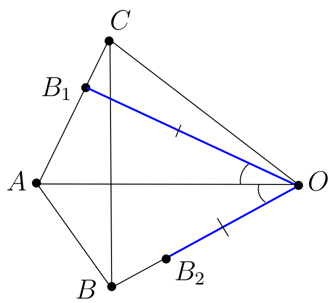


Figure 18: To the proof of triangle inequality for S^2.

4. Since $\angle AOC > \angle AOB + \angle BOC$ we have $\angle AOC > \angle AOB_2 + \angle B_2OC$.
   Also, $\angle AOC = \angle AOB_1 + \angle B_1OC$.
   Hence, $\angle B_2OC < \angle B_1OC$.

5. Recall the Cosine Rule in $\mathbb{E}^2$: $c^2 = a^2 + b^2 - 2ab\cos\gamma$.
   Note that given the sides a, b, for a larger angle γ between them we get a larger side c.

6. Applying results of steps 4 and 5 to $\triangle OB_1C$ and OB_2C, we get $B_2C < B_1C$.

7. $AB_2 + B_2C \overset{3,6}{<} AB_1 + B_1C = AC \leq AB_2 + B_2C$
   (here the last inequality is the triangle inequality on the plane).

8. The contradiction obtained in 7 shows that $\angle AOC \le \angle AOB + \angle BOC$ (where equality only holds when B lies in the plane ACO).

$\square$

2.2 Geodesics on S^2

Definition 2.5. A curve γ in a metric space X is a geodesic if γ is locally the shortest path between its points.

More precisely, $\gamma(t) : (0,1) \to X$ is geodesic if

$$\forall t_0 \in (0,1) \quad \exists \varepsilon : \quad l(\gamma(t)|_{t_0-\varepsilon}^{t_0+\varepsilon}) = d(\gamma(t_0-\varepsilon), \gamma(t_0+\varepsilon)).$$

Proposition 2.6. *Geodesics on S^2 are great circles.*

Proof. Use the (spherical) triangle inequality and repeat the proof of Proposition 1.24. $\square$

Definition 2.7. Given a metric space X, a geodesic $\gamma : (-\infty, \infty) \to X$ is called closed if $\exists T \in \mathbb{R}, T \neq 0 : \ \gamma(t) = \gamma(t+T) \ \ \forall t \in (-\infty, \infty)$, and open, otherwise.

Example. In $\mathbb{E}^2$, all geodesics are open, each segment is the shortest path.
In S^2, all geodesics are closed, one of the two segments of $\gamma \setminus \{A, B\}$ is the shortest path (another one is not shortest if A and B are not antipodal).
HW 4.1: describes a metric space containing both closed and open geodesics.

From now on: by lines in S^2 we mean great circles.

Proposition 2.8. *Every line on S^2 intersects every other line in exactly two antipodal points.*

Proof. Let $l_1 = \alpha_1 \cap S$ and $l_2 = \alpha_2 \cap S$ be two lines on S^2, see Fig. 19, left. Then

$$l_1 \cap l_2 = (\alpha_1 \cap \alpha_2) \cap S^2 = (\text{ line through origin }) \cap S^2,$$

as $O \in \alpha_1 \cap \alpha_2$.

$\square$

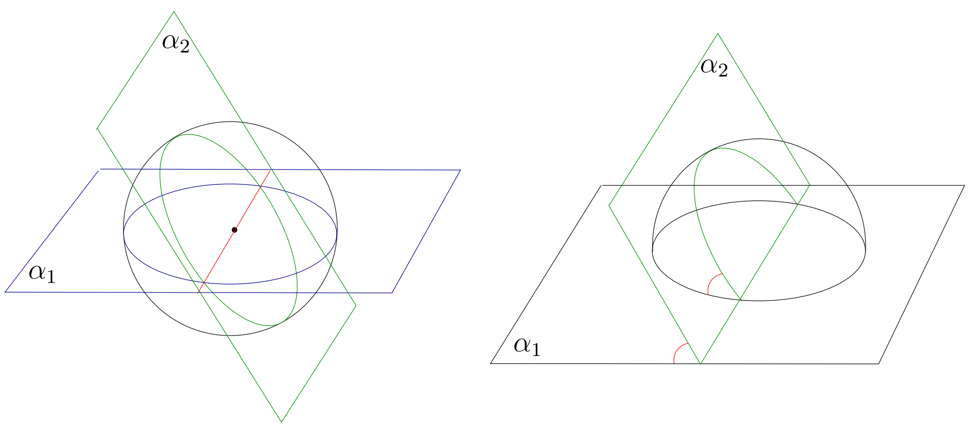


Figure 19: Intersection and angle between two lines on the sphere.

Definition 2.9. By the angle between two lines we mean the angle between the corresponding planes:
if $l_i = \alpha_i \cap S^2$, $i = 1, 2$ then $\angle(l_1, l_2) := \angle(\alpha_1, \alpha_2)$, see Fig. 19, right.

Equivalently, $\angle(l_1, l_2)$ is the angle between the lines $\hat{l}_1$ and $\hat{l}_2$, $\hat{l}_i \in \mathbb{R}^3$,
where $\hat{l}_i$ is tangent to the great circle l_i at $l_1 \cap l_2$ as to a circle in $\mathbb{R}^3$.

Proposition 2.10. *For every line l and a point $A \in l$ in this line there exists a unique line l' orthogonal to l and passing through A.*

Proof. Consider the plane $\alpha \in \mathbb{R}^3$ such that $l = \alpha \cap S^2$. We need to find another line $l' = \beta \cap S^2$, where $\beta \in \mathbb{R}^3$ is a plane orthogonal to α and such that $O, A \in \beta$. Let v_α be the normal vector at O to α, see Fig. 21, left. Since $\beta \perp \alpha$, we see that $v_\alpha \in \beta$. So, β is the plane spanned by the line OA and v_α. This construction shows both existence of l' and uniqueness.

□

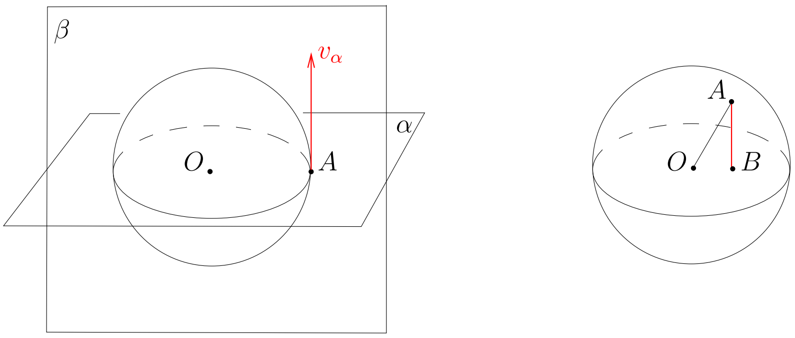


Figure 20: Existence and uniqueness of a perpendicular line on the sphere.

Proposition 2.11. *For every line l and a point $A \notin l$ in this line, s.t. $d(A, l) \neq \pi/2$ there exists a unique line l' orthogonal to l and passing through A.*

Proof. Let $B \in \alpha$ be the orthogonal projection of A to the plane α, see Fig. 21, right. Then $l' = \beta \cap S^2$, where $\beta = OAB$.

□

Notice that given the points A, B in the line l, one of the two segments $l \setminus \{A, B\}$ is the shortest path between them.

Definition 2.12. A triangle on S^2 is a union of three non-collinear points and a triple of the shortest paths between them.

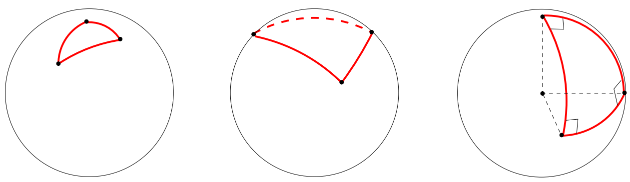


Figure 21: Spherical triangles

2.3 Polar correspondence

Definition 2.13. Let $l = S^2 \cap \Pi_l$ be a line on S^2, where Π_l is the corresponding plane through O in $\mathbb{R}^3$. The pole to the line l is the pair of endpoints of the diameter DD' orthogonal to Π_l, i.e. $Pol(l) = \{D, D'\}$.
A polar to a pair of antipodal points D, D' is the great circle $l = S^2 \cap \Pi_l$, s.t. the plane Π_l is orthogonal to DD', i.e. $Pol(D) = Pol(D') = l$.

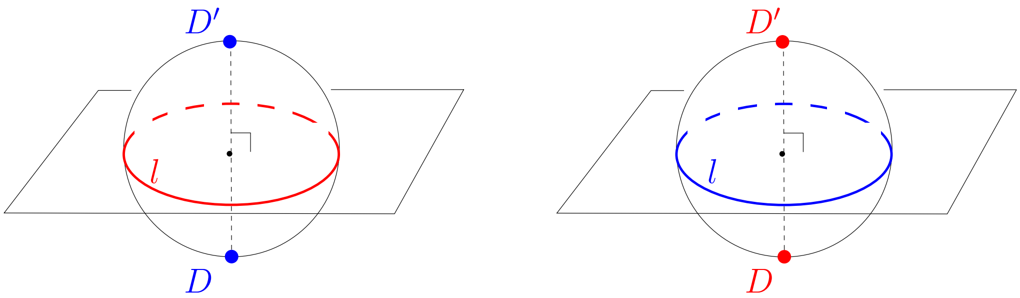


Figure 22: Polarity: $Pol(l) = \{D, D'\}$ (left) and $Pol(D) = Pol(D') = l$ (right).

Proposition 2.14. *If a line l contains a point A then the line $Pol(A)$ contains both points of $Pol(l)$.*

Proof. 1. Let $\{D, D'\} := Pol(l)$, i.e. $DD' \perp \alpha_l$, where $l = \alpha_l \cap S^2$. In particular, $OD \perp OA$ (see Fig. 23, left).

2. By definition, $Pol(A)$ is the line $l' = S^2 \cap \alpha_A$, where $\alpha_A \perp OA$.

3. We conclude that $OD \subset \alpha_A$ as $AD \perp OA$. Hence, $D \subset Pol(A)$. Similarly, $D' \subset Pol(A)$.

□

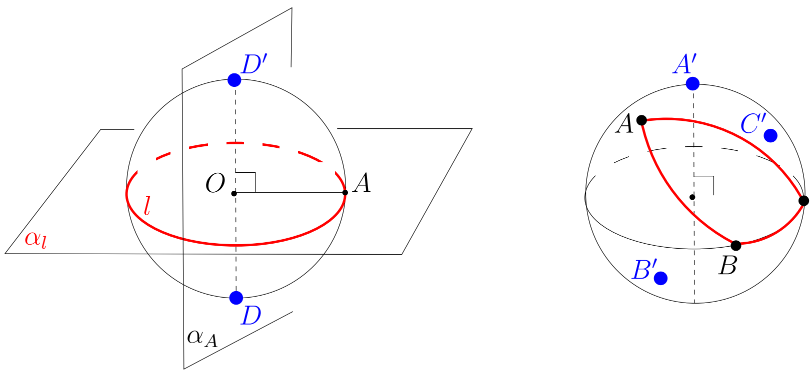


Figure 23: Left: $A \in l \;\Rightarrow\; Pol(l) \in Pol(A)$. Right: polar triangle.

Hence, polar correspondence transforms:

- points into lines;
- lines into points;
- the statement "A line l contains a point A" into
  "The points $Pol(l)$ lie on the line $Pol(A)$".

Definition 2.15. A triangle $A'B'C'$ is polar to ABC (denoted $A'B'C' = Pol(ABC)$) if $A' \in Pol(BC)$ and $\angle AOA' \le \pi/2$, and similar conditions hold for B' and C', see Fig. 23, right.

Remark. If $A' \in Pol(BC)$, then to say " $\angle AOA' \le \pi/2$" is the same as to say that A' lies on the same side with respect to BC as A.

Exercise. Is there a self-polar triangle ABC on S^2, i.e. a triangle ABC such that $Pol(ABC) = ABC$?

Theorem 2.16 (Bipolar Theorem)**.**

(a) If $A'B'C' = Pol(ABC)$ then $ABC = Pol(A'B'C')$.

(b) If $A'B'C' = Pol(ABC)$ and $\triangle ABC$ has angles α, β, γ and side lengths a, b, c, then $\triangle A'B'C'$ has angles $\pi - a, \pi - b, \pi - c$ and side lengths $\pi - \alpha, \pi - \beta, \pi - \gamma$.

Proof. (a) Since $A' \in Pol(BC)$, we have $OA' \perp OC, OB$. Since $B' \in Pol(AC)$, we have $OB' \perp OC, OA$. From this we conclude that $OC \perp OA', OB'$, i.e. OC is orthogonal to the plane $OA'B'$, which implies that $C \in Pol(A'B')$. Also, we have $\angle COC' < \pi/2$.

As similar conditions hold for A and B, we conclude that $ABC = Pol(A'B'C')$.

(b)
- Angle $\beta = \angle ABC$ between the spherical lines AB and BC is equal to the angle between corresponding planes α_{AB} and α_{BC} in $\mathbb{E}^3$.
- The length b' in the spherical triangle $A'B'C'$ is given by definition by $b' = \angle A'OC'$.
- As $OA' \perp \alpha_{BC}$, $OC' \perp \alpha_{AB}$, we see $\angle A'OC' = \pi - \beta$, see Fig. 24. So, we get $b' = \pi - \beta$.
- By symmetry, we get all other equations.

□

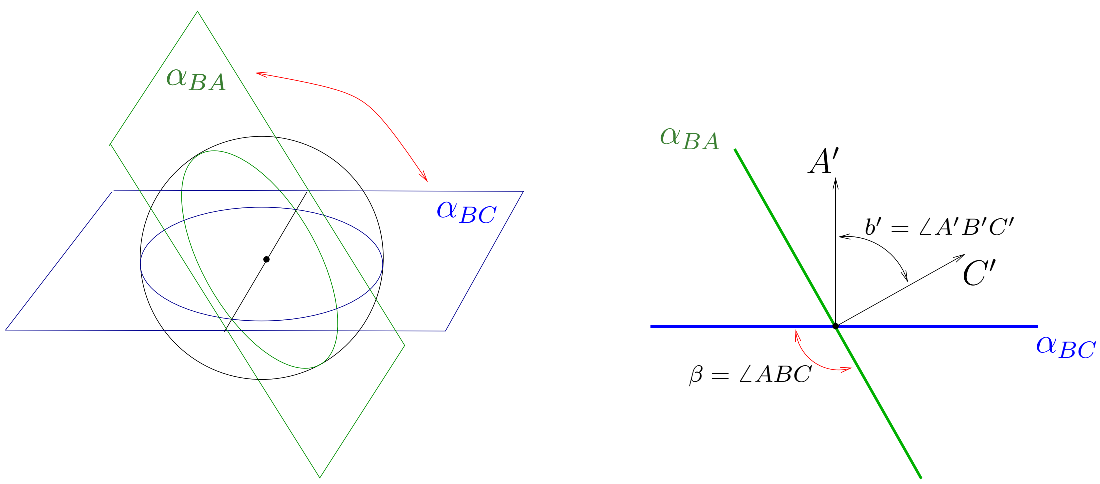


Figure 24: Proof of Bipolar Theorem.

2.4 Congruence of spherical triangles

Theorem 2.17. *SAS, ASA, and SSS hold for spherical triangles.*

Proof. The proofs are exactly the same as for similar statements in $\mathbb{E}^2$.

SAS: This is an axiom (of congruence of trihedral angles in $\mathbb{E}^3$).

ASA: 1. Suppose that $\angle BAC = \angle B'A'C'$, $AC = A'C'$, $\angle BCA = \angle B'C'A'$.
2. If $AB = A'B'$, then $\triangle ABC \cong \triangle A'B'C'$ by SAS.
3. If $AB \neq A'B'$, consider $B'' \subset A'B'$ such that $AB = AB''$.
4. Then $\triangle A'B''C' \cong \triangle ABC$ by SAS, which implies that $\angle BCA = \angle B''C'A'$. This means that the lines CB' and CB'' coincide, and hence $B = B'$ (as a unique intersection of two rays in the given half-space with respect to $A'C'$).

SSS: Assume that the corresponding sides of $\triangle ABC$ and $\triangle A'B'C'$ are equal but the triangles are not congruent, see Fig. 25. Consider a triangle ABC'' congruent to $A'B'C'$. Notice that $C'' \neq C$, but $AC = AC''$ and $BC = BC''$, which implies that the segment CC'' has two distinct perpendicular bisectors (one constructed as the altitude in the isosceles triangle ACC'', and another as an altitude in isosceles triangle BCC'', see Remark 2.18 below). This contradicts to Proposition 2.10.

□

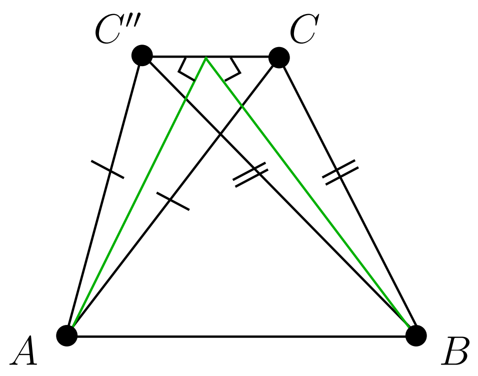


Figure 25: Proof of SSS.

Notice that as soon as we have SAS property, we can immediately deduce the following corollary:

Corollary 2.18. *(a) In a triangle ABC, if $AB = BC$ then $\angle BAC = \angle BCA$.*

(b) If $AB = BC$ and M is a midpoint of AC then $BM \perp AC$.

Proof. (a) Follows as $\triangle ABC \cong \triangle CBA$ by SAS.
Then (b) follows as $\triangle BAM \cong \triangle BCM$ by SAS in view of (a).

□

In Euclidean plane, triangles with three equal angles are not necessarily congruent, but only similar. This is not the case in S^2:

Theorem 2.19. *AAA holds for spherical triangles.*

Proof. Consider the polar triangles $Pol(ABC)$ and $Pol(A'B'C')$. By Bipolar Theorem (Theorem 2.16(b)) AAA for initial triangles turns into SSS for the polar triangles. Hence, $Pol(ABC)$ is congruent to $Pol(A'B'C')$. Applying Theorem 2.16 again, we conclude that ABC is congruent to $A'B'C'$.

□

2.5 Sine and cosine rules for the sphere

a. Sine and cosine rules on the plane

Before discussing spherical sine and cosine rules, lets recall the statements for Euclidean plane:

Consider a triangle on $\mathbb{E}^2$ with sides a, b, c and opposite angles α, β, γ, as in Fig. 26, left. Then:

$$\boxed{\begin{array}{rl} \text{sine rule:} & \dfrac{a}{\sin\alpha} = \dfrac{b}{\sin\beta} = \dfrac{c}{\sin\gamma} \\[2ex] \text{cosine rule:} & c^2 = a^2 + b^2 - 2ab\cos\gamma \end{array}}$$

Proof. Sine rule: Let A, B, C be the vertices of the triangle with the angles α, β, γ respectively. Drop the perpendicular BH from B to AC, see Fig. 26, right. Then $BH = c\sin\alpha = a\sin\gamma$, which implies $\frac{c}{\sin\gamma} = \frac{a}{\sin\alpha}$. The other equality is obtained by symmetry.

Cosine rule: With the same H as before, we have $BH = a\sin\gamma$, $CH = a\cos\gamma$, then

$$\begin{aligned} c^2 &= AH^2 + BH^2 = (b - CH)^2 + BH^2 \\ &\qquad\qquad = (b^2 - 2b\cdot a\cos\gamma + a^2\cos^2\gamma) + a^2\sin^2\gamma = a^2 + b^2 - 2ab\cos\gamma. \end{aligned}$$

□

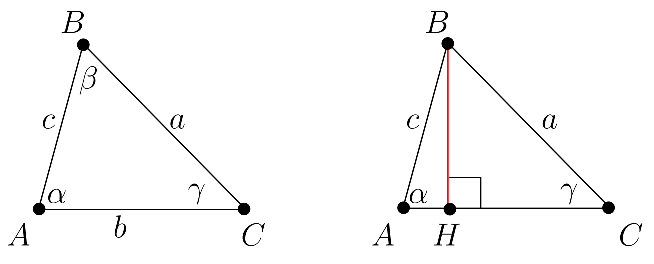


Figure 26: Triangle $\triangle ABC$.

b. Sine and cosine rules on the unit sphere

Theorem 2.20 (Sine rule for the unit S^2)**.** $\frac{\sin a}{\sin \alpha} = \frac{\sin b}{\sin \beta} = \frac{\sin c}{\sin \gamma}$.

Proof. - Let H be the orthogonal projection of A to the plane OBC.

- Let A_b and A_c be orthogonal projections of H to the lines OB, OC respectively, see Fig. 27, left.

- As $AH \perp OHC$ and $HA_c \perp OC$,
  Theorem of three perpendiculars (Theorem 1.42) implies that $AA_c \perp OC$.

- As $OC \perp A_cH$ and $OC \perp A_cA$,
  we see that $\angle AA_cH = \angle(OHC, OA_cA) = \angle(OBC, OAC) = \gamma$
  see Fig. 27, right.

- $AH \overset{\triangle AHA_c}{=} AA_c \sin\gamma \overset{\triangle AOA_c}{=} AO\sin(\pi - b)\sin\gamma = R\sin b \sin\gamma$.

- Similarly, $AH = AA_b \sin\beta = \cdots = R\sin c\sin\beta$.

- We conclude that $\frac{\sin b}{\sin\beta} = \frac{\sin c}{\sin\gamma}$.

□

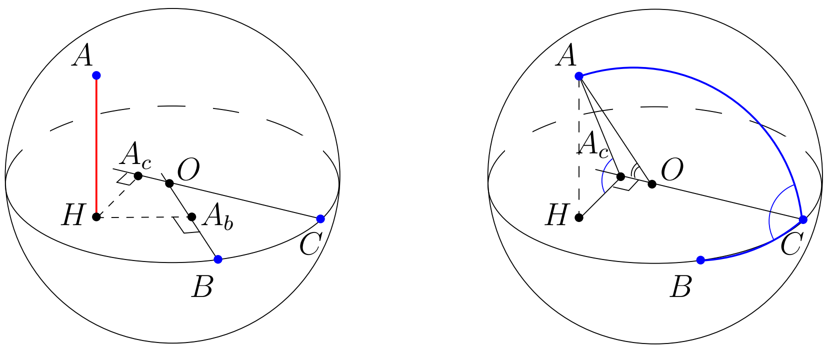


Figure 27: Proof of the sine rule on the sphere.

Remark. If a, b, c are small then $a \approx \sin a$ and the spherical sine rule transforms into Euclidean one.

Corollary. (Thales Theorem) If $a = b$ then $\angle\alpha = \angle\beta$, i.e. the base angles in isosceles triangles are equal.

Theorem 2.21 (Cosine rule for $\mathbb{S}^2$)**.** $\cos c = \cos a \cos b + \sin a \sin b \cos\gamma$.

Proof. We skip the proof in the class, but one can find it in any of the following:
- Prasolov, Tikhomirov: Section 5.1, p.87;
- Prasolov: p.48.

□

Remark. If a, b, c are small then $\cos a \approx 1 - a^2/2$ and the spherical cosine rule transforms into Euclidean one.

Theorem 2.22 (Second cosine rule). $\cos\gamma = -\cos\alpha\cos\beta + \sin\alpha\sin\beta\cos c$.

Proof. Let $A'B'C' = Pol(ABC)$ be the triangle polar to ABC. Then by Bipolar Theorem (Theorem 2.16) $a' = \pi - \alpha$, $\cos a' = -\cos\alpha$, $\sin a' = \sin\alpha$. Applying the first cosine rule (Theorem 2.21) to $\triangle A'B'C'$ we get

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

which implies

$$-\cos\gamma = \cos\alpha\cos\beta - \sin\alpha\sin\beta\cos c.$$

□

Remark.

(a) If a, b, c are small then $\cos a \approx 1$ and from the second cosine rule we have $\cos\gamma = -\cos\alpha\cos\beta + \sin\alpha\sin\beta = \cos(\alpha+\beta)$, which means that $\gamma = \pi - (\alpha+\beta)$. So, the second cosine rule transforms into $\alpha + \beta + \gamma = \pi$.

(b) For a right-angled triangle with $\gamma = \pi/2$ we have $\sin\gamma = 1$, $\cos\gamma = 0$. So we obtain:
sine rule: $\sin b = \sin c \cdot \sin\beta$,
cosine rule: $\cos c = \cos a \cos b$ (Spherical Pythagorean Theorem).

(c) Is there a "second sine rule" on the sphere?
Writing the sine rule for the polar triangle only changes the places of numerators and denominators in the sine rule and does not lead to anything new...

2.6 More about triangles

The following properties of spherical triangles are exactly the same as the corresponding properties of Euclidean triangles:

Proposition 2.23. *For any spherical triangle,*

1: angle bisectors are concurrent;

2,3,4: perpendicular bisectors, medians, altitudes are concurrent.

5,6: There exist a unique inscribed and a unique circumscribed circles for the triangle.

Proof.

- Parts 1,2 are discussed in HW 5.2 (and can be done as for $\mathbb{E}^2$).
- Parts 3,4 are discussed in HW 6.5 (here, one needs to use some projections to reduce the statement to similar statements on $\mathbb{E}^2$.
- Parts 5,6 follow directly from 1,2 respectively (as on $\mathbb{E}^2$, one needs to think about an angle bisector as a locus of points on the same distance from the sides of the angle and a perpendicular bisector as a locus of points on the same distance from the endpoints of the segment).

□

Remark. To define an altitude AH in a triangle $\triangle ABC$, we need to assume that at least one of angles $\angle B$ and $\angle C$ in $\triangle ABC$ is not a right angle.

So, There are many common properties for triangles in S^2 and $\mathbb{E}^2$, however, not everything about spherical triangles works exactly the same way as in Euclidean plane:

Example 2.24. *Let M, N be the midpoints of AB and AC in a spherical triangle ABC. Then $MN > AC/2$.*
One can use cosine law to prove the statement, see HW 6.6.

Moreover, for *some* triangles in the sphere one can even have $MN > AC$, or even $MN > 100AC$!
To see this take B to be the North Pole, and A and C to be the points on the same parallel very close to the South Pole.

2.7 Area of a spherical triangle

We will denote area of X by $S(X)$ or by S_X and will assume the following properties of the area:

- $S(X_1 \sqcup X_2) = S(X_1) + S(X_2)$
  where $\sqcup$ means a disjoint union, i.e. interior of X_1 is disjoint from interior of X_2.
- If f is an isometry of S^2 then $S(X) = S(f(X))$ for any domain $X \in S^2$.
- $S(S^2) = 4\pi R^2$ for a sphere of radius R.

Theorem 2.25. *The area of a spherical triangle with angles α, β, γ equals*

$$(\alpha + \beta + \gamma - \pi)R^2,$$

where R is the radius of the sphere.

Proof.
1. Consider a spherical digon, i.e. one of 4 figures obtained when $\mathbb{S}^2$ is cut along two lines. See Fig. 28, left. Let $S(\alpha)$ be the area of the digon of angle α.
2. $S(\alpha)$ is proportional to α. Indeed we can divide the whole sphere into $2n$ congruent digons, and obtain that $S(\pi/n) = 4\pi R^2/2n$. This will show the proportionality for π-rational angles. For others we will apply continuity of the area. As $S(2\pi) = S(\mathit{sphere}) = 4\pi R^2$, we conclude that $S(\alpha) = 2\alpha R^2$.
3. - The pair of lines AB and AC meeting at angle α determines two α digons.
   - Similarly, AB and BC gives two β-digons and AC, CB gives two γ-digons, see Fig. 28, middle.
   - The total area of all six digons is $S_{digons} = 2R^2(2\alpha + 2\beta + 2\gamma)$.
   - Triangle ABC is covered by three digons, also triangle $A'B'C'$ antipodal to ABC is covered by 3 digons.
   - All other parts of $\mathbb{S}^2$ are covered only by one digon each, see Fig. 28, right.
   - So,
     $$3(S_{ABC} + S_{A'B'C'}) + S_{\mathbb{S}^2 \setminus \{\triangle ABC \cup \triangle A'B'C'\}} = S_{digons}.$$
     Hence, $2(S_{ABC} + S_{A'B'C'}) + S_{\mathbb{S}^2} = S_{digons}$. Which implies
     $$4S_{ABC} + 4\pi R^2 = 2R^2(2\alpha + 2\beta + 2\gamma)$$
     and we get $S_{ABC} = R^2(\alpha + \beta + \gamma - \pi)$.

□

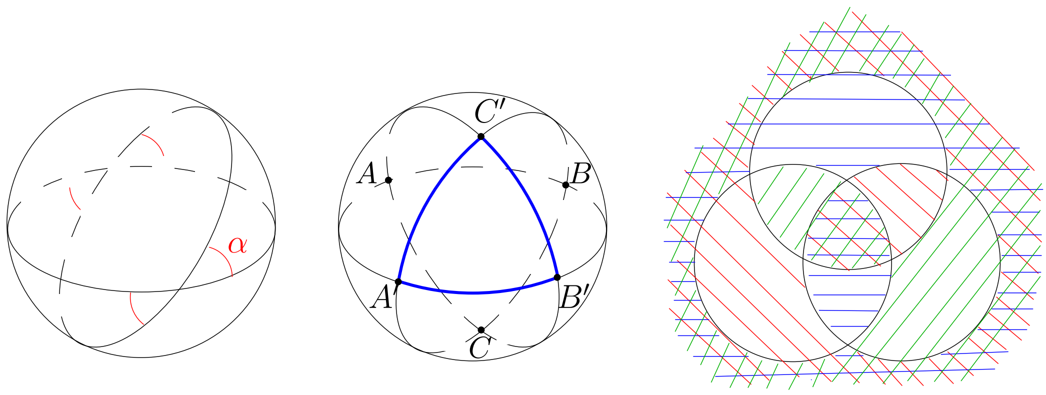


Figure 28: Computing the area of a triangle using digons

Corollary 2.26. $\pi < \alpha + \beta + \gamma < 3\pi$.

Proof. The area of triangle is positive. Also, every angle is smaller than π. □

Corollary 2.27. $0 < a + b + c < 2\pi$.

Proof. Let $A'B'C' = Pol(ABC)$. Then $\alpha' + \beta' + \gamma' > \pi$, and by Bipolar Theorem (Theorem 2.16) we have $(\pi - a) + (\pi - b) + (\pi - c) > \pi$, which implies $a + b + c < 2\pi$. □

Theorem 2.28. *No domain on S^2 is isometric to a domain on $\mathbb{E}^2$.*

Proof. One proof directly follows from sine or cosine rule, another from the sum of angles of a triangle.

The third proof is by comparing the length of circles of radius r: a spherical circle of radius r has length $2\pi \sin r$ while in $\mathbb{E}^2$ such a circle would have length $2\pi r$, see Fig. 29 (we leave the computation as an excercise). □

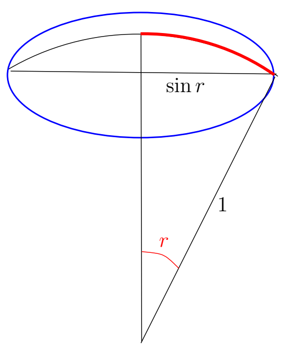


Figure 29: Computing length of spherical circle

2.8 Isometries of the sphere

Example 2.29. The following maps are isometries of $\mathbb{S}^2$ (as they are restrictions to S^2 of isometries in $\mathbb{E}^3$):

- Rotation about a point A on the sphere may be understood as a restriction of rotation of $\mathbb{E}^3$ about the corresponding diameter of the sphere.
- Reflection with respect a line l on $\mathbb{S}^2$ may be understood as a restriction of reflection in $\mathbb{E}^3$ with respect to the plane α s.t. $l = \alpha \cap \mathbb{S}^2$.
- Antipodal map is a restriction of the symmetry in $\mathbb{E}^3$ with respect to the point O.

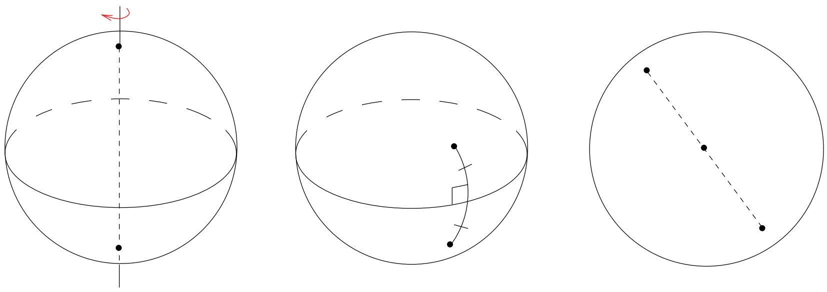


Figure 30: Examples of isometries on $\mathbb{S}^2$: rotation, reflection and antipodal map.

Proposition 2.30. *Every non-trivial isometry of S^2 preserving two non-antipodal points A, B is a reflection (with respect to the line AB).*

Proof.
- Suppose $f \in Isom(\mathbb{S}^2)$, such that $f(A) = A$, $f(B) = B$, $f(X) = X' \neq X$.
- Since f is an isometry, we see that $\triangle ABX$ is congruent to $\triangle ABX'$ (by SSS), see Fig. 31, left. Hence, $\angle ABX = \angle ABX'$.
- Since $X \neq X'$, this implies that X and X' lie in different hemispheres with respect to AB.
- Consider the point $H \in AB$ such that $\angle XHB = \pi/2$. Then $\triangle HCB \cong \triangle HX'B$ by SAS. This implies that $X' = r_{AB}(X)$ is a reflection image of X.

□

Proposition 2.31. *Given points A, B, C, satisfying $AB = AC$, there exists a reflection r such that $r(A) = A$, $r(B) = C$, $r(C) = B$.*

Proof. Let M be the midpoint of BC, let $r = r_{AM}$ be the reflection with respect to AM, see Fig. 31, right. Then $\triangle AMB \cong \triangle AMC$ by SSS, which implies that $\angle BMA = \angle AMC = \pi/2$, and hence r swaps B and C.

□

Exercise. The line through BC in the proof above contains 2 segments with endpoints B, C. Are there two distinct solutions for r?

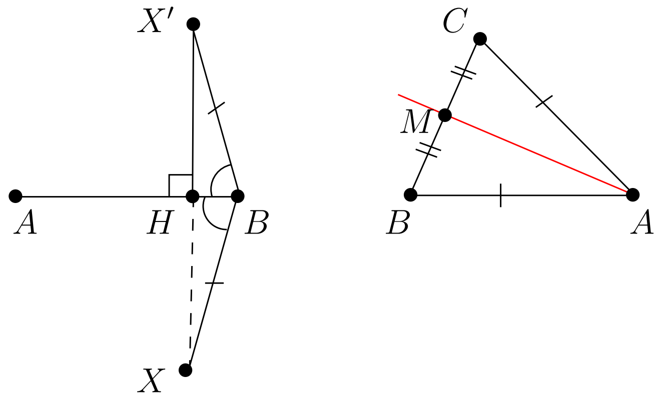


Figure 31: To the proofs of Propositions 2.30 and 2.31

Example 2.32. A glide reflection is an isometry defined by $f = r_l \circ R_{A,\varphi} = R_{A,\varphi} \circ r_l$, where r_l is a reflection with respect to a line l and $R_{A,\varphi}$ is a rotation about $A = Pol(l)$, see Fig. 32, left.

Theorem 2.33. *1. An isometry of S^2 is uniquely determined by the images of 3 non-collinear points.*

2. *Isometries act transitively on points of S^2 and on flags in S^2 (where a flag is a triples (A, l, h^+), where A is a point, l is a line containing A, and h^+ is a choice of hemisphere bounded by l).*
3. *The group $Isom(S^2)$ is generated by reflections.*
4. *Every isometry of S^2 is a composition of at most 3 reflections.*
5. *Every orientation-preserving isometry is a rotation.*
6. *Every orientation-reversing isometry is either a reflection or a glide reflection.*

Proof. 1-4 are proved similarly to their analogues in $\mathbb{E}^2$.

5: An orientation-preserving isometry of $\mathbb{S}^2$ is a composition of 2 reflections with respect to some lines l_1, l_2. As any two lines intersect non-trivially on $\mathbb{S}^2$, we conclude that it is a rotation.

6: See Lemma 2.34 below.

□

Lemma 2.34. *Let r_1, r_2, r_3 be distinct reflections not preserving the same point of S^2. Then $r_3 \circ r_2 \circ r_1$ is a glide reflection.*

Proof. To show the lemma we will use non-uniqueness of presentation of an isometry as a composition of reflections.
We will denote by r^*_X a reflection with respect to the line l^*_X. Also, denote $g = r_3 \circ r_2 \circ r_1$.

Notice, that the lines l_1, l_2, l_3 are all distinct and not passing through the same point.

- Let $A = l_1 \cap l_2$. Let l_2' be the line through A orthogonal to l_3. There exists a line l_1' through A such that $r_2 \circ r_1 = r_2' \circ r_1'$. Hence,

$$g = r_3 \circ r_2 \circ r_1 = r_3 \circ (r_2' \circ r_1') = (r_3 \circ r_2') \circ r_1',$$

  see Fig. 32 (the two diagrams in the middle).

- Similarly, let $B = l_3 \cap l_2'$. Let $l_3'' \perp l_1$ be the line through B orthogonal to l_1' and let l_2'' be the line such that $r_3 \circ r_2' = r_3'' \circ r_2''$ (i.e. $l_3'' \perp l_2''$), see Fig. 32 (the two diagrams on the right). Then we get

$$g = (r_3'' \circ r_2'') \circ r_1' = r_3'' \circ (r_2'' \circ r_1'),$$

  where r_3'' is the reflection in l_3'' and $(r_2'' \circ r_1')$ is the rotation about the point $l_2'' \cap l_1$ polar to l_3''. Hence, g is a glide reflection.

□

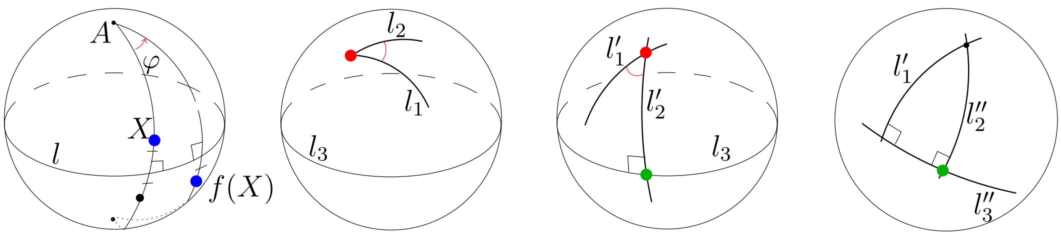


Figure 32: Glide reflection and a composition of three reflections

Remark. We could try to prove the lemma shorter by saying that $r_3 \circ r_2 \circ r_1 = (r_3 \circ r_2) \circ r_1$ is a composition of a rotation and reflection, as required. But we don't know (and it is not always true) that the centre of the rotation $(r_3 \circ r_2)$ is polar to the line of reflection r_3.

Exercise. What is the type of the antipodal map?

Remark 2.35. Fixed points of isometries on S^2 distinguish the types of isometries.

Indeed, fixed points of identity map, reflection r_l, rotation $R_{A,\alpha}$ and a glide reflection are the whole sphere, the line l, the pair of antipodal points A, A' and the empty set respectively.

Theorem 2.36. *(a) Every two reflections are conjugate in $Isom(S^2)$.*

(b) Rotations by the same angle are conjugate in $Isom(S^2)$.

Proof. Idea of proof:

(a) Let r_1 and r_2 be reflection with respect to the lines l_1 and l_2. Let l be an angle bisector for an angle formed by l_1 and l_2. Then $r_2 = r_l^{-1} \circ r_1 \circ r_l$ (indeed, r_l takes l_2 to l_1, then r_1 preserves l_1, then r_l^{-1} takes l_1 back to l_2, so, the composition $r_l^{-1} \circ r_1 \circ r_l$ preserves l_2 pointwise and changes the orientation, which means that it coincides with r_2).

(b) Let A and B be the centres of the two rotations $R_{A,\varphi}$,$R_{B,\varphi}$, let l be the orthogonal bisector of AB. Then $R^{-1}_{A,\varphi} = r_l^{-1} \circ R_{B,\varphi} \circ r_l$. Also, $R^{-1}_{A,\varphi}$ is conjugate to $R^{-1}_{A,\varphi}$ since the rotation $R_{A,\varphi} = r_2 \circ r_1$ is a composition of some reflections r_1, r_2, and the inverse is $R^{-1}_{A,\varphi} = r_1 \circ r_2 = r_1^{-1} \circ (r_2 \circ r_1) \circ r_1$.

□

Remark 2.37. As $\mathbb{S}^2 \subset E^3$, we have $Isom(\mathbb{S}^2) \subset Isom(\mathbb{E}^3)$ (more precisely, isometries of the sphere is the origin-preserving subgroup of isometries of $\mathbb{E}^3$). This is given by orthogonal 3×3 matrices (i.e. matrices satisfying $A^T A = AA^T = I$.)

Orientation reversing isometries correspond to matrices with $\det = -1$, while orientation-preserving to ones with $\det = 1$.
Orientation-preserving isometries form a subgroup given by

$$SO(3, \mathbb{R}) = \{A \in M_3 | A^T A = I, \det A = 1\}.$$

2.9 Platonic solids and their symmetry groups (NE)

(Non-examinable section)

We conclude our exposition of spherical geometry by a brief discussion of symmetry groups of Platonic solids, i.e. regular polyhedra known since antiquity, namely

tetrahedron, cube, octahedron, dodecaghedron and *icosohedron*

(see Fig.33, left to right).

Definition 2.38. By a regular polyhedron we mean a polyhedron P with largest possible group of symmetries G_P, i.e. the group G_P should act on P by isometries mapping its vertices to vertices, and the action $G_P : P$ should be transitive

- on vertices of P;
- on edges of P;
- on faces of P.

Moreover, G_P should act transitively on flags in P, i.e. on triples (V, E, F) where V is a vertex, and E is an edge such that $V \in E$, and F is a face of P such that $E \in F$.

To find a fundamental domain of the action, one needs to choose a flag (V_1, E_1, F_1) in P. Let $A = V_1$ be a vertex, and B be a midpoint of the edge E_1 and C be a centre of the face F_1. Then one can check that the triangle ABC is a fundamental domain of the action $G_P : P$.

Projecting P from its center O to a sphere centred at O one can turn the triangle ABC into a spherical triangle $A'B'C'$. One can check that the angles of this spherical triangle are

- $(\frac{\pi}{2}, \frac{\pi}{3}, \frac{\pi}{3})$ when P is a tetrahedron;
- $(\frac{\pi}{2}, \frac{\pi}{3}, \frac{\pi}{4})$ when P is a cube or an octahedron;
- $(\frac{\pi}{2}, \frac{\pi}{3}, \frac{\pi}{5})$ when P is a dodecahedron or an icosohedron.

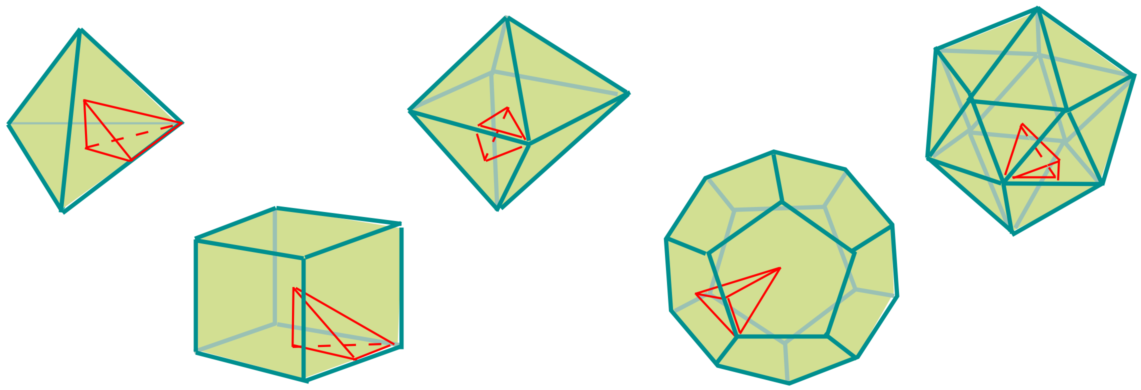


Figure 33: Regular polyhedra (from left to right): tetrahedron, cube, octahedron, dodecahedron and icosohedron.

One can also check that the group $G_P : S^2$ is generated by reflections with respect to the sides of the triangle $A'B'C'$.

Remark 2.39. Let $H : S^2$ be an action. As the sphere is a compact set, H acts on S^2 discretely if and only if H is a finite group.

Remark 2.40. A group H generated by reflections on S^2 is finite if and only if

- H is generated by 1 or 2 reflections;
- H is generated by reflections with respect to the sides of one of the following triangles with angles:
  - $(\frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{n})$, $n \in \mathbb{Z}$, $n \geq 2$;
  - $(\frac{\pi}{2}, \frac{\pi}{3}, \frac{\pi}{3})$, $(\frac{\pi}{2}, \frac{\pi}{3}, \frac{\pi}{4})$, $(\frac{\pi}{2}, \frac{\pi}{3}, \frac{\pi}{5})$.

Remark 2.41. Notice that the same group serves as the symmetry group for the cube and the octohedron - this is because the cube is <u>dual</u> to the octahedron (if we take a regular cube and mark the centeres of its faces, then the six marked points will be vertices of a regular octahedron; also, we can obtain a cube if we highlight the centeres of faces of the octahedron). Similarly, an icosohedron is dual to a dodecahedron, while a tetrahedron is dual to itself.

2.10 References

- In this section, we have mostly followed the exposition in V. V. Prasolov, *Non-Euclidean Geometry* (see Lecture I and pp. 48-49) or you can find the same material in pp. 83-87 of V. V. Prasolov, V. M. Tikhomirov *Geometry.*
- The spririt of our discussion of isometry group of the sphere follows the paper by Oleg Viro: O. Viro, Defining relations for reflections. I, arXiv:1405.1460v1.
- For another exposition concerning the isometry group of the sphere see G. Jones, Algebra and Geometry, Lecture notes (Section 2.2).
- More general notion of polarity comparing to the one considered in Section 2.3 is presented in Sections 16-17 of the following lecture notes: A. Barvinok, *Combinatorics of Polytopes.*
- One can read about tilings by triangles in V. V. Prasolov, *Non-Euclidean Geometry* , Lecture X, p. 34-36, or in V. V. Prasolov, V. M. Tikhomirov, "Geometry", Section 5.5, p 185-187.

3 Affine geometry

An affine space is a vector space whose origin we try to forget about.

Marcel Berger

We consider the same space $\mathbb{R}^2$ as in Euclidean geometry but with larger group acting on it.

3.1 Similarity group

Similarity group, $Sim(\mathbb{R}^2)$ is a group generated by all Euclidean isometries and scalar multiplications, i.e. transformations given by $(x_1, x_2) \mapsto (kx_1, kx_2)$, $k \in \mathbb{R}$.

Its elements may change size, but preserve the following properties:
angles, proportionality of all segments, parallelism, similarity of triangles.

This means that many problems in Euclidean geometry are actually problems about "similarity geometry".

Example 3.1. Consider the following theorem of Euclidean geometry:

A midline in a triangle is twice shorter than the corresponding side.

One can prove it as follows. Let M and N be the midpoints of AB and BC in the triangle ABC, see Fig. 34. Let $B = 0$ be the origin, consider the map $f : \mathbb{C} \to \mathbb{C}$ taking $z \to 2z$, i.e. the map which doubles every distance. Then for every segment I the length of $f(I)$ is twice the length of I. In particular, as $f(M) = A$ and $f(N) = C$, we get $|AC| = 2|MN|$.

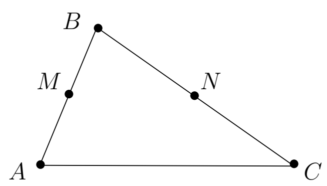


Figure 34: Length of midline using similarity

Remark. A map which may be written as a scalar multiplication in some coordinates in $\mathbb{R}^2$ is called homothety (with positive or negative coefficient depending on the sign of k).

Here, one can find the picture of a pantograph and a Sylvester machine - two mechanisms for implementing similarity (webpage by Rémi Coulon).

3.2 Affine geometry

Instead of scalar maps, as in "similarity geometry", now we will consider all non-degenerate linear maps.

Affine transformations are all transformations of the form $f(\boldsymbol{x}) = A\boldsymbol{x} + \boldsymbol{b}$ where $A \in GL(2, \mathbb{R})$.

Proposition 3.2. *Affine transformations form a group.*

Proof. We leave the proof as an exercise. You need to write $f(\boldsymbol{x}) = A\boldsymbol{x} + \boldsymbol{b}$ and to find the composition of two such maps, then to find f^{-1} and an identity map. The associativity will follow from associativity of composition.

□

Example 3.3. (a) Consider the map $f : \mathbb{C} \to \mathbb{C}$ given by $f(z) = 2z + 2 + i$. By definition $f \in Aff(\mathbb{R}^2)$, but also one can notice that $f \in Sim(\mathbb{R}^2)$.

(b) Now, consider $f : \begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} 2x + y + 1 \\ -x + y + 2 \end{pmatrix}$. As $\det \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix} = 3 \neq 0$, we conclude that $f \in Aff(\mathbb{R}^2)$. At the same time $f \notin Sim(\mathbb{R}^2$.

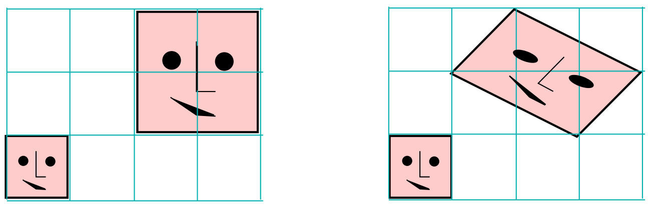


Figure 35: Examples of affine maps (see Example 3.3).

Affine transformations do not preserve length, angles, area.

Proposition 3.4. *Affine transformations preserve*

(1) collinearity of points;

(2) parallelism of lines;

(3) ratios of lengths on any line;

(4) concurrency of lines;

(5) ratio of areas of triangles (so ratios of all areas).

Proof. Linear maps preserve the properties (1)-(5), translations also preserve them. So, affine maps, as their compositions, also preserve all these properties.

□

Proposition 3.5. *(1) Affine transformations act transitively on triangles in* $\mathbb{R}^2$.

(2) *An affine transformation is uniquely determined by images of 3 non-collinear points.*

Proof. (1) Let ABC and $A'B'C'$. We <u>want</u> to find a map $f(\boldsymbol{x}) = A\boldsymbol{x} + \boldsymbol{b}$ such that $f(ABC) = A'B'C'$. We will find it as a composition $f = g \circ h$, where

$$A, B, C \xrightarrow{g} \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \xrightarrow{h} A', B', C'.$$

The map h is easy to find, and so is the map g^{-1}. This implies that the composition $f = g \circ h$ exists.

(2) Suppose there are two different affine transformations f and g taking the non-collinear points A, B, C to A', B', C'. Then the transformation $g^{-1} \circ f \neq id$ is a non-trivial transformation preserving all three points A, B, C. Let h be the affine transformation taking the points $\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ to A, B, C. Then the affine transformation $h^{-1} \circ (g^{-1} \circ f) \circ h$ preserves the points $\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ (and it is a non-trivial transformation, since it is conjugate to a non-trivial one). Which is a contradiction, as a transformation $A\boldsymbol{x} + \boldsymbol{b}$ taking the points $\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ to themselves clearly has $\boldsymbol{b} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

□

Example 3.6. We will use the affine group to show the following statement of Euclidean geometry:

The medians of a triangle in $\mathbb{E}^2$ are concurrent.

Proof.

- The statement is trivial for a regular triangle (as each of the three medians passes through the centre of the triangle).
- Apply an affine transformation f which takes some regular triangle to the given triangle ABC.
- f takes the medians of the regular triangle to the medians of ABC (as it maps vertices to vertices and midpoints to midpoints).
- So, it takes the intersection of the three medians to the intersection of the three medians of ABC.

□

Theorem 3.7. *Every bijection $f : \mathbb{R}^2 \to \mathbb{R}^2$ preserving collinearity of points, betweenness and parallelism is an affine map.*

Proof.

- Let g be an affine map which takes the points $\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ to $f(\begin{pmatrix} 0 \\ 0 \end{pmatrix}), f(\begin{pmatrix} 1 \\ 0 \end{pmatrix}), f(\begin{pmatrix} 0 \\ 1 \end{pmatrix})$.
  (this map exists by Theorem 3.5).

- We want to show $f(\boldsymbol{x}) = g(\boldsymbol{x})$ for all $\boldsymbol{x} \in \mathbb{R}^2$.

- We will denote the points by their complex coordinates, so by now we know the desired property for $0, 1, i$.

- As affine maps take parallel lines to parallel lines and f also preserves collinearity, we conclude that $f(\boldsymbol{x}) = g(\boldsymbol{x})$ also for $\boldsymbol{x} = 1+i$ (as $1+i$ lies on the line though 1 parallel to the line through O and i and on also it lies on the line through i parallel to the line through 0 and 1), see Fig. 36, left.

- Similarly, we use the points $i, 1+i, 1$ to conclude the property for the point 2, see Fig. 36 middle left.

- Applying this procedure, one can show the property for all integer points $a+bi$, $a, b, \in \mathbb{Z}$.

- Every half-integer point $a+bi$, $a, b, \in \frac{1}{2}\mathbb{Z}$ can be obtained as an intersection of two segments with integer endpoints, so the property also holds for half-integer points, see Fig. 36 middle right and right..

- Applying the previous step again, we obtain the property for $\frac{1}{4}$-integer points, then for $\frac{1}{8}$-integer points, and so on... We will get smaller and smaller lattices.

- As f preserves betweenness and coincides with g on a dense set of points, we conclude that f is continuous and $f(\boldsymbol{x}) = g(\boldsymbol{x})$ for all $\boldsymbol{x} \in \mathbb{R}^2$.

  (More precisely, we first conclude this for all horisontal and vertical lines $x_1 = a$ and $x_2 = b$, where $a, b \in \mathbb{Z}/2^n$ for some n, and then extend it to any point (x_1, x_2) by looking at any non-horizontal and non- vertical line l through it).

$\square$

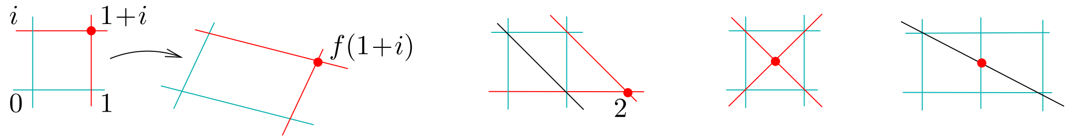


Figure 36: To the proof of Theorem 3.7.

Remark. If f is a bijection $\mathbb{R}^2 \to \mathbb{R}^2$ preserving collinearity, then it preserves parallelism and betweenness.

Proof. Parallelism: of f takes parallel lines to the lines intersecting at the point A, consider $f^{-1}(A)$. It exists because f is a bijection, and it would lie on both of the parallel lines as f preserves collinearity. The contradiction shows that f preserves parallelism.
Betweenness: the argument here is much more involved, we will skip it. You can find the argument on pp.40-41 in the book by Prasolov and Tikhomirov. $\square$

This allows as to reformulate Theorem 3.7 as follows.

Theorem 3.7'. (The fundamental theorem of affine geometry).
Every bijection $f : \mathbb{R}^2 \to \mathbb{R}^2$ preserving collinearity of points is an affine map.

Corollary 3.8. *If $f : \mathbb{R}^2 \to \mathbb{R}^2$ is a bijection which takes circles to circles, then f is an affine map.*

Proof.

(1) The transformation f^{-1} maps three collinear points $f(A), f(B), f(C)$ to 3 three collinear points A, B, C.

   - Indeed, if the points A, B, C are not collinear, then they are pairwise distinct and there is a circle through A, B, C.
   - Hence, $f(A), f(B), f(C)$ are also pairwise distinct (as f is bijective) and lie on a circle (since f maps circles to circles).
   - Then $f(A), f(B), f(C)$ cannot lie on one line.

(2) From (1) and Theorem 3.7' we conclude that f^{-1} is affine, which implies that f is also affine.

□

Remark. An affine transformation takes ellipses to ellipses. So, in Corollary 3.8 we can change the circles to ellipses.

Example 3.9 (Parallel Projection)**.** Consider two copies α and β of a two-dimensional plane in $\mathbb{R}^3$, let suppose that each of α and β are endowed with coordinates. Project from α to β by parallel rays (the rays should not be parallel to any of α and β!). Then we get a bijection between the two planes, and one can see that this bijection is preserving parallelism (indeed, if two parallel lines $l, m \in \alpha$ are mapped to intersecting lines $l', m' \in \beta$, then what is the preimage of the intersection $l' \cap m' \in \beta$?). Applying the fundamental theorem of affine geometry, we conclude that the parallel projection is an affine map.

Proposition 3.10. *Every parallel projection is an affine map, but not every affine map is a parallel projection.*

Proof. It is already shown in Example 3.9 that the parallel projections are affine maps. To see the second statement, consider the affine map $f : z \to 2z$:

- Suppose that f is a parallel projection $f : \alpha \to \beta$.
- The planes α and β are not parallel (otherwise, f would be an isometry, which is not the case).
- Consider the line of intersection $\alpha \cap \beta$. Every point of this line is mapped by f to itself, so the distance between two points on that line is preserved.
- At the same time $z \to 2z$ makes all distances twice longer. So, $f : z \to 2z$ cannot be a parallel projection.

□

Exercise 3.11. Every affine map can be obtained as a composition of two parallel projections. (See also p.18 in *Geometry, Lecture notes*, by Norbert Peyerimhoff).

3.3 References

- Most of the material above (and more information on affine geometry) may be found in
  G. Jones, *Algebra and Geometry*, Lecture notes (Section 3).

- For fundamental theorem of affine geometry and its corollaries see
  V. V. Prasolov, V. M. Tikhomirov, *Geometry* , Section 2.1. pp.39-42.

- For another exposition of affine geometry, based on parallel projection, see
  N. Peyerimhoff, *Geometry, Lecture notes*, (Section 2, Section 2.1 and 2.2.).

- Illustrating Mathematics by Rémi Coulon: a panthograph and a Sylvester machine.

Albrecht Dürer, *The Draughtsman of the Lute*. Woodcut.
From Dürer's " Unterweysung der Messung mit dem Zyrkel und Richtscheyd", 1525.
Image from https://www.metmuseum.org/art/collection/search/387741
(OA public domain)

4 Projective geometry

Projective geometry is all geometry.
Arthur Caley

Motivation: We have considered larger and larger groups acting on the same space $\mathbb{R}^2$, now we are going to consider even larger group $Proj(a)$ of projective transformations:

$$Isom(\mathbb{E}^2) \subset Sim(2) \subset Aff(2) \ \subset \ Proj(2).$$

The bigger is the group acting, the smaller is the set of properties it preserves. Now, we will extend the group so that is will only preserve collinearity (but not parallelism or betweenness).

The group $Proj(2)$ of projective transformations will act transitively on the pairs of lines, in particular there will be transformations taking intersecting lines into parallel. The intersection point of the lines in this case still needs to be mapped somewhere. This motivates the idea of adding some points to the plane, namely "points at infinity" (we will have infinitely many of them, more precisely, one point for each direction).

4.1 Projective line, $\mathbb{RP}^1$

Model:

- **Points** of the projective line are lines though the origin O in $\mathbb{R}^2$.

  On the plane with coordinates (x_1, x_2) consider the line l_0 given by the equation $x_2 = 1$. Then every line l through the origin O can be represented by the coordinates of the intersection $l \cap l_0 = (x, 1)$, except for the line Ox_1 which does not intersect l_0, see Fig. 37.
  We will assign to Ox_1 a special point, "point at infinity" and will denote it x_∞.

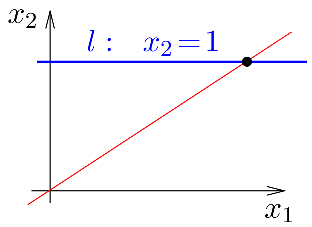


Figure 37: Projective line: set of lines through O in $\mathbb{R}^2$.

- **Group action**: $GL(2, \mathbb{R})$ acts on $\mathbb{R}^2$ by mapping a line though O to another line through O: a matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with $ad - bc \neq 0$ maps the point $(\lambda x, \lambda) \in l$ to

  $$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \lambda x \\ \lambda \end{pmatrix} = \lambda \begin{pmatrix} ax + b \\ cx + d \end{pmatrix}.$$

  If $cx + d \neq 0$, we can write $A : (x, 1) \to (\frac{ax+b}{cx+d}, 1)$.
  The point $(-d/c, 1)$ is mapped to x_∞. So, $GL(2, \mathbb{R})$ acts on $\mathbb{RP}^1$.

- **Homogeneous coordinates**: a line through O is determined by a pair of numbers (ξ_1, ξ_2), where $(\xi_1, \xi_2) \neq (0, 0)$.

  The pairs (ξ_1, ξ_2) and $(\lambda\xi_1, \lambda\xi_2)$ determine the same line, so are considered as equivalent.

  The ratio $(\xi_1 : \xi_2)$ determines the line and is called homogeneous coordinates of the corresponding point in $\mathbb{RP}_1$.

  The $GL(2, \mathbb{R})$-action in homogeneous coordinates writes as

  $$A: \ (\xi_1 : \xi_2) \mapsto (a\xi_1 + b\xi_2 : c\xi_1 + d\xi_2), \text{ where } A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

  and is called a projective transformation.

Remark. Projective transformations are called this way since they are compositions of projections (of one line to another line from a point not lying on the union of that lines). The following several statements will help us to prove that projective transformations are exactly the set of all possible compositions of such projections.

Lemma 4.1. *Let points $A_2.B_2, C_2, D_2$ of a line l_2 correspond to the points A_1, B_1, C_1, D_1 of the line l_1 under the projection from some point $O \notin l_1 \cup l_2$. Then*

$$\frac{|C_1A_1|}{|C_1B_1|} \Big/ \frac{|D_1A_1|}{|D_1B_1|} = \frac{|C_2A_2|}{|C_2B_2|} \Big/ \frac{|D_2A_2|}{|D_2B_2|}.$$

Proof. For a triangle Δ let S_Δ denote the Euclidean area of Δ. Recall that given a Euclidean triangle ABC with altitude BH one has

$$S_{ABC} = \frac{1}{2}|BH| \cdot |AC| = \frac{1}{2}|AB| \cdot |AC| \sin \angle BAC. \tag{4.1}$$

In particular, $S_{OC_1A_1} = \frac{A_1C_1 \cdot h}{2}$, $S_{OC_1B_1} = \frac{A_1B_1 \cdot h}{2}$, where h is the distance from O to the line l_1. Hence, we have

$$\frac{|C_1A_1|}{|C_1B_1|} = \frac{S_{OC_1A_1}}{S_{OC_1B_1}} \overset{(4.1)}{=} \frac{|OC_1||OA_1| \sin \angle A_1OC_1}{|OC_1||OB_1| \sin \angle B_1OC_1} = \frac{|OA_1| \sin \angle A_1OC_1}{|OB_1| \sin \angle B_1OC_1},$$

which implies that

$$\begin{aligned}\frac{|C_1A_1|}{|C_1B_1|} \Big/ \frac{|D_1A_1|}{|D_1B_1|} &= \frac{|OA_1| \sin \angle A_1OC_1}{|OB_1| \sin \angle B_1OC_1} \Big/ \frac{|OA_1| \sin \angle A_1OD_1}{|OB_1| \sin \angle B_1OD_1} \\ &= \frac{\sin \angle A_1OC_1}{\sin \angle A_1OD_1} \cdot \frac{\sin \angle B_1OD_1}{\sin \angle B_1OC_1} = \frac{\sin \angle A_2OC_2}{\sin \angle A_2OD_2} \cdot \frac{\sin \angle B_2OD_2}{\sin \angle B_2OC_2} = RHS.\end{aligned}$$

□

Definition 4.2. Let A, B, C, D be four points on a line l, and let a, b, c, d be their coordinates on l. The value $[A, B, C, D] := \frac{c-a}{c-b} \big/ \frac{d-a}{d-b}$ is called the cross-ratio of these points.

So, we can reformulate Lemma 4.1 as follows.

Lemma 4.1'. Projections preserve cross-ratios of points.

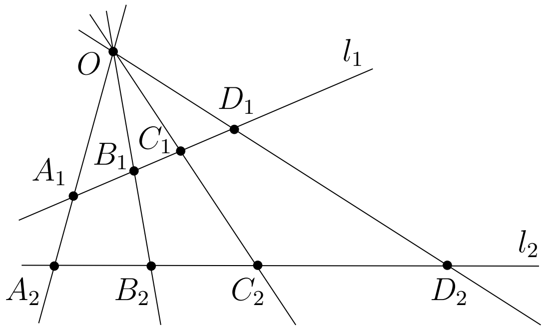


Figure 38: Projection preserves cross-ratio.

Definition 4.3. The cross-ratio of four lines lying in one plane and passing through one point is the cross-ratio of the four points at which these lines intersect an arbitrary line l.

Remark. By Lemma 4.1', Definition 4.3 does not depend on the choice of the line l.

Proposition 4.4. *Any composition of projections is a linear-fractional map.*

Proof. Let f be a composition of projections. Let a', b', c' be images of points a, b, c under a composition of projections. By Lemma 4.1', $[a, b, c, x] = [a', b', c', f(x)]$, i.e.

$$\frac{c-a}{c-b} \Big/ \frac{x-a}{x-b} = \frac{c'-a'}{c'-b'} \Big/ \frac{f(x)-a'}{f(x)-b'}.$$

Expressing $f(x)$ from this equation we get $f(x) = \frac{\alpha x+\beta}{\gamma x+\delta}$ for some $\alpha, \beta, \gamma, \delta$.

□

Proposition 4.5. *A composition of projections preserving 3 points is an identity map.*

Proof. We leave the proof as an exercise.
Hint: use $f(x) = \frac{\alpha x+\beta}{\gamma x+\delta}$ and show that if f fixes three points then either $f(x) = x$ or there is a quadratic equation with 3 roots.

□

Lemma 4.6. *Given $A, B, C \in l$ and $A', B', C' \in l'$, there exists a composition of projections which takes A, B, C to A', B', C'.*

Proof.

- Consider any line l'' such that $A' \in l''$ and $l'' \neq l'$. Let $O \in AA'$ be any point, see Fig. 39.
- Project B, C from O to l''. This will define points B'' and C'' respectively.
- Let $P = B'B'' \cap C'C''$. Project l'' to l' from P. The composition of the two projections takes points A, B, C to $A'B'C'$.

□

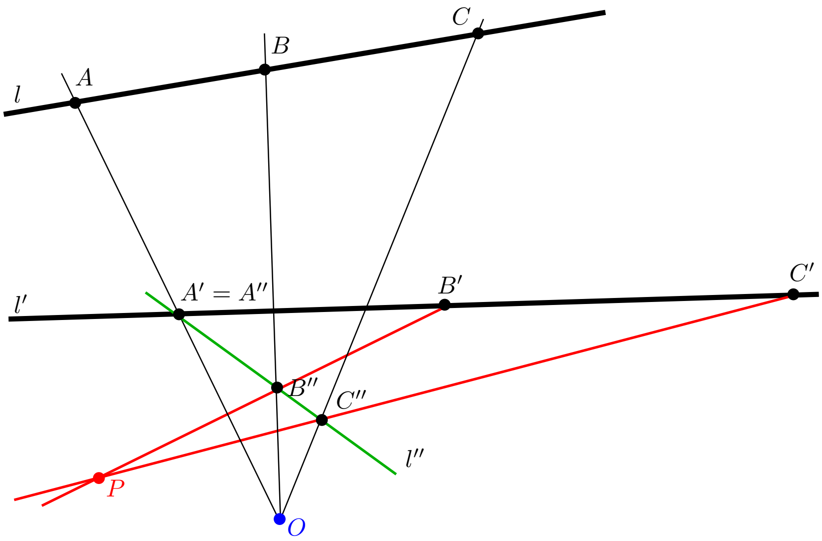


Figure 39: To the proof of Lemma 4.6.

Remark. If in the proof above $B'B''||C'C''$ we can chose another line l'' so that the lines will not be parallel (in particular, if we move l'' so that it crosses BO and CO closer to the point O, then the intersection $P = B'B'' \cap C'C''$ moves also closer to O).

Theorem 4.7.

(a) The following two definitions of projective transformations of $\mathbb{RP}^1$ are equivalent:

(1) Projective transformations are compositions of projections;

(2) Projective transformations are linear-fractional transformations.

(b) A projective transformation of a line is determined by images of 3 points.

Proof. First, we will prove part (a) of the theorem.

(1)⇒(2) Compositions of projections are linear-fractional transformations by Proposition 4.4.

(1)⇐(2) We will prove this in three steps.

(i) We will now show that

linear-fractional transformations preserve cross-ratios.

Indeed, if $y_i = \frac{\alpha x_i + \beta}{\gamma x_i + \delta}$, then one can check that

$$y_i - y_j = \frac{(\alpha\gamma - \beta\delta)(x_i - x_j)}{(\gamma x_i + \delta)(\gamma x_j + \delta)}.$$

Denote $u_i = \frac{1}{\gamma x_i + \delta}$. Then

$$\frac{y_3 - y_1}{y_3 - y_2} \Big/ \frac{y_4 - y_1}{y_4 - y_2} = [x_1, x_2, x_3, x_4] \frac{u_3 \cdot u_1}{u_3 \cdot u_2} \Big/ \frac{u_4 \cdot u_1}{u_4 \cdot u_2} = [x_1, x_2, x_3, x_4].$$

(ii) Hence, a linear-fractional transformation is determined by the images of 3 points. Indeed, if there are two linear-fractional transformations f and g which take A, B, C to A', B', C', then $g^{-1} \circ f$ is a non-triavial linear-fractional transformaion preserving three points A, B, C, which is impossible as would lead to a quadratic equation with 3 roots (compare to the proof of Proposition 4.5).

(iii) Let f be a linear-fractional transformation. By Lemma 4.6, there exists a composition of projections φ which takes $A, B, C \in \mathbb{R}$ to $f(A), f(B), f(C)$. In view of the part ((1)⇒(2)), the map φ is linear-fractional. Then Step (ii) implies that $\varphi = f$ (i.e. a linear-fractional map f is the composition of projection φ).

This completes the proof of part (a) of the theorem. Part (b) follows now from Step (ii).

□

4.2 Projective plane, $\mathbb{RP}^2$

<u>**Model:**</u>

- **Points** of $\mathbb{RP}^2$ are lines through the origin O in $\mathbb{R}^3$.
  Let x_1, x_2, x_3 be coordinates in $\mathbb{R}^3$ and let $\alpha \in \mathbb{R}^3$ be the plane $x_3 = 1$.
  For each line $l \notin Ox_1x_2$ take a point $l \cap \alpha$, see Fig. 40.
  For each line in the plane Ox_1x_2 assign a <u>"point at infinity"</u>.

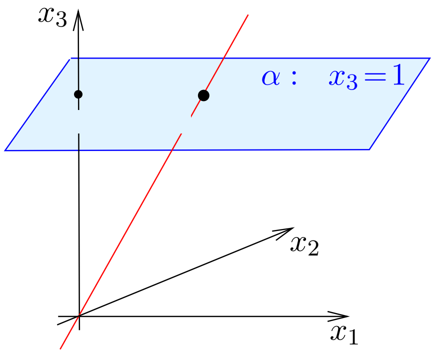


Figure 40: Projective plane: set of lines through O in $\mathbb{R}^3$.

- **Lines** of $\mathbb{RP}^2$ are planes through O in $\mathbb{R}^3$.
  All points at infinity form a <u>line at infinity</u> (a copy of $\mathbb{RP}^1$).

- **Group action:** $GL(3, \mathbb{R})$ (acts on $\mathbb{R}^3$ mapping a line though O to another line through O).

- **Homogeneous coordinates:**

  · A line though O is determined by a triple of numbers (ξ_1, ξ_2, ξ_3), where $(\xi_1, \xi_2, \xi_3) \neq (0, 0, 0)$.

- Triples (ξ_1, ξ_2, ξ_3) and $(\lambda\xi_1, \lambda\xi_2, \lambda\xi_3)$ determine the same line, so are considered equivalent.
- So, lines are in bijection with ratios $(\xi_1 : \xi_2 : \xi_3)$ called homogeneous coordinates.

- **Projective transformations** in homogeneous coordinates:

  $A: \ (\xi_1, \xi_2, \xi_3) \mapsto (a_{11}\xi_1 + a_{12}\xi_2 + a_{13}\xi_3 : a_{21}\xi_1 + a_{22}\xi_2 + a_{23}\xi_3 : a_{31}\xi_1 + a_{32}\xi_2 + a_{33}\xi_3)$, where $A = (a_{ij}) \in GL(3, \mathbb{R})$.

- **Points and lines in $\mathbb{RP}^2$:**
  - Points are lines through O in $\mathbb{R}^3$;
  - Lines are 2-dimensional planes through O in $\mathbb{R}^3$, see Fig. **??**.
  - A plane through O can be written as

$$c_1x_1 + c_2x_2 + c_3x_3 = 0, \tag{4.2}$$

    where $(c_1, c_2, c_3) \neq (0, 0, 0)$.
  - If $(c_1, c_2) \neq (0, 0)$ then the plane defined by Equation 4.2 makes a trace on the plane $x_3 = 1$; this trace if the line given by

$$\begin{cases} \frac{c_1}{c_3}x_1 + \frac{c_2}{c_3}x_2 = -1 & \text{for } c_3 \neq 0 \\ c_1x_1 + c_2x_2 = 0 & \text{for } c_3 = 0 \end{cases}$$

  - The plane $x_3 = 0$ gives a "line at infinity".

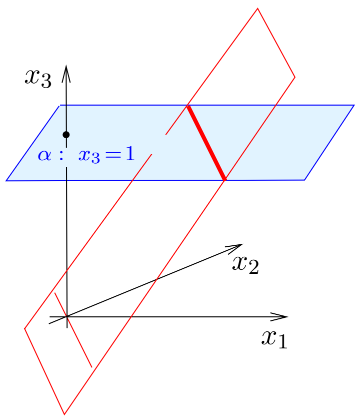


Figure 41: Projective plane: lines are planes through O in $\mathbb{R}^3$.

Remark.

(1) A unique line passes through any given two points in $\mathbb{RP}^2$ (as a unique plane through the origin passes through any two lines intersecting at the origin).

(2) Any two lines in $\mathbb{RP}^2$ intersect at a unique point (as any two planes through O in $\mathbb{R}^3$ intersect by a line through O).

(3) Relation 4.2 establishes **duality** between points and lines in $\mathbb{RP}^2$:
(the point $(c_1 : c_2 : c_3)$ is dual to the plane $c_1x_1 + c_2x_2 + c_3x_3 = 0$).
So, for any theorem about points in $\mathbb{RP}^2$ there should be a dual theorem about lines.

Theorem 4.8. *Projective transformations of $\mathbb{RP}^2$ preserve cross-ratio of 4 collinear points.*

Proof.
- Let f be a projective transformation and let $\beta \in \mathbb{R}^3$ be the plane through the origin containing the four collinear points whose cross-ratio we consider.
- Find an isometry $i \in Isom(\mathbb{R}^3)$ which takes β to the plane $f(\beta)$.
- Let $\varphi = f \circ i^{-1}$, i.e. $f = \varphi \circ i$. Notice that φ is a projective transformation of the projective line β (as φ is a composition of a projective transformation and an isometry).
- i preserves cross-ratios (as it is an isometry), and φ preserves cross-ratios by Theorem 4.7. This implies that f preserves cross-ratio of the considered points (as a composition of cross-ratio preserving maps).
- As the quadruple of collinear points was chosen randomly, we conclude that f preserves all cross-ratios.

□

Definition. A <u>triangle</u> in $\mathbb{RP}^2$ is a triple of non-collinear points.

Proposition 4.9. *All triangles of $\mathbb{RP}^2$ are equivalent under projective transformations.*

Proof. There exists an element of $GL(3, \mathbb{R})$ which takes three given linearly independent vectors to three other given linearly independent vectors.

□

Definition 4.10. A <u>quadrilateral</u> in $\mathbb{RP}^2$ is a set of four points, no three of which are collinear.

Proposition 4.11. *For any quadrilateral Q in $\mathbb{RP}^2$ there exists <u>a unique</u> projective transformation which takes Q to a given quadrilateral Q'.*

Proof.
- It is sufficient to prove the statement for the fixed quadrilateral

$$Q' = Q_0 = [(1:0:0), (0:1:0), (0:0:1), (1:1:1)].$$

Indeed, if we have projective transformations $f : Q \to Q_0$ and $g : Q' \to Q_0$, then $g^{-1} \circ f$ is a projective transformation mapping $Q \to Q'$. Moreover, if $\varphi \neq g^{-1} \circ f$ is another projective transformation taking Q to Q' then $g \circ \varphi \neq f$ is another projective transformation mapping Q to Q_0.

- By Proposition 4.9 we may assume that

$$Q = [(1:0:0), (0:1:0), (0:0:1), (a:b:c)].$$

Then $f = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}$ is the unique map taking Q_0 to Q, which implies that f^{-1} is the unique map taking Q to Q_0.

□

Theorem 4.12. *A bijective map from $\mathbb{RP}^2$ to $\mathbb{RP}^2$ preserving projective lines is a projective map.*

Proof. Consider a bijection $f : \mathbb{RP}^2 \to \mathbb{RP}^2$. Let l_∞ be the line at infinity and $f(l_\infty)$ be its image under f. Consider a projective map φ which maps $f(l_\infty)$ to l_∞ (it does exists as there is a projective map taking any two points in $\mathbb{RP}^2$ to any other two points in $\mathbb{RP}^2$). Then the map $\psi = \varphi \circ f$ takes l_∞ to itself (so, one can restrict it two $\psi : \mathbb{R}^2 \to \mathbb{R}^2$). Also, ψ preserves collinearity (as a composition of the transformation f preserving collinearity with a projective transformation).

Hence, by Fundamental Theorem of affine geometry the map $\psi = \varphi \circ f$ is affine. This implies that the map $f = \varphi^{-1} \circ \psi$ is projective (as a composition of an affine and projective transformations).

□

Corollary 4.13. *A projection of a plane to another plane is a projective map.*

Proof. As a projection preserves the lines, Theorem 4.12 implies that it is a projective map.

□

Remark. A projection of a plane α to another plane β is not an affine map if α is not parallel to β, as in this case some line from α will not be mapped to β.

4.3 Some classical theorems on $\mathbb{RP}^2$

Remark on projective duality:

point $A = (a_1 : a_2 : a_3)$	$\longleftrightarrow$	line $l_A : \quad a_1x_1 + a_2x_2 + a_3x_3 = 0$
$A \in l_B$	$\longleftrightarrow$	$B \in l_A$
line through A, B	$\longleftrightarrow$	point of intersection: $l_A \cap l_B$
3 collinear points	$\longleftrightarrow$	3 concurrent lines
...	$\longleftrightarrow$	...

Proposition 4.14 (On dual correspondence)**.** *The interchange of words "point" and "line" in any statement about configuration of points and lines related by incidence does not affect validity of the statement.*

Proof. The relation $a_1x_1 + a_2x_2 + a_3x_3 = 0$ is symmetric with respect to the coordinates of the point X and the line l_A, applying duality we only change the geometric interpretation of the equations. Algebra remains the same.

□

Theorem 4.15 (Pappus' theorem)**.** *Let a and b be lines, $A_1, A_2, A_3 \in a$, $B_1, B_2, B_3 \in b$. Let $P_3 = B_1A_2 \cap A_1B_2$, $P_2 = B_1A_3 \cap A_1B_3$, $P_1 = B_3A_2 \cap A_3B_2$. Then the points P_1, P_2, P_3 are collinear.*

Proof.

- Let $P_2' = B_1A_3 \cap P_1P_3$, let $C = B_1A_3 \cap A_1B_2$.
  We need to show that $P_2 = P_2'$, see Fig. 44, left.

- Consider a composition f of 3 projections:

$$B_1A_3 \xrightarrow{A_1} b \xrightarrow{A_2} B_2A_3 \xrightarrow{P_3} B_1A_3,$$

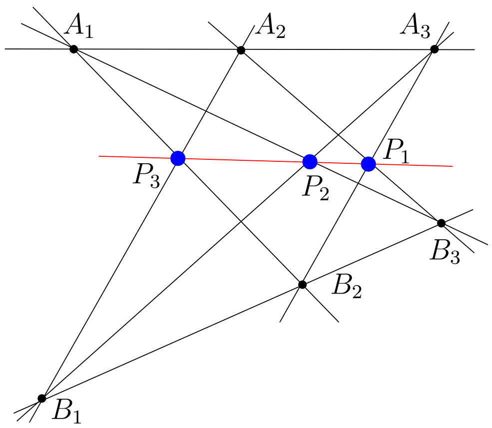
where $l_1 \xrightarrow{A} l_2$ denotes a projection of l_1 to l_2 from A, see Fig. 44, right.

- Notice that f takes $C \to B_2 \to B_2 \to C$, so $f(C) = C$.
Also it takes $B_1 \to B_1 \to B_1A_2 \cap B_2A_3 \to B_1$, so $f(B_1) = B_1$.
One can check similarly that $f(A_3) = A_3$ and $f(P_2) = P_2'$.

- So, this is a projective transformation of the line B_1A_3 preserving the points C, B_1, A_3. By Theorem 4.7 (b), f is identity map.

- Since $f(P_2) = (P_2')$ we conclude that $P_2 = P_2'$.

□



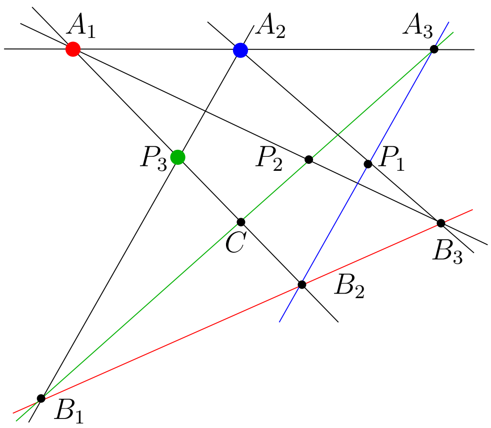


Figure 42: Pappus' Theorem and its proof by composition of 3 projections.

Remark. Sketch of another proof of Pappus' Theorem:

- By Proposition 4.11 there exists a projective map taking the points $A_1A_2B_2B_1$ to vertices of a unit square.

- So, we may assume that the points A_1, A_2, A_3 are $(0,1), (1,1), (a,1)$ and the points B_1, B_2, B_3 are $(0,0), (1,0), (b,0)$.

- Then it is easy to compute the coordinates of the points P_1, P_2, P_3 and check that the points are collinear.

- To establish collinearity of the points, check that the vectors P_1P_2 and P_1P_3 are proportional.

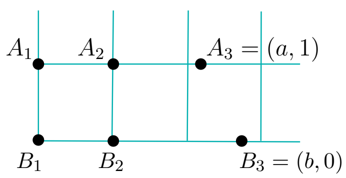


Figure 43: Another proof of Pappus' Theorem.

Remark 4.16 (Dual statement to Pappus' theorem)**.** *Let A and B be points and a_1, a_2, a_3 be lines through A, and b_1, b_2, b_3 be lines through B.*
Let p_1 be a line through $b_2 \cap a_3$ and $a_2 \cap b_3$,
p_2 be a line through $b_1 \cap a_3$ and $a_1 \cap b_3$,
p_3 be a line through $b_2 \cap a_1$ and $a_2 \cap b_1$.
Then the lines p_1, p_2, p_3 are concurrent.

(This is actually the same statement as Pappus' theorem itself!)

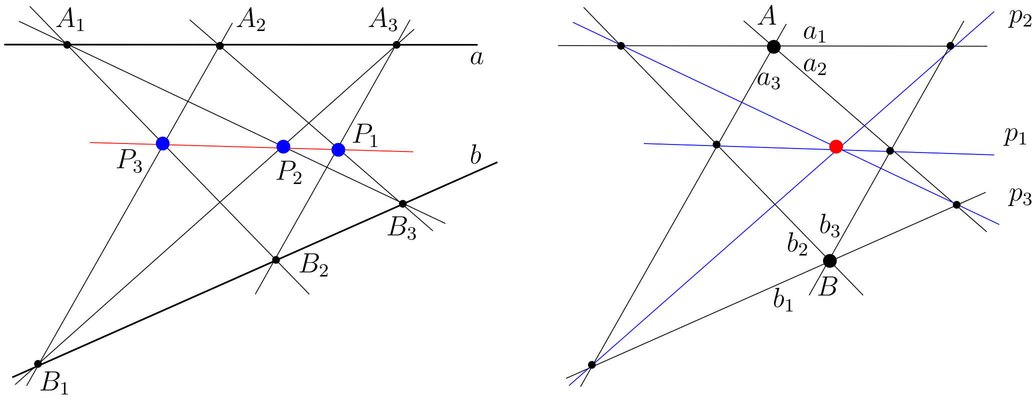


Figure 44: Pappus' Theorem and the dual statement.

Remark 4.17. Pappus' theorem is a special case of Pascal's Theorem (see Fig. 45):
If A, B, C, D, E, F lie on a conic then the points $AB \cap DE$, $BC \cap EF$, $CD \cap FA$ are collinear.

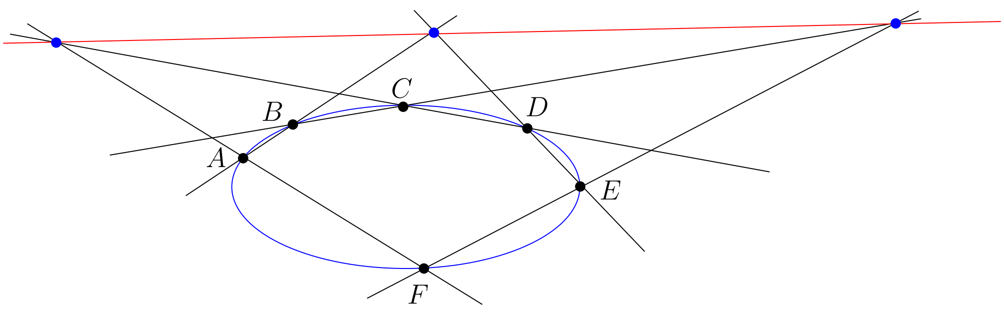


Figure 45: Pascal's Theorem.

We leave Pascal's Theorem without proof, you can find the proof in

- V. V. Prasolov, V. M. Tikhomirov. *Geometry*, (2001). Section 4.2, p. 71.

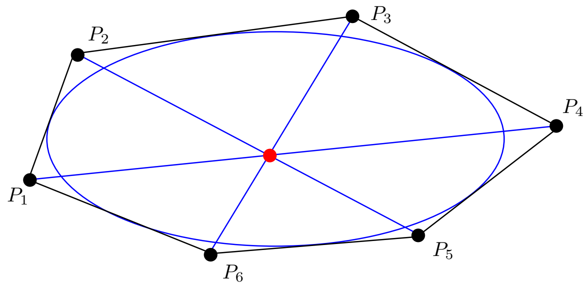


Figure 46: Brianchon's Theorem.

Remark 4.18. Dual to Pascal's Theorem is Brianchon's Theorem (see Fig. 46):
Let $P_1P_2P_3P_4P_5P_6$ be a hexagon formed by 6 tangent lines to a conic. Then the lines P_1P_4, P_2P_5, P_3P_6 are concurrent.

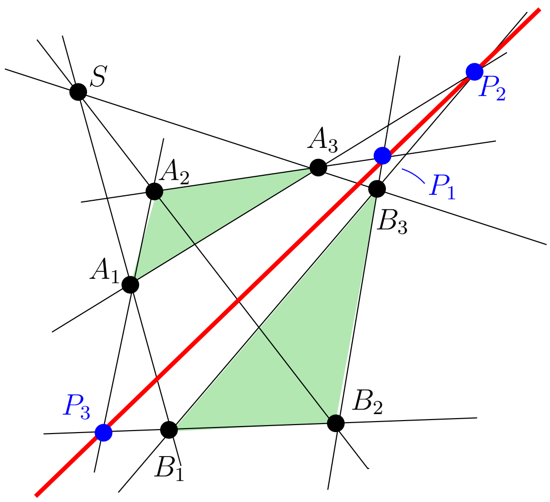


Figure 47: Desargues' Theorem.

Theorem 4.19 (Desargues' theorem)**.** *Suppose that the lines joining the corresponding vertices of triangles $A_1A_2A_3$ and $B_1B_2B_3$ intersect at one point S Then the intersection points $P_1 = A_2A_3 \cap B_2B_3$, $P_2 = A_1A_3 \cap B_1B_3$, $P_3 = A_1A_2 \cap B_1B_2$ are collinear.*

Proof. The idea of the proof is as follows. First, we will show a 3-dimensional analogue of the statement (and this will be short and easy part (a)). Then, in part (b) of the proof, we will get the 2-dimensional statement as a limit of deformation of the 3-dimensional configuration.

(a) Let α be a plane in $\mathbb{R}^3$ containing points A_1, A_2, A_3, and β be a plane containing points B_1, B_2, B_3. Let $l = \alpha \cap \beta$ be the intersection line. And suppose that the lines joining the corresponding vertices of triangles $A_1A_2A_3$ and $B_1B_2B_3$ intersect at one point S. Notice that the lines A_iA_j and B_iB_j lie in one plane (passing through P, A_i, A_j and B_i, B_j), so, they are either parallel or intersect. The intersection point of $A_iA_j \in \alpha$ and $B_iB_j \in \beta$ can only lie on $l = \alpha \cap \beta$), see Fig. 48, left. (In particular, if A_iA_j is parallel to B_iB_j, we may understand this as intersection at the point at infinity on the line l). So, all three points $P_k = A_iA_j \cap B_iB_j$, $(k = 1, 2, 3, k \neq i \neq j)$ belong to l.

(b)
- Now, we consider the 2-dimensional configuration (we place it into a horizontal plane γ in 3-dimensional space).
- Let $O \notin \gamma$ be any point such that the plane $OA_2B_2 \perp \gamma$,, see Fig. 48, right.
- Choose a point $A_2' \in OA_2$, and consider a point $B_2' = OB_2 \cap SA_2$.
- Consider the triangle $A_1A_2'A_3$ and $B_1B_2'B_3$, denote the planes containing them by α and β respectively. By part (a) of the proof, the three intersection points constructed for these triangles lie on the line $l = \alpha \cap \beta$.
- Now, we start to move the point A_2' towards A_2. The planes α and β approach the initial horizontal plane γ. The intersection line $l = \alpha \cap \beta$ approaches some line in γ. This line at the limit will be the line containing all three points $P_1, P_2, P_3 \in \gamma$.

$\square$

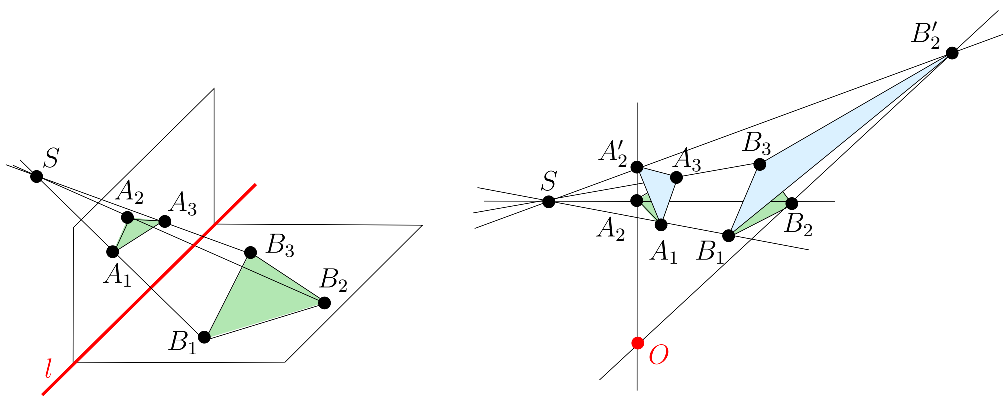


Figure 48: Proof of Desargues' Theorem.

4.4 Topology and metric on $\mathbb{RP}^2$

Remark 4.20 (Topology of $\mathbb{RP}^2$)**.** $\mathbb{RP}^2$ is a set of lines through O in $\mathbb{R}^3$, in other words $\mathbb{RP}^2 = \mathbb{S}^2/\sim$, i.e. the sphere with antipodal points identified, which is equivalent to a disc with the opposite points identified.

It includes a Möbius band, so, it is one-sided and non-orientable.

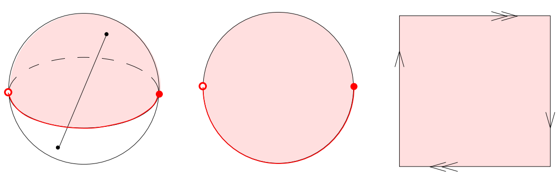


Figure 49: Topology of $\mathbb{RP}^2$.

Remark 4.21 (Elliptic geometry).

- As $\mathbb{RP}^2 = \mathbb{S}^2/\sim$, one can use the spherical metric to introduce the metric on the set of points of $\mathbb{RP}^2$. Then $\mathbb{RP}^2$ with this metric will be locally isometric to $\mathbb{S}^2$, i.e. a small domain on $\mathbb{RP}^2$ is isometric to a small domain on $\mathbb{S}^2$.

- However, most projective transformations to not preserve this metric. So, this metric is not a notion of projective geometry.

- The geometry of $\mathbb{RP}^2$ with spherical metric (and a group of isometries acting on it) is called elliptic geometry and has the following properties:

  (1) For any two distinct points there exists a unique line through these points;

  (2) Any two distinct lines intersect at a unique point;

  (3) For any line l and point p (which is not a pole for l) there exists a unique line l' such that $p \in l'$ and $l \perp l'$.

  (4) The group of isometries acts transitively on the points (and lines) of this geometry.

Remark 4.22 (Conic sections).

- Quadrics, i.e. the curves of second order on $\mathbb{R}^2$ (such as ellipse, parabola and hyperbola) may be obtained as conic sections (sections of a round cone by a plane, see Fig. 50).

- Ellipse, parabola and hyperbola are equivalent under projective transformations (to see this, one can use projections of one plane to another from the tip of the cone).

- To find out more about conic sections see

  V. Prasolov, V. M. Tikhomirov. *Geometry*, (2001). Chapter 4. Conics and Quadrics.Section 4.1. Plane curves of second order. pp.61-69.

This will constitute **Additional 4H reading** and will be examinable for students enrolled to Geometry V (MSc students).

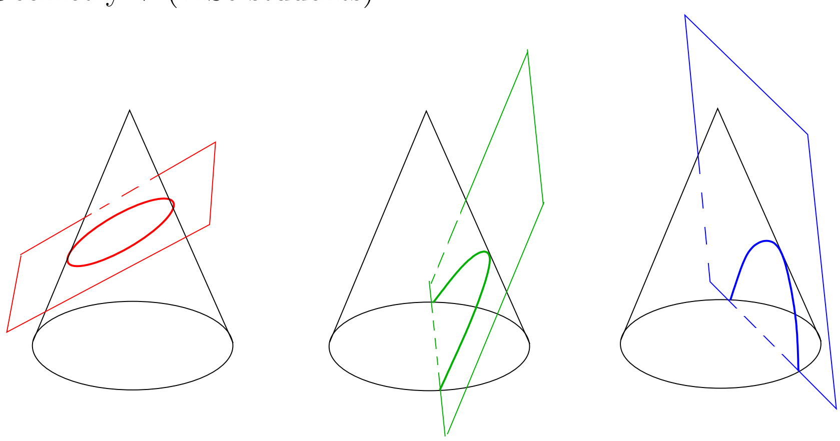


Figure 50: Conic sections: ellipse, parabola and hyperbola.

4.5 Polarity on $\mathbb{RP}^2$ (NE)

(Non-examinable section!)

Consider a trace of a cone $\mathbf{C} = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 = z^2\}$ on the projective plane $\mathbb{RP}^2$ - a conic.

Definition. Points $A = (a_1, a_2, a_3)$ and $B = (b_1, b_2, b_3)$ of $\mathbb{RP}^2$ are called polar with respect to $\mathbf{C}$ if $a_1 b_1 + a_2 b_2 = a_3 b_3$.

Example:

1. Points of $\mathbf{C}$ are self-polar.
2. Point $(2 : 1 : 2)$ is polar to $(1 : 2 : 2)$.

Definition. Given a point $A \in \mathbb{RP}^2$, the set of all points X polar A is the line $a_1 x_1 + a_2 x_2 - a_3 x_3 = 0$, it is called the polar line of A.

Example. Let $A = (0, 0, 1)$ - the North Pole of the sphere, then its polar is the line defined by $x_3 = 0$, i.e. all points with coordinates $(a_1, a_2, 0)$. So, the line $a_1 x_1 + a_2 x_2 = 0$ is the polar line for the point $A = (0, 0, 1)$.

How to find the polar line:

Lemma 4.23. *A tangent line to* $\mathbf{C}$ *at a point* $B = (b_1, b_2, b_3)$ *is* $x_1 b_1 + x_2 b_2 = x_3 b_3$.

We skip the proof of the lemma.

Proposition 4.24. *Let A be a point "outside" $\mathbf{C}$, let l_P and l_Q be tangents to $\mathbf{C}$ at P and Q, where $P, Q \in \mathbf{C}$, s.t. $A = l_P \cap l_Q$. Then PQ is the line polar to A.*

Proof. As $A \in l_P$, we have $a_1 p_1 + a_2 p_2 = a_3 p_3$, so P is polar to A.
As $A \in l_Q$, we have $a_1 q_1 + a_2 q_2 = a_3 q_3$, so Q is polar to A.
Therefore, PQ is the line polar to A, see Fig. 51, left.

□

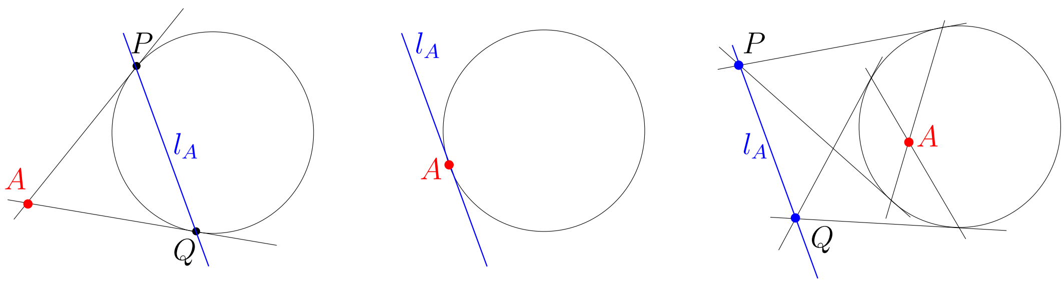


Figure 51: Polar line l_A for a point A inside, on and outside of the conic.

Proposition 4.25. *If $A \in \mathbf{C}$ then the tangent l_A at A is the polar line to A.*

Proof. The line $x_1a_1 + x_2a_2 = x_3a_3$ is tangent at A by Lemma 4.23 and is polar to A by definition of the polar line.

□

Proposition 4.26. *Let A be a point "inside" of the conic* $\mathbf{C}$*. Let p and q be two lines through A. Let P and Q be the points polar to the lines p and q. Then PQ is the line polar to A with respect to* $\mathbf{C}$*.*

Proof. P is polar to A, Q is polar to A, hence, PQ is polar to A, see Fig. 51, right.

□

Remark 4.27.

1. Polarity generalise the notion of orthogonality.
2. More generally, for a conic $\mathbf{C} = \{\mathbf{x} \in R^3 \mid \mathbf{x}^T A\mathbf{x} = 0\}$, where A is a symmetric 3×3 matrix, the point $\mathbf{a}$ is polar to the point $\mathbf{b}$ if $\mathbf{a}^T A\mathbf{b} = 0$.
3. We worked with a diagonal matrix $A = diag\{1, 1, -1\}$.
4. If we take an identity diagonal matrix $A = diag\{1, 1, 1\}$ we get an empty conic $x^2 + y^2 + z^2 = 0$, which gives exactly the same notion of polarity as we had on S^2.
   (Indeed, the point $(a_1 : a_2 : a_3)$ is polar to $a_1x_1 + a_1x_2 + a_3x_3 = 0$ which is the orthogonal plane $(\boldsymbol{x}, \boldsymbol{a}) = 0$).

4.6 Hyperbolic geometry: Klein model

Historic remarks:

- Parallel postulate (or Euclid's Vth postulate) claims that

  > *Given a line l and a point $A \notin l$, there exists a unique line l' such that $l || l'$ and $A \in l'$.*

- For centuries, people tried to derive Euclid's Vth postulate from other postulates.
- In 1870s it turned out that Euclid's Vth postulate is independent of others, i.e. there exists a geometry where
  - all other postulates hold;
  - parallel postulate is substituted by
    "Given a line l and a point $A \notin l$, there exists more than one (infinitely many) line l' such that $l \cap l' = \emptyset$ and $A \in l'$.
- Names:
  - Gauss, Lobachevsky, Bolyai - derived basic theorems of hyperbolic geometry;
  - Beltrami, Cayley, Klein, Poincaré - constructed various models.

More detailed exposition of history can be found in many books, for example in A. B Sossinsky, *Geometries*, Providence, RI : American Mathematical Soc. 2012. One can find the book in the library, see also Chapter 11 (p.119) here.

Klein Model: in interior of unit disc.

- Points of the model are interior points of the unit disc;
- Lines are chords.
- Distance between two points is defined by:

$$d(A,B) = \frac{1}{2}\big|ln[A,B,X,Y]\big|,$$

where

- X, Y are the endpoints of the chord through AB, see Fig. 52, left;
- $[A,B,X,Y] = \frac{|XA|}{|XB|}\Big/\frac{|YA|}{|YB|}$ is the cross-ratio;
- $|PQ|$ denotes the Euclidean length of the segment PQ.

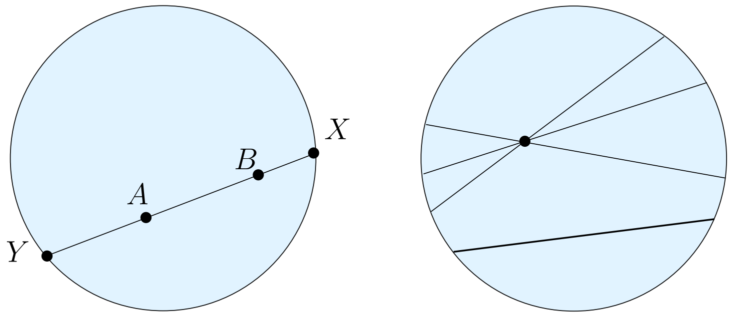


Figure 52: Klein model.

Remark:

1. Axioms of Euclidean geometry are satisfied in the model (except for Parallel Axiom!).
2. Parallel Axiom is obviously not satisfied (see Fig. 52, right):

   Given a line l and a point $A \notin l$, there are infinitely many lines l' s.t. $A \in l$ and $l \cap l' = \emptyset$.

Remark: We will spend a large part of the next term looking at hyperbolic geometry. Our closest aims are to show that

(1) The distance introduced above satisfies axioms of metric;

(2) Isometries act transitively on the points in this model.

Theorem 4.28. *The function $d(A,B)$ satisfies axioms of distance, i.e.*

(1) $d(A,B) \geq 0$ *and* $d(A,B) = 0 \Leftrightarrow A = B$;

(2) $d(A,B) = d(B,A)$;

(3) $d(A,B) + d(B,C) \geq d(A,C)$.

Proof. (1) $d(A,B) \geq 0$ by definition.
Let us show that $d(A,B) = 0$ if and only if $A = B$. Indeed,

$$d(A,B) = 0 \quad \Leftrightarrow \quad ln[A,B,X,Y] = 0 \quad \Leftrightarrow \quad [A,B,X,Y] = 1$$
$$\Leftrightarrow \quad \frac{x-a}{x-b} \Big/ \frac{y-a}{y-b} = 1 \quad \Leftrightarrow \quad \frac{x-a}{x-b} \cdot \frac{y-b}{y-a} = 1,$$

where a, b, x, y are coordinates of the points A, B, X, Y on the line AB.
Notice that $\frac{x-a}{x-b} \geq 1$ and $\frac{y-b}{y-a} \geq 1$, which implies that the product of these numbers equal to 1 if and only if both of them are equal to 1, which is equivalent to the condition $a = b$, i.e. $A = B$.

(2) $d(A,B) = d(B,A)$ since $[A,B,X,Y] = -[B,A,Y,X]$ (which we know from HW 7.8).

(3) We are left to show the triangle inequality $d(A,B) + d(B,C) \geq d(A,C)$, this will be done in Lemma 4.30 below.

□

Remark 4.29. On hyperbolic line:

- Let $[y,x] \in \mathbb{R}$ be an interval. For $a, b \in [y,x]$ (as in Fig. 52, left) we define
$$d(a,b) = \frac{1}{2}\big|\ln[a,b,x,y]\big| = \frac{1}{2}\Big|ln(\frac{x-a}{x-b} / \frac{y-a}{y-b})\Big|.$$
Notice that the logarithm makes sense as the argument is positive for all $a, b \in [y,x]$.
- $d(a,a) = 0$.
- $d(a,b) \to \infty$ when $b \to x$ or $a \to y$.
- Since $ln[a,b,x,y] = -ln[a,b,y,x]$ (as $[a,b,x,y] = 1/\lambda$ when $[a,b,y,x] = \lambda$), we conclude that the endpoints X and Y are "equally good", i.e. the line is not oriented.
- For $c \in [y,x]$ we have $\pm d(a,b) \pm (b,c) \pm d(c,a) = 0$, since
$$\left(\frac{x-a}{x-b} \cdot \frac{y-b}{y-a}\right)\left(\frac{x-b}{x-c} \cdot \frac{y-c}{y-b}\right)\left(\frac{x-c}{x-a} \cdot \frac{y-a}{y-c}\right) = 1.$$
- If $c \in [a,b]$ then $d(a,c) + d(c,b) = d(a,b)$.

Lemma 4.30 (Triangle inequality)**.** *Let A, B, C be three points in Klein model. Then $d(A,B) + d(B,C) \geq d(A,C)$.*

Proof. (1) We start the proof with the following additional construction:

- Extend the sides of the triangle ABC till the boundary of the disc to obtain the chords XY, X_1Y_2 and Y_2X_1 respectively (see Fig. 53, left).

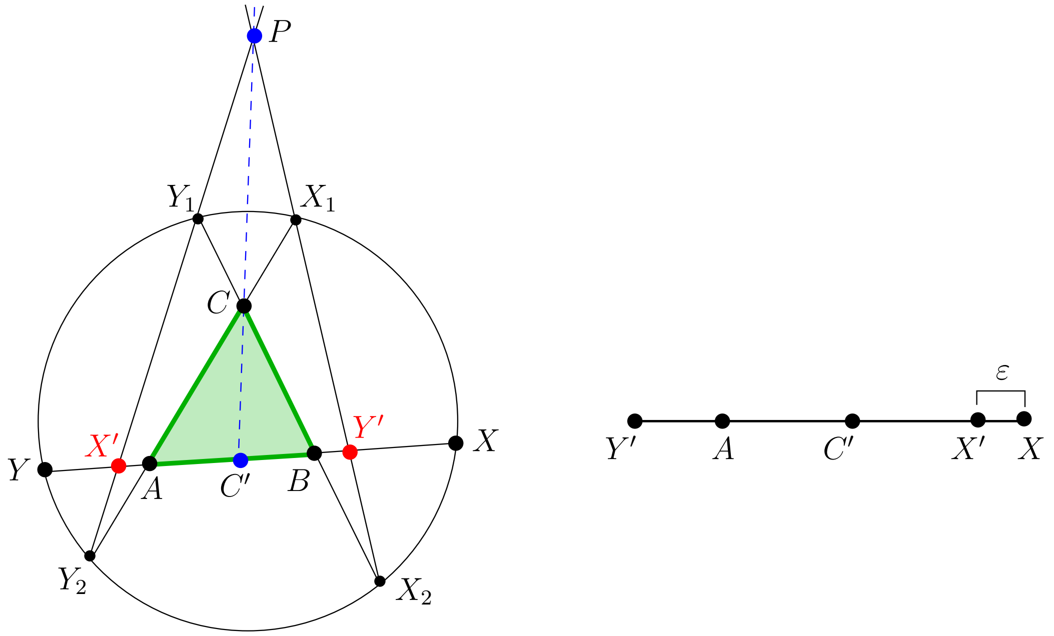


Figure 53: Proof of triangle inequality.

- Define $P := X_1X_2 \cap Y_1Y_2$.
- Define $X' = YX \cap X_1X_2$ and $Y' = YX \cap Y_1Y_2$.
- Define $C' = PS \cap XY \in [AB]$.

(2) Consider the projection from P to the segment XY. As it preserves cross-ratios, we get $[A, C, X_1, Y_2] = [A, C', X', Y']$ and $[C, B, X_2, Y_1] = [C', B, X', Y']$.

(3) <u>Claim:</u> $[A, C', X', Y'] > [A, C', X, Y]$ and $[C', B, X', Y'] > [C', B, X, Y]$.

Proof of the claim. We need to move the endpoints of the segments to the outside of the segment. We will show $[A, C', X', Y'] > [A, C', X, Y']$ and then applying similar movement (i.e. shifting Y to Y') we will get the statement.

Let a, c', x', y', x denote the coordinates of the points A, C', X', Y, X and suppose $x - x' = \varepsilon$, see Fig. 53, right. Then

$$\begin{aligned}
[a, c', x', y'] - [a, c', x, y'] &= \frac{y' - c'}{y' - a}\left(\frac{x' - a}{x' - c'} - \frac{x' - a + \varepsilon}{x' - c' + \varepsilon}\right) \\
&= \frac{y' - c'}{y' - a}\frac{\varepsilon(c' - a)}{(x' - c)(x' - c' + \varepsilon)} > 0,
\end{aligned}$$

which proves the claim.

(4) Finally, we compute:

$$\begin{aligned}
d(A,C)+d(C,B) &\overset{\text{def}}{=} \frac{1}{2}ln[A,C,X_1,Y_2]+\frac{1}{2}ln[C,B,X_2,Y_1]\\
&= \frac{1}{2}ln([A,C,X_1,Y_2]\cdot[C,B,X_2,Y_1])\\
&\overset{(2)}{=} \frac{1}{2}ln([A,C',X',Y']\cdot[C',B,X',Y'])\\
&\overset{(3)}{>} \frac{1}{2}ln([A,C',X,Y]\cdot[C',B,X,Y])\\
&= \frac{1}{2}ln(\frac{x-a}{x-c'}\cdot\frac{y-c'}{y-a}\cdot\frac{x-c'}{x-b}\cdot\frac{y-b}{y-c'})\\
&= \frac{1}{2}ln[a,b,x,y]=d(A,B).
\end{aligned}$$

□

Isometries of Klein model

By isometries we mean transformations of the model preserving the distance, i.e. preserving the disc and the cross-ratios.

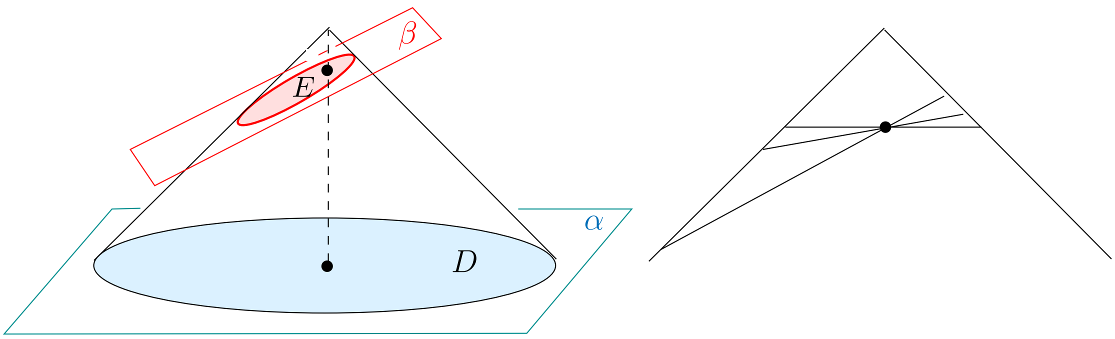


Figure 54: To the proof of Theorem 4.31.

Theorem 4.31. *There exists a projective transformation of the plane that*

- *maps a given disc to itself;*
- *preserves cross-ratios of collinear points;*
- *maps the centre of the disc to an arbitrary inner point of the disc.*

Proof. We will give a sketch of a proof here.

1. Let C be the cone $x^2+y^2=z^2$, let the disc $D=C\cap\alpha$ be the horizontal section of the cone C by a plane α defined by $z=const$.
2. Let β be a plane s.t. $\beta\cap C$ is an ellipse E, see Fig. 54, left.
3. Let $\mathcal{P}$ be the projection of the disc D to the plane β from the apex S of the cone: the projection takes the disc D to the ellipse E, this map is a projective transformation (due to Corollary 4.13).

4. Let $i \in Isom(\mathbb{E}^3)$ be an isometry such that $i(\beta) = \alpha$, suppose also that i takes the centre of the ellipse to the centre of the disc D.

5. Consider an affine transformation $\mathcal{A}$ of the plane α which takes the ellipse $i(E)$ to the disc D.

6. Then the composition $\mathcal{A} \circ i \circ \mathcal{P}$ takes D to D. The map $\mathcal{A} \circ i$ takes the centre of the ellipse to the centre of the disc, while $\mathcal{P}(0)$ lies as far from the centre as we want depending on the choice of plane β, see Fig. 54, right.

7. The map $\mathcal{A} \circ i \circ \mathcal{P}$ is a projective transformation, as it is the composition of a projective transformation, isometry and affine transformation (i.e. of three projective maps), see Fig. 55.

□

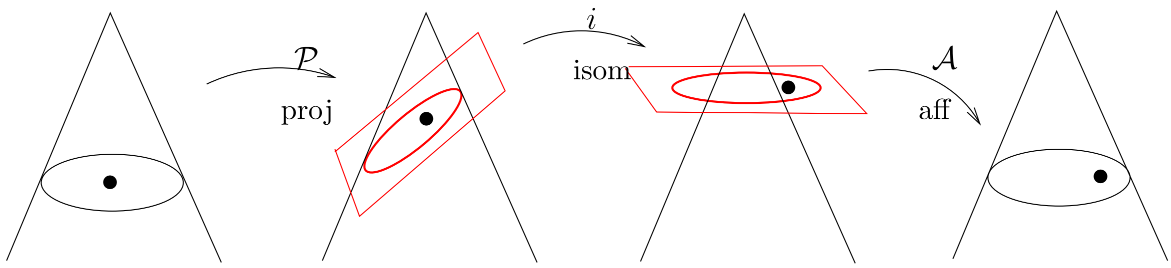


Figure 55: To the proof of Theorem 4.31.

Corollary 4.32.

- *Isometries act transitively on the* <u>*points*</u> *of Klein model.*

- *Isometries act transitively on the* <u>*flags*</u> *in Klein model.*

Proof. The theorem shows transitivity on points. To show transitivity on flags one can:

- map a given point to the centre of the disc;

- then rotate the disc about the centre (it is an isometry in the sense of the model, since it clearly preserves all cross-ratios, and hence preserves the distance).

- reflect the disc (in Euclidean sense) with respect to a line through O (again, it is an isometry as cross-ratios are preserved).

□

Remark.

1. In general, angles in Klein model are not represented by Euclidean angles.

2. Angles at the centre are Euclidean angles.
   Indeed, two orthogonal (in Euclidean sense) chords make equal hyperbolic angles (as one can take one of them to another by an isometry of the hyperbolic plane), so, these angles are $\pi/2$. Similarly, all (Euclidean) angles of size π/n, $n \in \mathbb{Z}$ represent hyperbolic angles of size π/n, and moreover, the angles coincide with Euclidean ones for all π-rational angles. Finally, by continuity we conclude that all angles at the centre of the disc coincide with Euclidean angles.

3. Right angles are shown nicely everywhere in the Klein model (see Proposition 4.33).

Proposition 4.33. *Let l and l' be two intersecting lines in the Klein model. Let t_1 and t_2 be tangent lines to the disc at the endpoints of l. Then* $l \perp l' \Leftrightarrow t_1 \cap t_2 \in \tilde{l}'$, *where $\tilde{l}'$ is the Euclidean line containing the chord representing l'.*

Proof. - We know that at the centre of the disc right angles are shown by two perpendicular diameters l_0 and $l_0^\perp$. Consider the lines p_1, p_2 tangent to the disc at the endpoints of l_0, see Fig. 56, left. Then $l_0^\perp$ is the line through O parallel to the lines p_1, p_2. In other words, $l^\perp$ is the line through O and the intersection $p_1 \cap p_2$ (which does not exist in $\mathbb{E}^2$ but is well-defined in $\mathbb{RP}^2$.

- Let f be a projective transformation which maps the disc to itself, takes l_0 to l and O to $l \cap l'$ (it does exist in view of Corollary 4.32). Notice that f is an isometry of the model (as it preserves the disc and the cross-ratios). Hence, it takes a pair of perpendicular lines to perpendicular (in the sense of hyperbolic geometry) lines.

- Notice that the lines $f(p_1) = t_1$ and $f(p_2) = t_2$ are the tangent lines to the disc at the endpoints of l (indeed, they should contain the endpoints of l but should only have one intersection with the disc, being the images of the tangent lines p_1 and p_2). So, $f(l_0^\perp)$ is the line through $f(O)$ and $f(p_1) \cap f(p_1)$, which exactly means that $l' \perp l$ if and only if it passes through $t_1 \cap t_2$. See Fig. 56, right.

$\square$

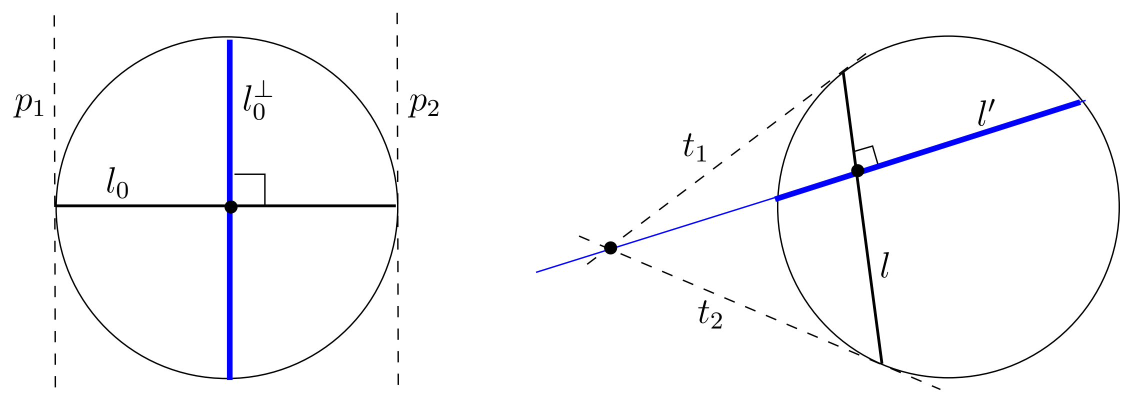


Figure 56: Right angles in the Klein model.

Pairs of lines in hyperbolic geometry: two lines in hyperbolic geometry are called

- intersecting if they have a common point inside hyperbolic plane;
- parallel if they have a common point on the boundary of hyperbolic plane;
- divergent or ultra-parallel otherwise.

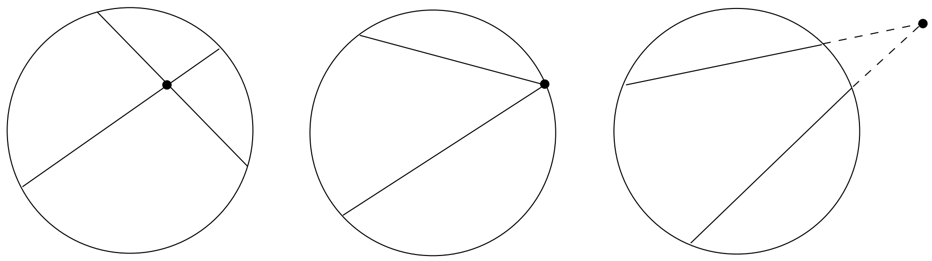


Figure 57: Pairs of lines in the Klein model: intersecting, parallel and ultra-parallel.

Proposition 4.34. *Any pair of divergent lines has a unique common perpendicular.*

Proof. See Fig. 58.

□

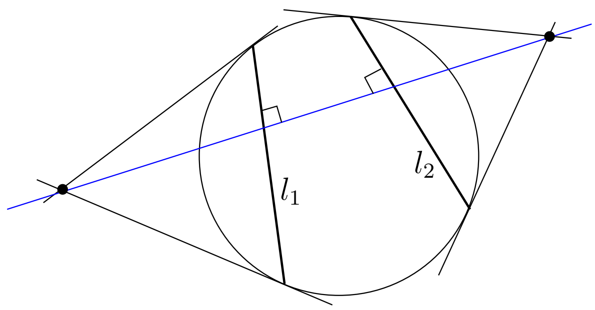


Figure 58: Common perpendicular for any ultra-parallel lines l_1 and l_2.

4.7 References

- Sections 4.1 and 4.2 (on projective line and projective plane) closely follow Lecture II and Lecture III of
  V. V. Prasolov, *Non-Euclidean Geometry*.
  You can find the same material in Section 3.1 of
  V. V. Prasolov, V. M. Tikhomirov, *Geometry*.

- Section 4.3 "Some classical theorems" follows the section on Pappus' and Desargues' theorems in Chapter 3 of
  V. V. Prasolov, V. M. Tikhomirov, *Geometry*.

- Most part of the material of Sections 4.4 and 4.5 (topology of projective plane and polarity on projective plane) may be found in Part II of
  E. Rees, *Notes on Geometry*, Universitext, Springer, 2004.
  (the book is available on DUO in Other Resources).

- Section 4.6 follows Lecture IV of Prasolov's book (or see pp.89-93 in Prasolov, Tikhomirov).

- A very nice overview of projective geometry is provided by
  R. Schwartz, S. Tabachnikov, *Elementary Surprises in Projective Geometry*

- Elliptic geometry is briefly described in
  A. B Sossinsky, *Geometries*, Providence, RI : American Mathematical Soc. 2012.
  One can find the book in the library, see also Section 6.7 (p.75) here.

- A very nice course on projective geometry is
  N. Hitchin, *Projective Geometry*
  (the notes are available on DUO in Other Resources).

- For an overview of history of non-Euclidean geometry see
  A. B Sossinsky, *Geometries*, Providence, RI : American Mathematical Soc. 2012.
  One can find the book in the library, see also Chapter 11 (p.119) here.

- Video:

  - *Why slicing cone gives an ellipse* - video on Grant Sanderson's YouTube channel 3Blue1Brown.

5 Möbius geometry

Hierarchy of geometries

By now, we have considered a number of geometries - Euclidean, spherical, affine, projective, even a bit of hyperbolic. But how are they related to each other?

One answer to this is given by Arthur Caley: "Projective geometry is all geometry." And indeed, as one can see from Fig. 59, Euclidean, affine spherical and the Klein model of hyperbolic geometry are all subgeometries of projective geometry.

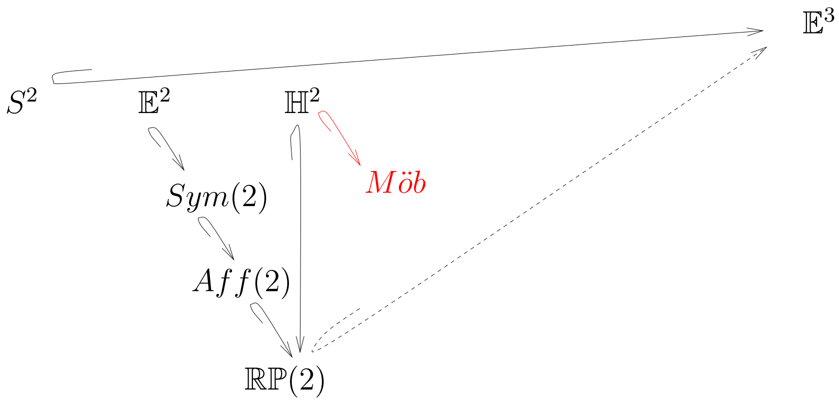


Figure 59: Hierarchy of geometries

At the same time, when hyperbolic geometry is considered in the Klein model, it allows to nicely see the lines, but is not very convenient for working with angles, which are not represented well there. Our first aim now will be to consider Möbius geometry - geometry of linear fractional maps on $\overline{\mathbb{C}}$ which are angle-preserving. This geometry will provide other models for hyperbolic geometry - the models where the lines look more complicated, but the angles are just Euclidean angles.

(And Möbius geometry will not be a part of projective geometry - so, projective geometry is not all geometry after all!).

5.1 Group of Möbius transformations

Definition 5.1. A map $f : \mathbb{C} \cup \{\infty\} \to \mathbb{C} \cup \{\infty\}$ given by $f(z) = \frac{az+b}{cz+d}$, $a, b, c, d \in \mathbb{C}$, $ad - bc \neq 0$ is called a Möbius transformation or a linear-fractional transformation.

Remark. It is a bijection of the Riemann sphere $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ to itself.

Theorem 5.2. *(a) Möbius transformations form a group (denoted Möb) with respect to the composition, this group is isomorphic to*

$$PGL(2, \mathbb{C}) = GL(2, \mathbb{C})/\{\lambda I \mid \lambda \neq 0\}.$$

(b) This group is generated by $z \to \alpha z$, $z \to z + 1$ and $z \to 1/z$, where $\alpha, \beta \in \mathbb{C}$.

Proof. (a) Given a matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL(2, \mathbb{C})$, let $f_A(z) = \frac{az+b}{cz+d}$. In this way we can obtain any Möbius transformation. Moreover, since $\frac{\lambda az+\lambda b}{\lambda cz+\lambda d} = \frac{az+b}{cz+d}$, we may assume that $ad - bc = \pm 1$. Furthermore, we get a bijection between elements of $PGL(2, \mathbb{C})$ and linear-fractional maps. It is straight-forward to check that this bijection respects the group structure, i.e.

$$f_B \circ f_A = f_{BA}.$$

(b) Consider any linear-fractional transformation $f = \frac{az+b}{cz+d}$. We can write

$$f(z) = \frac{az+b}{cz+d} = \frac{\frac{a}{c}(cz+d) + b - \frac{ad}{c}}{cz+d} = \frac{a}{c} + \frac{bc-ad}{c(cz+d)} = f_3 \circ f_2 \circ f_1(z),$$

where

$$f_1(z) = cz + d, \qquad f_2(z) = \frac{1}{z}, \qquad f_3(z) = \frac{bc-ad}{c} z + \frac{a}{c}.$$

Clearly, each of f_1, f_2, f_3 can be obtained as a composition of transformations $z \to \alpha z$, $z \to z + \beta$ and $z \to 1/z$. Furthermore, $z \to z + \beta = \beta(\frac{z}{\beta} + 1)$ is a composition of $z \to \alpha z$ and $z \to z + 1$. So, we conclude that f (and, hence, any linear-fractional transformation) is a composition of $z \to \alpha z$, $z \to z + 1$ and $z \to 1/z$.

□

Example 5.3. The generators az, $z + 1$ and $1/z$ $(a, b \in \mathbb{C})$ can be represented by matrices

$$\begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

respectively.

Theorem 5.4. *(a) Möbius transformations act on $\mathbb{C} \cup \{\infty\}$ triply-transitively.*

(b) A Möbius transformation is uniquely determined by the images of 3 points.

Proof. We need to construct a map $f \in M\ddot{o}b$ taking three given distinct points z_1', z_2', z_3' in $\mathbb{C}\cup\{\infty\}$ to any other three given distinct points z_1, z_2, z_3. We will construct a Möbius transformation $f_0 : (0, 1, \infty) \to (z_1, z_2, z_3)$. Then $f = f_0 \circ g_0^{-1}$, where $g_0 : (0, 1, \infty) \to (z_1', z_2', z_3')$.

<u>Construction:</u> we will construct $f_0 = \frac{az+b}{cz+d}$.

- We will assume that $z_1, z_2, z_3 \neq \infty$, otherwise, we will precompose with $1/(z+d)$.
- $f_0(0) = b/d = z_1$, which is equivalent to $b = z_1 d$.
- $f_0(\infty) = a/c = z_3$, which is equivalent to $a = z_3 c$.
- Hence, $f_0(1) = \frac{z_3 c + z_1 d}{c+d} = z_2$, and we get $c = \frac{(z_2 - z_1)d}{(z_3 - z_2)}$.
- We have obtained a, b, c (all of them proportional to d), so we can cancel d (i.e. assume $d = 1$) to get representative of f_0 which takes $(0, 1, \infty)$ to the required points.

- The constructed map is a Möbius transformation since

$$ad - bc = z_3cd - z_1cd = (z_3 - z_1)\frac{z_2 - z_1}{z_3 - z_2}d^2 \neq 0,$$

(as $z_i \neq z_j$ by assumption). This proves part (a).

Uniqueness of the Möbius transformation $f_0 : (0, 1, \infty) \to (z_1, z_2, z_3)$ follows immediately from the computation. If there are two maps f and h taking $(z_1', z_2', z_3') \to (z_1, z_2, z_3)$ then $f \circ g_0$ and $h \circ g_0$ are two maps taking $(0, 1, \infty) \to (z_1, z_2, z_3)$, which is impossible. This implies part (b).

□

Theorem 5.5. *Möbius transformations*

(a) take lines and circles to lines and circles;

(b) preserve angles between curves.

Proof. It is sufficient to check the statements for the generators:

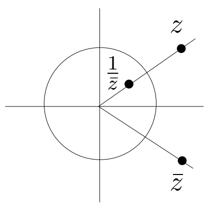
- $z \to az$: is a rotation about 0 by argument of a composed with a dilation by $|a|$;
- $z \to z + 1$: translation by 1;
- $z \to 1/z$: composition of a reflection $z \to \bar{z}$ and an inversion $z \to 1/\bar{z}$ (see Fig. 60, left for the action of $z \to 1/\bar{z}$).

All these transformations satisfy (a) and (b) (for $z \to 1/z$ recall the results from Complex Analysis II - we will also show it independently below in Theorems 5.14 and 5.15).

□

Example.

1. See Fig. 60, left, for the action of $z \to 1/\bar{z}$.
2. Transformation $z \to 1/z$ takes the real line to itself and the circle $(z - \frac{1}{2})^2 = (\frac{1}{2})^2$ to the line $Re\ z = 1$. The right angle between these two curves is preserved (see Fig. 60, right).



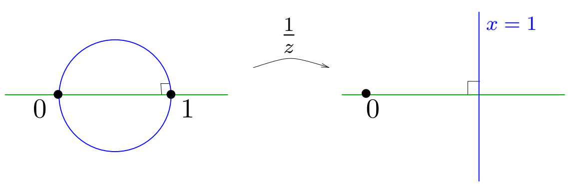


Figure 60: Left: Transformation $z \to 1/\bar{z}$. Right: an angle preserved by $z \to 1/z$.

5.2 Types of Möbius transformations

Consider the fixed points of the transformation $f(z) = \frac{az+b}{cz+d}$, i.e., the points satisfying

$$z = \frac{az+b}{cz+d}.$$

This is a quadratic equation with respect to z, so it has exactly two complex roots (these roots may coincide, in which case f has a unique fixed point).

Definition 5.6. A Möbius transformation with a unique fixed point is called parabolic.

Example: $z \to z + b$, where $b \in \mathbb{C}^*$ is a parabolic transformation (with a unique fixed point ∞).

Proposition 5.7. *Every parabolic Möbius transformation is conjugate in the group* $M\ddot{o}b$ *to* $z \to z + 1$.

Proof. Suppose that f is a parabolic transformation with $z_0 = f(z_0)$.

Let $g(z) = \frac{1}{z-z_0}$, notice that $g(z_0) = \infty$. Then the transformation

$$f_1(z) := g \circ f \circ g^{-1}(z)$$

has a unique fixed point at ∞ (here we use the same reasoning as in Proposition 1.18(a)). This implies that $f_1(z) = \frac{az+b}{cz+d}$ with $c = 0$ (as $f_1(\infty) = \infty$). By scaling a and b we may assume $f_1(z) = az + b$. Since f_1 has a (double) root at infinity (and no other roots) we see that the equation $z = az + b$, has the only solution $z = -\frac{b}{a-1}$ at infinity, which is only possible when $a = 1$. We conclude that $f_1 = z + b$, so f is conjugate to $z \to z + b$.

Finally, let $h(z) = bz$. Then

$$f_2(z) := h^{-1} \circ f_1(z) \circ h(z) = \frac{1}{b}(bz + b) = z + 1,$$

So, we conclude that f is conjugate to $z \to z + 1$.

□

Proposition 5.8. *Every non-parabolic Möbius transformation is conjugate in* $M\ddot{o}b$ *to* $z \to az$, $a \in \mathbb{C} \setminus \{0\}$.

Proof. Let z_1, z_2 be the fixed points of a Möbius transformation f. The transformation $g(z) = \frac{z-z_1}{z-z_2}$ sends them to 0 and ∞. So, $f_1(z) = gfg^{-1}(z)$ has fixed points at $0, \infty$. Hence, $f_1(z) = az$, and we see that f is conjugate to $z \to az$.

□

Definition 5.9. A non-parabolic Möbius transformation conjugate to $z \to az$ is called

(1) elliptic, if $|a| = 1$;

(2) hyperbolic, if $|a| \neq 1$ and $a \in \mathbb{R}$;

(3) loxodromic, otherwise.

Remark 5.10. Consider the dynamics of various types of elements when they are iterated many times (see Fig. 61). We draw each type twice: in the first row all fixed points a visible, while in the second row one fixed point is mapped to ∞ (but the picture is more simple).

Parabolic elements are best understood when the fixed point is ∞ - then it is translation of all points by the same vector. Applying a Möbius transformation we see that iterations of such a transformation move points a long circles through the fixed point.

All other elements are best viewed when the fixed points are 0 and ∞: elliptic elements just rotate points around two equally good fixed points, while hyperbolic and loxodromic elements have one one attracting fixpoint and one repelling.

Two fixpoints of a hyperbolic or a loxodromic transformation have different properties: one is attracting another is repelling.

Elliptic transformations have two similar fixpoints (neither attracting nor repelling).

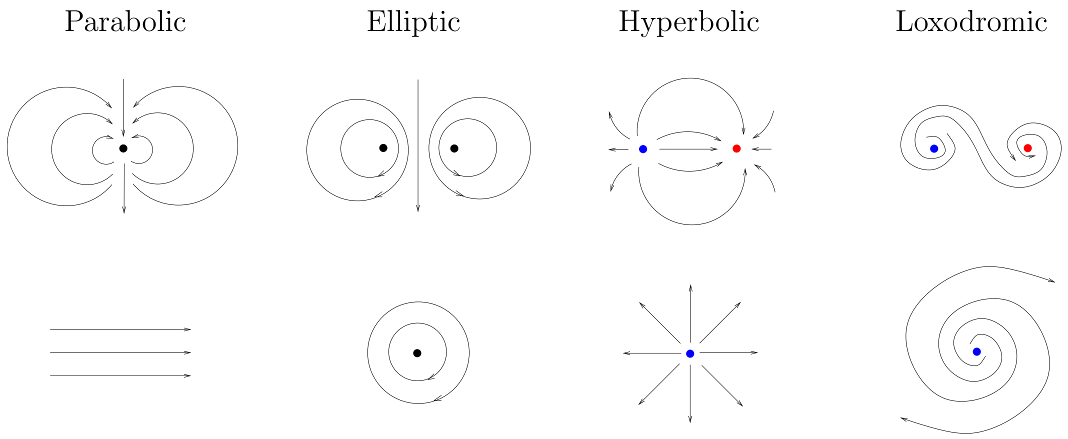


Figure 61: Dynamics of parabolic, elliptic, hyperbolic and loxodromic elements.

Dynamics of Möbius transformations is nicely illustrated in the 2-minute video by Douglas Arnold and Jonathan Rogness.

5.3 Inversion

Definition 5.11. Let $\gamma \in \mathbb{C}$ be a circle with centre O and radius r. An inversion I_γ with respect to γ takes a point A to a point A' lying on the ray OA s.t. $|OA| \cdot |OA'| = r^2$, see Fig. 62.

Proposition 5.12. *(a)* $I_\gamma^2 = id$.

(b) Inversion in γ *preserves* γ *pointwise* ($I_\gamma(A) = A$ *for all* $A \in \gamma$).

Proof. This immediately follows from the definition. □

Lemma 5.13. *If* $P' = I_\gamma(P)$ *and* $Q' = I_\gamma(Q)$ *then* $\triangle OPQ$ *is similar to* $\triangle OQ'P'$.

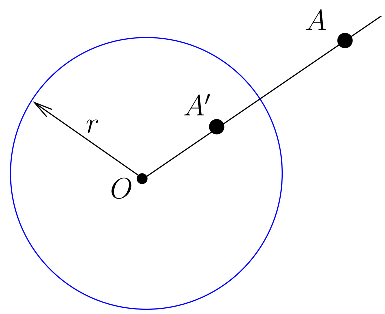


Figure 62: Inversion: $|OA| \cdot |OA'| = r^2$.

Proof. Since $|OP| \cdot |OP'| = r^2 = |OQ| \cdot |OQ'|$ we have

$$\frac{|OP|}{|OQ|} = \frac{|OQ'|}{|OP'|},$$

see Fig. 63. As $\angle POQ = \angle P'OQ'$, we conclude that $\triangle POQ \sim \triangle Q'OP'$ (by sAs). $\square$

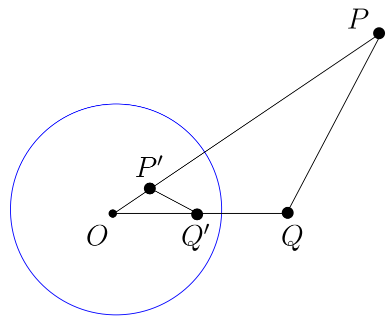


Figure 63: Inversion: $\triangle OPQ \sim \triangle OQ'P'$.

Theorem 5.14. *Inversion takes circles and lines to circles and lines. More precisely,*

1. *lines through O are mapped to lines through O;*
2. *lines not through O are mapped to circles through O*
3. *circles not through O are mapped to circles not through O.*

Proof. Consider an inversion I_γ with respect to a circle γ.

1. This part is evident from the definition.
2. Let l be a line, $O \notin l$. Let $Q \in l$ be a point such that $OQ \perp l$, see Fig. 64. Let $P \in l$ be any point of l and let $P' = I\gamma(P)$, $Q' = I_\gamma(Q)$.

   By Lemma 5.13, $\triangle POQ \sim \triangle Q'OP'$, so $\angle OP'Q' = \pi/2$. This implies that P' lies on the circle with diameter OQ' (by converse of E26). This implies that $I_\gamma(l)$ is the circle with the diameter OQ'.

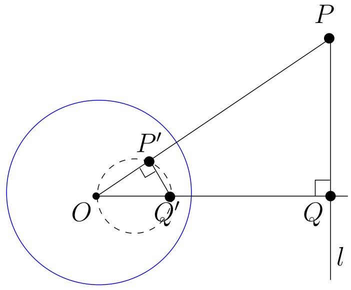


Figure 64: Inversion takes lines not through origin to the circles through origin.

3. Let γ_0 be a circle $O \notin \gamma_0$. Let l be a line through O and the centre of γ_0. Let $\{P, Q\} = l \cap \gamma_0$, $R \in \gamma_0$, and let I_γ takes the points P, Q, R to P', Q', R' respectively, see Fig. 65.

   Be Lemma 5.13, we have $\angle OPR = \angle OR'P'$ which implies $\angle RPQ = \angle P'R'R$. Also, we have $\angle OQR = \angle OR'Q'$. Since PQ is the diameter of γ_0, we have $\angle PRQ = \pi/2$, which implies that $\angle RPQ + \angle OQR = \pi/2$. Therefore, $\angle Q'R'P' = \pi/2$, and hence, R' lies on the circle with diameter $Q'R'$.

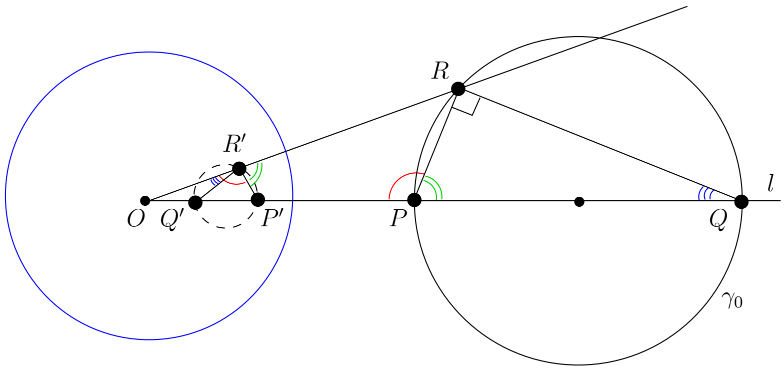


Figure 65: Inversion takes circles not through origin to the circles not through origin.

$\square$

See Inversion Tool on Cut-The-Knot portal for hands-on illustration of Theorem 5.14.

Theorem 5.15. *Inversion preserves angles.*

Proof. Let I_γ be the inversion with respect to the circle γ. Let l be a line such that $O \notin l$, see Fig. 66. Then $I_\gamma(l)$ is a circle γ through O and the tangent line to γ line at the point O is parallel to l (one can see it for example from the symmetry with respect to the line orthogonal to l dropped from O). This implies that if l_1, l_2 are two lines not through the origin, then the angle between them is preserved by the inversion.

For two circles (or a line and a circle) we measure the angles between tangent lines to them (and this angle is preserved as shown above).

If one or both of l_1, l_2 pass through O then it the image of such line is still parallel to initial line, so the angle is still preserved.

$\square$

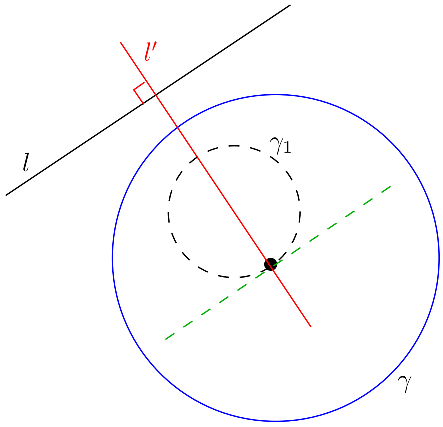


Figure 66: Inversion preserves angles.

Remark. Inversion may be understood as "reflection with respect to a circle".

Example 5.16. Let I_1 be inversion with respect to the unit circle centred at the origin, and $I_{\sqrt{2}}$ be the inversion with respect to the circle of radius $\sqrt{2}$ centred at $-i$, see Fig. 67. Notice that $I_{\sqrt{2}}$ takes the unit circle to the real line.

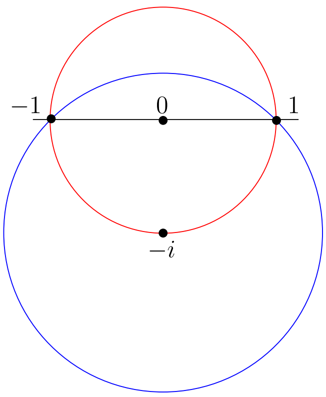


Figure 67: Reflection is a conjugated inversion.

Define $r := I_{\sqrt{2}} I_1 I_{\sqrt{2}}$. Then $r(x) = x$ for every $x \in \mathbb{R}$, and it is easy to see that r swaps the half-planes defined by the real line.

As r is a composition of inversions, it preserves the angles, which (together with preserving all points of real line) implies that r is a reflection, see Fig. 68.

Theorem 5.17. *Every inversion is conjugate to a reflection by another inversion.*

Proof. As in Example 5.16, given an inversion I_γ with respect to a circle γ, consider an inversion I with respect to a circle forming angle $\pi/4$ with γ: then $I \circ I_\gamma \circ I$ is a reflection.

□

Theorem 5.18. *Every Möbius transformation is a composition of even number of inversions and reflections.*

Proof. By Theorem 5.2 Every Möbius transformation is a composition of transformations az, $z+1$, $1/z$. We will check that each of these transformations is a composition of inversions and reflections.

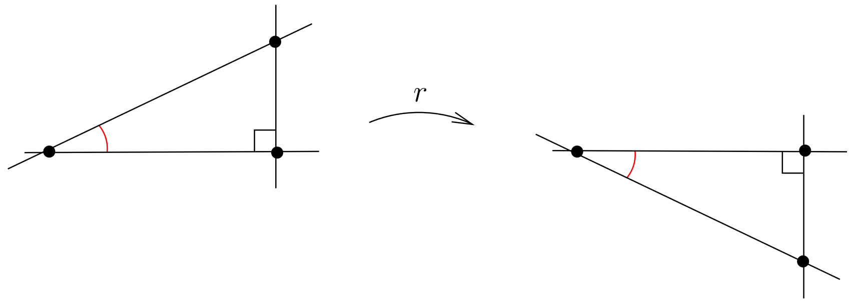


Figure 68: r preserves all real points and preserves angles, hence r is a reflection.

- $az = |a|e^{i\,\mathrm{Arg}\,a}z$ is a composition of a dilation and a rotation. A rotation by angle α is a composition of reflections with respect to the lines meeting at angle $\alpha/2$ (by Example 1.13). Dilation $D : z \to r^2 z$ $(r \in \mathbb{R})$ is a composition of two inversions I_1 and I_r with respect to circles of radius 1 and r centred at the origin:

$$I_r \circ I_1(z) = I_r(\frac{1}{\bar{z}}) = r^2 z = D(z).$$

- Translation $z + 1$ is a composition of two reflections (again by Example 1.13).

- The map $f : z \to 1/z$ is a composition of an inversion $1/\bar{z}$ and a reflection $\bar{z}$: $f = \bar{z} \circ \frac{1}{\bar{z}}$.

Notice that we described each of the generators as a composition of even number of reflections and inversions.

□

Remark 5.19. Inversion and inversion change orientation of the plane, but Theorem 5.18 says that a Moöbius transformation is expressed through even number of them. Hence, it shows that Möbius transformations preserve orientation.

See here for an animation demonstrating properties of inversion (by M. Christersson).

5.4 Möbius transformations and cross-ratios

Definition 5.20. For $z_1, z_2, z_3, z_4 \in \mathbb{C} \cup \{\infty\}$, the complex number

$$[z_1, z_2, z_3, z_4] = \frac{z_3 - z_1}{z_3 - z_2} / \frac{z_4 - z_1}{z_4 - z_2} \in \mathbb{C} \cup \{\infty\}$$

is called the cross-ratio.

Theorem 5.21. *Möbius transformations preserve cross-ratios.*

Proof. This is an easy computation for each of the generators az, $z + 1$, $1/z$ (check!).

□

Corollary 5.22. *A Möbius transformation is determined by images of 3 points.*

Proof. If $f \in M\ddot{o}b$, $f : a, b, c \to a', b', c'$ and $y = f(x)$, then y can be computed from the linear equation $[a, b, c, x] = [a', b', c', y]$. □

Remark 5.23. Points $z_1, z_2, z_3 \in \mathbb{C}$ are collinear if and only if $\frac{z_1-z_2}{z_1-z_3} \in \mathbb{R}$ (i,e. when vectors $z_1 - z_2$ and $z_1 - z_3$ are proportional over $\mathbb{R}$).

Proposition 5.24. *Points $z_1, z_2, z_3, z_4 \in \mathbb{C} \cup \{\infty\}$ lie on one line or circle if and only if $[z_1, z_2, z_3, z_4] \in \mathbb{R}$.*

Proof. By Theorem 5.4 there exists a Möbius transformation f which takes z_1, z_2, z_3 to $0, 1, \infty$. Let $x \in \mathbb{C}$ and $y = f(x)$. It is easy to see (using Remark 5.23) that y lies on a real line if and only if $[0, 1, \infty, y]$ is real. Hence, x lies on the same line or circle as z_1, z_2, z_3 if and only if $[z_1, z_2, z_3, z_4] \in \mathbb{R}$.

□

Remark 5.25. Geometric proof of Proposition 5.24:
Consider 4 points on the same circle. By E28, $\angle z_1 z_4 z_2 = \angle z_1 z_4 z_3$, which means that $\text{Arg}(\frac{z_3-z_1}{z_3-z_2}) = \text{Arg}(\frac{z_4-z_1}{z_4-z_2})$. This implies that $[z_1, z_2, z_3, z_4] = \frac{z_3-z_1}{z_3-z_2} \big/ \frac{z_3-z_1}{z_3-z_2} \in \mathbb{R}$. Conversely, if z_4 does not lie on the same circle as z_1, z_2, z_3, then the angles at z_3 and z_4 are different and the cross-ratio is not real.

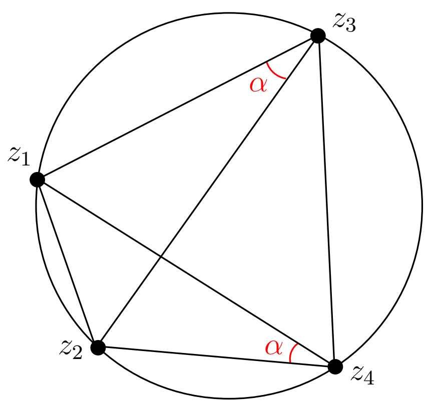


Figure 69: Geometric meaning of real cross-ratio.

Proposition 5.26. *Given four distinct points $z_1, \dots, z_4 \in \mathbb{C} \cup \infty$, one has*

$$[z_1, z_2, z_3, z_4] \neq 1.$$

Proof. Suppose that $[z_1, z_2, z_3, z_4] = 1$. Then by Proposition 5.24 the points lie on one line or circle, so we may assume that $[z_1, z_2, z_3, z_4] = [x, 0, 1, \infty]$, where $x \in \mathbb{R}$ (here we use triple transitivity of $M\ddot{o}b$). So, $[z_1, z_2, z_3, z_4] = \frac{1-x}{1-0} \big/ \frac{\infty-x}{\infty-0} = 1 - x$, this only equals to 1 when $x = 0$, which is impossible as the points z_1, z_2, z_3, z_4 (and hence, the points $x, 0, 1, \infty$) are distinct by assumption.

□

Example 5.27. (a) Two parallel lines are not $M\ddot{o}b$-equivalent to two concentric circles (as circles are disjoint while lines are tangent at ∞, i.e. sharing one point).

(b) Let l_x be a line given by $Re(z) = x$, $x \in \mathbb{R}$. Is there a Möbius transformation taking l_0, l_1, l_2 to l_0, l_1, l_3?

To answer the question consider a line or circle γ orthogonal to all three of l_0, l_1, l_2. It is easy to see that γ is a line orthogonal to l_i (justify this!). Let A, B, C, D be the points where γ intersects respectively l_0, l_1, l_2 (where $D = \infty$). Then

$$[A, B, C, D] = [0, 1, 2, \infty] = \frac{2-0}{2-1}\Big/\frac{\infty - 0}{\infty - 1} = 2/1 = 2.$$

Similarly, let A', B', C', D' be the points where γ intersects respectively l_0, l_1, l_3 (where $D' = \infty$). Then

$$[A', B', C', D'] = [0, 1, 3, \infty] = \frac{3-0}{3-1}\Big/\frac{\infty - 0}{\infty - 1} = \frac{3}{2}.$$

As for $\lambda = 2$ none of $\lambda, 1-\lambda, \frac{1}{\lambda}, \frac{1}{1-\lambda}, \frac{-\lambda}{1-\lambda}, \frac{1-\lambda}{-\lambda}$ coincides with $\frac{3}{2}$, we conclude that there is no Möbius transformation taking l_0, l_1, l_2 to l_0, l_1, l_3.

Remark 5.28. Does reflection/inversion preserve cross-ratio?
Example: reflection $z \to \bar{z}$ takes the cross-ratio $[z_1, z_2, z_3, z_4]$ to $\overline{[z_1, z_2, z_3, z_4]}$.

The same will happen for every inversion/reflection f: we can find a Möbius transformation g which takes Fix_f to $\mathbb{R} \cup \{\infty\}$, then $f = g \circ \bar{z} \circ g^{-1}$, where g preserves cross-ratios, and $\bar{z}$ conjugates.

We conclude that inversions and reflections take cross-ratios to conjugate numbers.

Corollary 5.29. *Cross-ratios of four points lying on a line or a circle are preserved by inversions and reflections.*

5.5 Inversion in space

Definition 5.30. Let $S \in \mathbb{R}^3$ be a sphere of radius r centred at O. The inversion I_S with respect to S takes a point $A \in \mathbb{R}^3$ to a point A' on the ray OA s.t. $|OA| \cdot |OA'| = r^2$, see Fig.70.

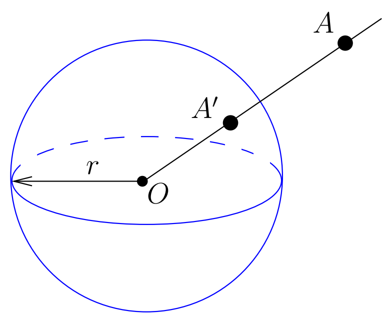


Figure 70: Inversion in space: $|OA| \cdot |OA'| = r^2$.

Theorem 5.31 (Properties of inversion)**.**

(1) Inversion takes spheres and planes to spheres and planes.

(2) Inversion takes lines and circles to lines and circles.

(3) Inversion preserves angles between curves.

(4) Inversion preserves cross-ratio of four points $[A, B, C, D] = \frac{|CA|}{|CB|}\Big/\frac{|DA|}{|DB|}$.

Proof. Let I be inversion in the sphere S centred at O of radius r.

(1) Let l be a line through the centre O of S orthogonal to a plane or sphere α, see Fig. 71. Let Π be any plane containing l. Restricting I to Π we obtain an inversion on the plane Π. So the circle or line $\Pi \cap \alpha$ will be mapped to some circle of line q. Now, notice that all data (including the definition of inversion, the sphere S, the line l and the set α) are preserved by rotation around l. So, if we rotate Π, we will see exactly the same image of $\Pi \cap \alpha$ as in Π (but rotated). Rotating a circle or a line around the line l through the centre of the circle (or a point on the line) we get a sphere or a plane.

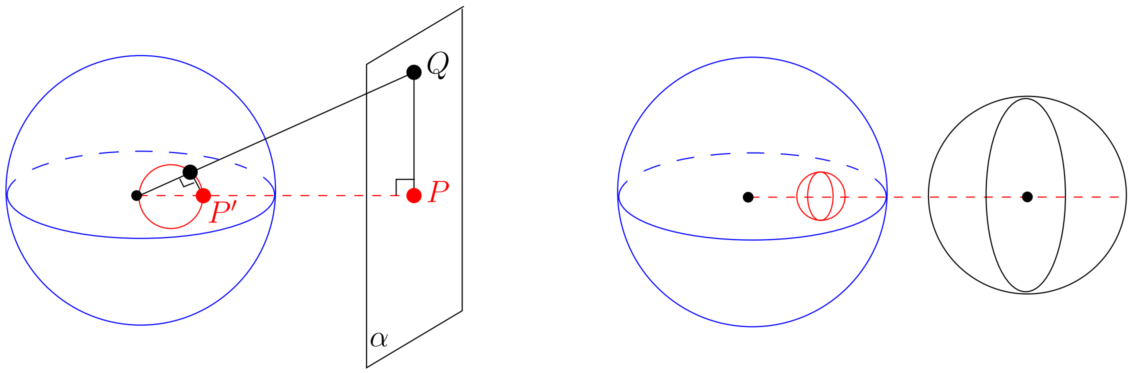


Figure 71: Inversion in space: rotating the circle get a sphere.

(2) A line is an intersection of two planes, a circle is an intersection of a plane and a circle. Applying the result of part (1) we see that the image of a line of a circle is a line or a circle.

(3) Similarly to 2-dimensional case, a line l is mapped by an inversion to a circle through O with a tangent line at O parallel to l.

(4) Let $P' = I(P)$ and $Q' = I(Q)$. Then in the same way as in 2-dimensional case we see that $\triangle OPQ \sim \triangle OQ'P'$, one can see it by considering the restriction of I to to the plane OPQ (or to any plane containing O, P, Q if they are collinear). Therefore,
$$\frac{|PQ|}{|P'Q'|} = \frac{|OP|}{|OQ'|} = \frac{|OP|}{|OQ|} \cdot \frac{|OQ|}{|OQ'|} = \frac{|OP| \cdot |OQ|}{r^2}.$$
This implies that
$$\frac{[A,B,C,D]}{[A',B',C',D']} = \frac{\frac{|CA|}{|C'A'|} \cdot \frac{|DB|}{|D'B'|}}{\frac{|CB|}{|C'B'|} \cdot \frac{|DA|}{|D'A'|}} = \frac{\frac{|OA|\cdot|OB|\cdot|OC|\cdot|OD|}{r^4}}{\frac{|OA|\cdot|OB|\cdot|OC|\cdot|OD|}{r^4}} = 1.$$

□

5.6 Stereographic projection

Definition 5.32. Let S be a sphere centred at O, let α be a plane through O. Let $N \in S$ be a point with $NO \perp \alpha$. The map $\pi : S \to \alpha$ s.t. $\pi(A) = \alpha \cap NA$ for all $A \in S$ is called a stereographic projection, see Fig. 72, left.

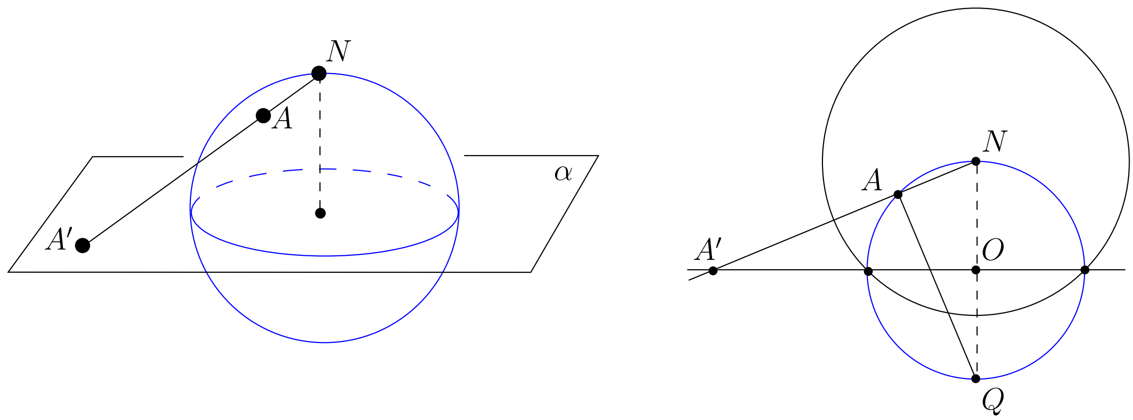


Figure 72: Stereographic projection: definition (left) and as a restriction of inversion (right).

Proposition 5.33 (Properties of stereographic projection)**.**

(a) Stereographic projection takes circles to circles and lines.

(b) Stereographic projection preserves angles.

(c) Stereographic projection preserves cross-ratio.

Proof. Properties (a)-(c) listed in the proposition are properties (2)-(4) of inversion with respect to sphere. Hence, it is sufficient to show that the stereographic projection is a restriction of some inversion to the sphere.

For the unit sphere S centred at the origin consider the sphere $S_{\sqrt{2}}$ of radius $\sqrt{2}$ centred a the point N, see Fig. 72, right. Notice that $S_{\sqrt{2}}$ makes angle $\pi/4$ both with the sphere S and the plane α. This means that the inversion $I_{S_{\sqrt{2}}}$ maps S to α, which implies that $I_{S_{\sqrt{2}}}(A) = A' = NA \cap \alpha$. □

Remark. Another argument to see that $I_{S_{\sqrt{2}}}$ maps A to A' is based on similarity of triangles $\triangle A'ON \sim \triangle QAN$, where $Q = S \cap ON$ (the triangles are similar by AAA), see Fig. 72,right. Then

$$\frac{|A'N|}{2|ON|} = \frac{|ON|}{|NA|},$$

which implies that $|A'N| \cdot |NA| = 2|ON|^2 = (\sqrt{2}|ON|)^2$.

Remark. We used stereographic projection when proved that the formula for the area of spherical triangle $S_\triangle = \alpha + \beta + \gamma - \pi$, see Fig. 28.

Remark. Another way to define stereographic projection, is to project from N to the plane α tangent to S at point opposite to N; see Fig 74, left. This projection has the same properties as the one in Definition 5.32. To see, that the properties of the new projection are the same, just notice that the images on the two planes only differ by a scaling (projection from N).

Remark. This version of stereographic projection is nicely illustrated in a 1-minute video by Henry Segerman.

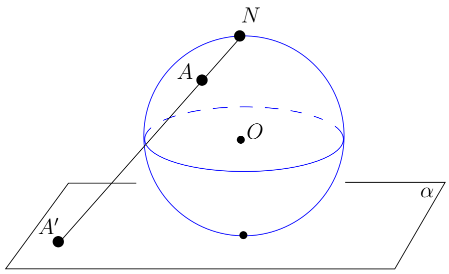


Figure 73: Another way to define stereographic projection.

Remark. Why do we need all these properties of 3-dimensional inversions and stereographic projection? We used 3-dimensional inversion to show the properties of the stereographic projection, and the later will be used to show that different models of $\mathbb{H}^2$ give rise to the same geometry.

Example 5.34 (Steiner Porism)**.** A circle γ_1 lies inside another circle γ_2. A circle $\mathcal{C}_0$ is tangent to both γ_1 and γ_2. A circle $\mathcal{C}_i$ is tangent to three circles: γ_1, γ_2 and $\mathcal{C}_{i-1}$, for $i = 1, 2, 3 \dots$. It may happen that either all circles $\mathcal{C}_i$, $i \in \mathbb{N}$ are different, or $\mathcal{C}_n = \mathcal{C}_1$ for some n. Show that the outcome does not depend on the choice of the initial circle $\mathcal{C}_0$ (but only depends on γ_1 and γ_2).

Proof. First, we need to show that every two disjoint circles are Möbius-equivalent to two concentric circles. This can be done using an appropriate sequence of inversions. We will skip the proof here, as we will see another, shorter explanation of this later based on hyperbolic geometry (see Example 6.18 below).

Once the circles γ_1 and γ_2 are mapped to concentric circles, the statement follows trivially (as any choice of initial circle $\mathcal{C}_0$ may be transformed to any other choice by a rotation around the common centre of the concentric circles).

□

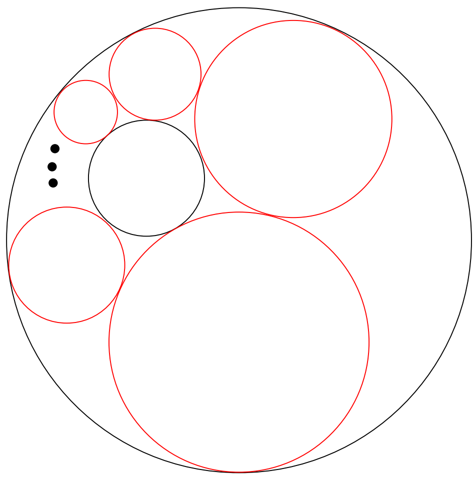


Figure 74: Steiner Porism.

5.7 References

- One can read about hierarchy of geometries (with more examples than we had at the start of Section 5) in
  A. B Sossinsky, *Geometries*, Providence, RI : American Mathematical Soc. 2012.
  One can find the book in the library, see also Section 1.4 (pp.119–124) here.

- Our exposition of Möbius transformations, Inversion and Stereographic projection is expansion of Lecture V in
  V. V. Prasolov, *Non-Euclidean Geometry*.
  You can find the same material on pp.93-95 of
  V. V. Prasolov, V. M. Tikhomirov, *Geometry*.

- Another exposition about Möbius transformations is contained it Section 1.1 of *Hyperbolic Geometry* by Caroline Series.

- For introduction and discussion of inversion see the following sources

  - Malin Christersson, *Circle inversion* (Illustrated introduction with proofs).
  - Tom Davis, *Inversion in a circle*. (An article showing how to use inversions (but not giving the proofs of basic properties of inversion).

- I have borrowed the "Proof without words" for Ptolemy theorem (you can find it in the Problems Classes notes) from the cut-the-knot portal. Which in its turn refers to the following paper:

  - W. Derrick, J. Herstein, *Proof Without Words: Ptolemy's Theorem*, The College Mathematics Journal, v. 43, n. 5, November 2012, p 386.

- Animations:

  - Inversion Tool, hands-on demonstration of inversion on cut-the-knot partal.
  - Loxodromic transformation in the page by Paul Nylander. Also, scroll down to find the animation of the coloured version.

- Videos:

  - Dynamics of Möbius transformations is illustrated in the 2-minute video by D. Arnold and J. Rogness.
  - Animation demonstrating Inversion in circles, by M. Christersson.
  - 1-minute video illustrating stereographic projection by Henry Segerman.

6 Hyperbolic geometry: conformal models

6.1 Poincaré disc model

Model: $\mathbb{H}^2$ =unit disc $D = \{|z| < 1, z \in \mathbb{C}\}$;
$\partial\mathbb{H}^2 = \{|z| = 1\}$, boundary, called absolute;
- lines: parts of circles or lines orthogonal to $\partial\mathbb{H}^2$, see Fig. 75, left;
- distance: a function of cross-ratio;
- angles: same as Euclidean angles.

Group: Isometries (i.e. Möbius transformation, inversions, reflections - preserving the disc).

Notice that these transformations indeed preserve distance (when distance is a function of cross-ratio), angles, set of lines.

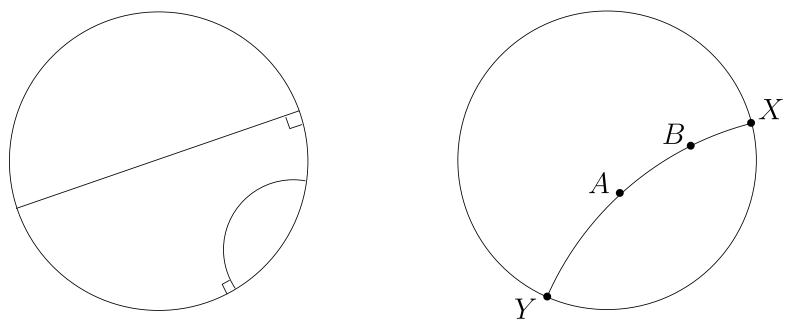


Figure 75: Poincaré disc model.

Proposition 6.1. *For any two points $A, B, \in \mathbb{H}^2$ there exists a unique hyperbolic line through A, B.*

Proof. Let I be the inversion with respect the absolute. Consider $I(A)$. Let γ be the (Euclidean) circle/line through $A, B, I(A)$ (it does exist and is unique as every Euclidean triangle has a unique circumscribed circle), see Fig. 76, left. Let $X = \gamma \cap \partial\mathbb{H}^2$ (exists as A is inside absolute and $I(A)$ is outside). Then $I(\gamma) = \gamma$ (as it swaps A with $I(A)$ and preserves X). This implies tat $\gamma \perp \partial\mathbb{H}$, and hence, $\gamma \cap D$ is the hyperbolic line through A, B.

Notice that any (Euclidean) line/circle containing a hyperbolic line through A, B should be preserved by inversion I (as the Euclidean line/circle should be orthogonal to $\partial\mathbb{H}^2$), so, it should contain $A, I(A), B$, and hence, coincide with γ.

$\square$

Remark 6.2. The same holds for $A, B, \in \mathbb{H}^2 \cup \partial\mathbb{H}^2$. Indeed, if one of the points A, B (say, A) is not lying on the absolute, we can use the same proof as before. If $A, B, \in \partial\mathbb{H}^2$, then any (hyperbolic) line through A, B should be orthogonal to the absolute at A and B, or, in other words, orthogonal to the tangent lines l_A and l_B to the absolute at points A, B, so the point $Q = l_A \cap l_B$ is the (Euclidean) centre of the corresponding (Euclidean) circle, see Fig. 76, right. If $A, B \in \partial\mathbb{H}^2$ are diametrically opposed points of the absolute, then the same reasoning shows that the diameter AB is the unique hyperbolic line through A and B.

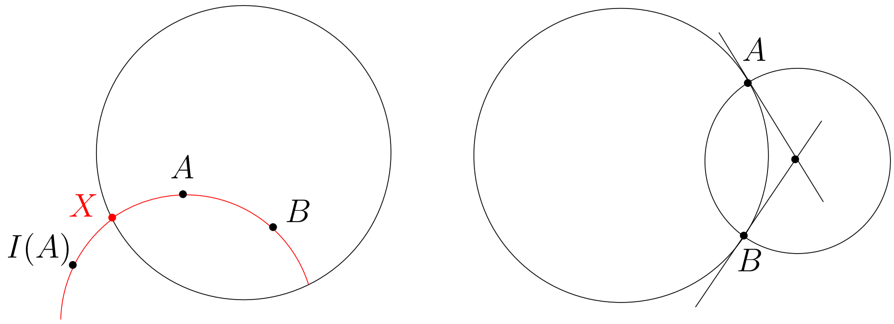


Figure 76: Construction of a line through given points A and B.

Definition 6.3. $d(A,B) = \big|ln|[A,B,X,Y]|\big| = \big|ln\frac{|XA|}{|XB|}/\frac{|YA|}{|YB|}\big|$,
where X, Y are the points of the absolute contained in the (hyperbolic) line AB, see Fig. 75, right.

Theorem 6.4. *$d(A,B)$ satisfies axioms of the distance.*

Proof.

1. $d(A,B) \geq 0$ and $d(A,B) = 0$ if and only if $A = B$: this is evident (in the same way as for Klein model). More precisely, logarithm is zero if and only if the cross-ratio equals one, which in view of Proposition 5.26.
2. $d(A,B) = d(B,A)$ since $[A,B,X,Y] = [B,A,Y,X]$.
3. Triangle inequality $d(A,B) + d(B,C) \geq d(A,C)$ will by proved below in Corollary 6.12.

□

Example 6.5. Examples of isometries in the Poincaré disc model:

1. Rotation about the centre of the model;
2. reflection with respect to a diameter;
3. inversion with respect to a circle representing a hyperbolic line.
(Notice that such an inversion preserves the boundary of the disc, as a circle representing a hyperbolic line is orthogonal to the absolute; also, it takes inside of the disc to inside as the hyperbolic line is preserved pointwise).

All of these transformations preserve the disc and preserve the cross-ratios, so are isometries. Inversions and reflections described above play a role of hyperbolic reflections - they fix a hyperbolic line pointwise and swap the half-planes.

Proposition 6.6. *Let $l \in \mathbb{H}^2$ be a (hyperbolic) line, $A \in \mathbb{H}^2$ or $A \in \partial\mathbb{H}^2$ be a point, $A \notin l$. Then there exists a unique line l' through A orthogonal to l.*

Proof. Existence: Let I_l be an inversion with respect the line l. I_l swaps the points A and $A' = I(A)$ ($A' \in \mathbb{H}^2$ as I_l takes the points of $\mathbb{H}^2$ to the points of $\mathbb{H}^2$), see Fig. 77, left. This implies that I_l preserves the hyperbolic line AA' (as I_a swaps A and A' and preserves the point $l' \cap \partial\mathbb{H}^2$). So, AA' is orthogonal to l and we can take $l' = AA'$.

Uniqueness If $l' \perp l$ then $I_l(l') = l'$, so l' contains the point $A' = I_l(A)$, which implies $l' = AA'$. □

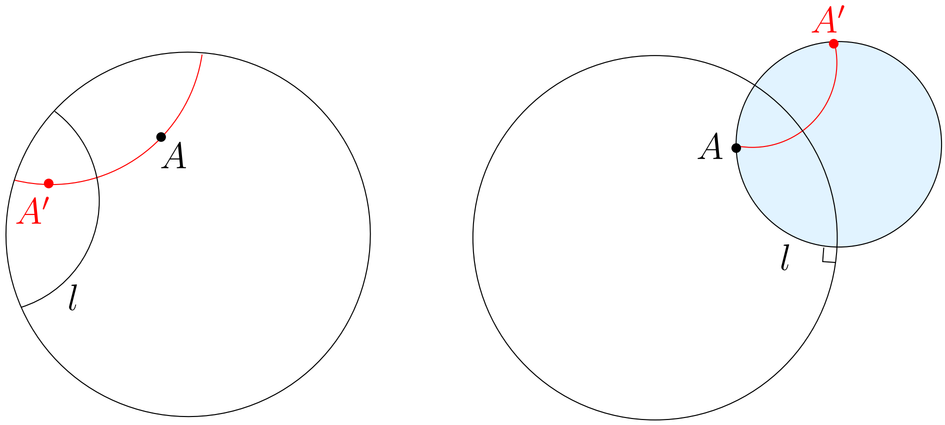


Figure 77: Construction of a perpendicular line to l through A.

Proposition 6.7. *Let $l \in \mathbb{H}^2$ be a (hyperbolic) line, $A \in \mathbb{H}^2$ be a point, $A \in l$. Then there exists a unique line l' through A orthogonal to l.*

Proof. Let I_a be the inversion in the absolute. Then $I_a(l) = l$ (as Euclidean circles/lines representing hyperbolic lines are orthogonal to the absolute). Let $A' = I_a(A)$ (notice that $A' \notin \mathbb{H}^2$), see Fig. 77, right.

Let l' be the line or circle through A and A' such that $l' \perp l$ (it does exist in view of Remark 6.2 applied to the shaded disc bounded by l). Then $I_a(l') = l'$ (as I_a swaps A and A' and preserves the point $l' \cap \partial\mathbb{H}^2$). From this we conclude that $l' \perp l$ and hence, l' represents a hyperbolic line orthogonal to l and containing A. □

Proposition 6.8. *Every hyperbolic segment has a midpoint.*

Proof. Let $X, Y \in \partial\mathbb{H}^2$ be the endpoints of the hyperbolic line AB. Let $f \in M\ddot{o}b$ be a map of the unit disc to itself such that $f(X) = X$, $f(Y) = Y'$ where XY' is a diameter of the disc, and $f(Z) = Z$ for some $Z \in \partial\mathbb{H}^2$ (this map exists by triple transitivity of Möbius transformations on the points). We will show that the segment $f(A)f(B)$ has a midpoint, and this will imply the same for AB as f preserves the cross-ratio.

From now on we assume that A, B lie on a diameter. Consider a point B' on the same diameter such that B lies between A and B'. Then

$$d(A,B) = |ln\frac{|XA|}{|XB|}/\frac{|YA|}{|YB|}| = |ln\frac{|XA|}{|XB|}\frac{|YB|}{|YA|}| < |ln\frac{|XA|}{|XB'|}\frac{|YB'|}{|YA|}| = d(A,B'),$$

which means that $d(A, B)$ is a strictly monotone function. It is also clearly continuous.

Now, consider a point $T = T(t)$ moving from the point $A = T(0)$ to the point B so that $d(A, T) = t$ (i.e. $B = T(d_0)$) where $d_0 = (A, B)$)). Then $d(A, T)$ grows monotonically from 0 to d_0 while $d(T, B)$ declines monotonically from d_0 to 0, which implies that in some intermediate point M these distances coincide. □

Remark. When $B = B(t)$ runs along a ray AX from A to X, the distance $d(A, B(t))$ grows monotonically from 0 to ∞.

Theorem 6.9. *The isometry group of $\mathbb{H}^2$ acts transitively*

(1) on triples of points of the absolute;

(2) on points in $\mathbb{H}^2$.

Proof. (1) There exists $f \in M\ddot{o}b$ taking any three given points of the absolute to any other three given points. Then f takes the absolute to itself. If it takes the inside of the disc to the outside, consider $f' = I_a \circ f$ where I_a is the inversion with respect to the absolute. Then f' preserves the disc and preserves the cross-ratio, so it is an isometry.

(2) Let O be the centre of the disc and A be a point. It is enough to find an isometry f_A which takes A to O (and when we need to map A to B we will consider a composition $f_B^{-1} \circ f_A$). Let M be the (hyperbolic) midpoint of the hyperbolic segment OA (the midpoint exists in view of Proposition 6.8). Let l' be the (hyperbolic) line orthogonal to OA and containing M (exists by Proposition 6.6), see Fig. 78. Then inversion $I_{l'}$ with respect to l' preserves the line OA (as $l' \perp OA$) and swaps the points of OA lying on the same hyperbolic distance from M, i.e. $I_{l'}(A) = O$. So, $I_{l'}$ is the required isometry (as $I_{l'}(D) = D$).

□

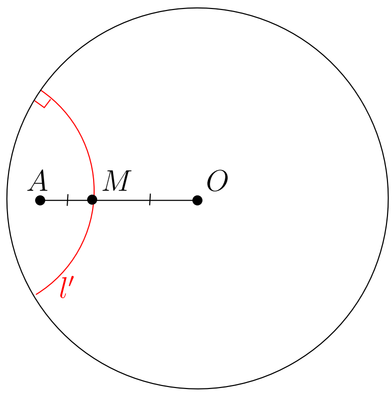


Figure 78: Mapping the point A by isometry to the centre O of the disc.

Remark. Isometries act transitively on flags (recall that a *flag* in $\mathbb{E}^2$ or S^2 or $\mathbb{H}^2$ is a triple (P, r, h^+), where P is a point, r is a straight ray starting from P and h^+ is a choice of half-plane bounded by the line containing r). Indeed, one can map a point to the centre of the disc, then rotate around the centre and reflect with respect to a line through the centre).

Proposition 6.10. *For $C \in AB$, $d(A, C) + d(C, B) = d(A, B)$.*

Proof. Take A to O by an isometry. Then the segment AB is represented by a segment of Euclidean line. For points on a line we have checked that in Remark 4.29 when considered the Klein model, where $d(A, B) = \frac{1}{2}|ln[a, b, x, y]|$.

□

Lemma 6.11. *In a right-angled $\triangle ABC$ with $\angle C = \pi/2$, holds $d(B, C) < d(B, A)$.*

Proof. By transitivity of isometries on the points of $\mathbb{H}^2$ we may assume that B is the centre of the disc. Let γ be a (Euclidean) circle centred at B passing through C, see fig 79. Notice that it is also a hyperbolic circle (i.e. the set of points on the same distance from the centre of the disc). As BC is the radius of γ, we see $CB \perp \gamma$ (as the Euclidean sets). Since the side BC of the hyperbolic triangle ABC is represented by a line or circle orthogonal to the absolute and orthogonal to CB, we conclude that A lies outside of γ. Hence, $BA > BA' = CB$ where $A' = \gamma \cap BA$.

□

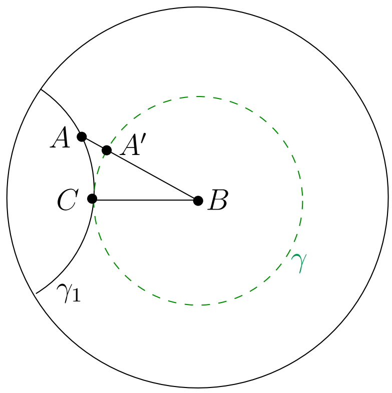


Figure 79: In a right-angled $\triangle ABC$ with right $\angle C$, holds $d(B,C) < d(B,A)$.

Corollary 6.12. *Triangle inequality: for $C \notin AB$,* $\quad d(A,C) + d(C,B) > d(A,B)$.

Proof. Let H be orthogonal projection of B to AC (in hyperbolic sense, i.e. $H = l \cap AC$ where l is the hyperbolic line orthogonal to AC through B). Then

$$d(A,B) + d(B,C) > d(A,H) + d(H,C) \geq (A,C).$$

□

Remark. Triangle inequality implies that (a) distance is well-defined, and
(b) hyperbolic lines are geodesics in the model.

By a hyperbolic circle centred at a point A we mean a set of points on the same (hyperbolic) distance from A.

Proposition 6.13. *(a) Hyperbolic circles are represented by Euclidean circles in the Poincaré disc model.*

(b) Every Euclidean circle in the disc represents a hyperbolic circle.

Proof. (a) Let O be the centre of the disc. A hyperbolic circle centred at O is represented by a Euclidean circle (with the same centre). To obtain a circle centred at a point $A \in H^2$, $A \neq O$, we apply an inversion $I_{l'}$ which swaps O and A (as in the proof of Theorem 6.9). The inversion takes the circle centred at O to a circle centred at A (in hyperbolic sense). In Euclidean sense the image of a circle under an inversion is either a circle or a line; it is a circle as it is contained inside the disc D.

(b) Let $\mathcal{C} \subset D$ be a (Euclidean) circle with Euclidean centre Q. Let let A and B be the points of intersection of the circle $\mathcal{C}$ with the Euclidean lien OQ. Let M be the hyperbolic midpoint of the hyperbolic segment AB. We can map M to O by an isometry f of $\mathbb{H}^2$ (in view of Theorem 6.9). Then $f(\mathcal{C})$ is orthogonal to the (Euclidean) line AB (as $OQ \perp \mathcal{C}$), it is a circle, and passes through two points $f(A)$ and $f(B)$ on the same Euclidean distance from $O = f(M)$. We conclude that as a Euclidean set $f(\mathcal{C})$ is the circle centred at the origin. So, it is a hyperbolic circle. And hence $\mathcal{C}$ is also a hyperbolic circle.

□

Notice that the Euclidean centre of the circle with hyperbolic centre $A \neq O$ is different from A.

Theorem 6.14. *An isometry of $\mathbb{H}^2$ is uniquely determined by the image of a flag.*

Proof. Let $O \in \mathbb{H}^2$ be a point, l be a (hyperbolic) ray from O and h^+ be a choice of a half-plane with respect to the (hyperbolic) line containing l, let $F = (O, l, h^+)$ be a flag. It is sufficient to show that if f is an isometry of $\mathbb{H}^2$ and $f(F) = F$ then $f = id$.

First, notice that if $f(F) = F$ then $f(l) = l$ pointwise, as f preserves the cross-ratio $[O, B, X, Y]$ where $B \in l$ and X, Y are the endpoints of the line containing l.

Now, Let $C \notin l$ be any point. As $d(O, C) = d(O, f(C))$ and $d(B, C) = d(B, f(C))$, we see that $f(C)$ lies on the intersection of two hyperbolic circles: one centred at O of radius OC and another centred at B of (hyperbolic) radius BC. This implies that $f(C)$ lies on the intersection of two Euclidean circles (in view of Proposition 6.13). Since $f(h^+) = h^+$ we conclude that $f(C) = C$ for any choice of the point C, i.e. $f = id$.

□

Theorem 6.15. *Every isometry of the Poincaré disc model can be written as either $\frac{az+b}{cz+d}$ (Möbius transformation) or $\frac{a\bar{z}+b}{c\bar{z}+d}$ (anti-Möbius transformation).*

Proof. Let $g \in Isom(\mathbb{H}^2)$.Let F be a flag. Proving transitivity of isometry on flags we have constructed a Möbius or anti-Möbius transformation $f : F \to g(F)$. Clearly, it is an isometry. Uniqueness shown in Theorem 6.14 implies that $g = f$.

□

Corollary 6.16. *An isometry of $\mathbb{H}^2$ is uniquely determined by the images of three points of the absolute.*

Proof. It follows as Möbius and anti-Möbius transformations are determined uniquely by images of the points.

□

Here, you can find some Hyperbolic Geometry Artworks by Paul Nylander.

Corollary 6.17. *"Isometries preserve the angles", i.e. hyperbolic angles coincide with Euclidean ones.*

Proof. This follows since Möbius and anti-Möbius transformations preserve angles.

□

Exercise 6.18. We can use the results of hyperbolic geometry to show the following statement:

> *Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be two disjoint circles. Then there exists a Möbius transformation which takes them to two concentric circles.*

To show this, first use some Möbius transformation to map $\mathcal{C}_1$ inside of $\mathcal{C}_2$. Then imagine that $\mathcal{C}_2$ is the Poincare disc model and $\mathcal{C}_1$ is a circle inside of it...

Proposition 6.19. *The sum of angles in a hyperbolic triangle is less than* π.

Proof. Since the isometries act transitively on points of hyperbolic plane, it is sufficient to show the statement for a triangle with one vertex at the centre of the disc. For such a triangle we compare angles of Euclidean triangle with angles of the hyperbolic one: the angle at the centre coincide, while the angles at other points a strictly smaller in hyperbolic case see Fig. 80.

□

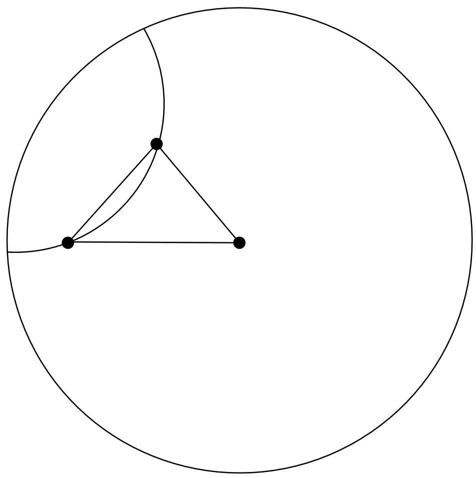


Figure 80: Sum of angles in hyperbolic triangle is smaller than in the Euclidean one.

Remark. On can show that if $\alpha+\beta+\gamma<\pi$ then there exists a triangle with angles α, β, γ.

6.2 Upper half-plane model

Model: $\mathbb{H}^2 = \{z \in \mathbb{C},\ Imz > 0\}$;
$\partial\mathbb{H}^2 = \{Imz = 0\}$, absolute;
- lines: rays and half-circles orthogonal to $\partial\mathbb{H}^2$;
- distance: $d(A,B) = \big|\ln[A,B,X,Y]\big|$;
- angles: same as Euclidean angles.

Group: isometries, i.e. Möbius transformation, inversions, reflections (preserving the half-plane).

Proposition 6.20. *This defines the same geometry as the Poincaré disc model.*

Proof. Consider a Möbius transformation $f = \frac{1}{i}\frac{z+1}{z-1}$ (inverse to Cayley map) taking the unit disc to the upper half-plane. f preserves cross-ratio, and hence, preserves the distance.

□

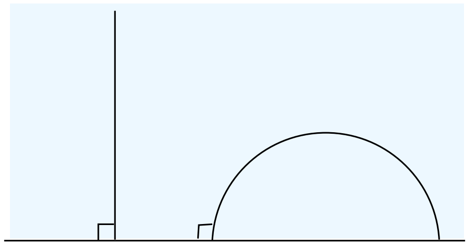


Figure 81: Upper half-plane model.

This implies that we can use in this model all results obtained in the Poincaré disc model.

Proposition 6.21. *In the upper half-plane, hyperbolic circles are represented by Euclidean circles.*

Proof. We know this for the Poincaré disc, so applying a Möbius transformation we get the same for the upper half-plane model.

□

Remark. If the two models are so similar, why do we need both?
This is because different properties are better viewed in different models.
Example:

- In the disc model it is easier to see that (hyperbolic) circles are represented by Euclidean circles.

- In the upper half-plane it is easier to see that there is a hyperbolic line through every two points on the absolute (if two points A and B are given by $a, b \in \mathbb{R}$ then the corresponding line is given by a (Euclidean) circle of radius $|a-b|/2$ centred at $(a+b)/2$).

The distance between two points is also easy to compute in the upper half-plane.

Theorem 6.22. *In the upper half-plane model,*

$$\cosh d(z,w)) = 1 + \frac{|z-w|^2}{2Im(z)Im(w)}.$$

Proof. We start by checking the formula for the points $z = i$, $w = ki$, $k > 1$:

$$d(z,w) = |\ln[i, ki, \infty, 0]| = |\ln \frac{\infty - i}{\infty - ki} / \frac{0-i}{0-ki}| = \ln k.$$

So, we obtain that

$$\cosh(d(z,w)) = \cosh(\ln k) = \frac{1}{2}(e^{\ln k} + e^{-\ln k}) = \frac{1}{2}(k + \frac{1}{k}).$$

At the same time,

$$1 + \frac{|z-w|^2}{2Im(z)Im(w)} = 1 + \frac{(k-1)^2}{2 \cdot 1 \cdot k} = 1 + \frac{k^2}{2k} - \frac{2k}{2k} + \frac{1}{2k} = \frac{1}{2}(k + \frac{1}{k}).$$

To check the formula for general z, w, apply a Möbius transformation taking the points z, w to $i, \lambda i$ for some $\lambda \in \mathbb{R}_+$ (such a transformation exists in view of transitivity of isometries on flags). Then the left hand side is preserved as it is a function of a cross-ratio, while the right hand side is preserved since it is preserved by each of the generators of Möbius transformations, i.e. by $z \to az$, $z \to z+1$, $z \to 1/z$.

$\square$

Theorem 6.23. *Every isometry of the upper half-plane model can be written as either* $z \mapsto \frac{az+b}{cz+d}$ *or* $z \mapsto \frac{a(-\bar{z})+b}{c(-\bar{z})+d}$ *with* $a, b, c, d \in \mathbb{R}$, $ad - bc > 0$.

Proof. Isometries of the upper half-plane are isometries of the Poincaré disc conjugated by Möbius transformations. So, from Theorem 6.15 we conclude that all isometries of the upper half-plane are either Möbius or anti-Möbius transformations.

To conclude about the coefficients a, b, c, d, we will first consider orientation-preserving isometries. Notice that is f is an isometry of the upper half-plane, then f takes the real line to itself. This implies that one can choose the coefficients a, b, c, d to be real: indeed, we have $f(\infty) = a/c \in \mathbb{R}$, $f(0) = b/d \in \mathbb{R}$, since also $f(1) \in \mathbb{R}$ one can conclude that $a/b \in \mathbb{R}$ (check!). Finally,

$$f(i) = \frac{ai+b}{ci+d} = \frac{(ai+b)(ci-d)}{-c^2-d^2} = \frac{-bd-ac+i(bc-ad)}{-c^2-d^2}.$$

As f maps the upper half-plane to the upper half-plane, $Im(f(i)) = -\frac{bc-ad}{c^2+d^2} > 0$ which is equivalent to $ad - bc > 0$. This finishes the proof for orientation-preserving isometries.

An orientation reversing isometry g may be considered as a composition of the map $r : z \to -\bar{z}$ (which is a reflection with respect to the imaginary axis, see Fig. 82) with an orientation- preserving isometry $f = g \circ r$. Then $g = f \circ r^{-1} = f \circ r$. As $f(z) = \frac{az+b}{cz+d}$, $a, b, c, d \in \mathbb{R}$, $ad - bc > 0$ by above, we conclude that $g(z) = \frac{a\bar{z}+b}{c\bar{z}+d}$ with the same restrictions on the coefficients.

$\square$

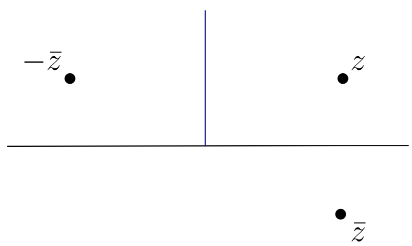


Figure 82: Reflection $z \to -\bar{z}$.

Equivalently, orientation-preserving isometries can be written as $z \mapsto \frac{az+b}{cz+d}$ where $a, b, c, d \in \mathbb{R}$, $ad - bc = 1$. Hence, for the group of or.-preserving isometries we have

$$Isom^+(\mathbb{H}^2) = PSL(2, \mathbb{R}) = SL(2, \mathbb{R})/\pm I.$$

Orientation-reversing isometries can be written as $z \mapsto \frac{a\bar{z}+b}{c\bar{z}+d}$ with $a, b, c, d \in \mathbb{R}$ and $ad - bc = -1$.

Remark. One can see that the distance is "larger" near the absolute. To see this consider a line l and two points on it A_0A_1. We can map A_0 to A_1 so that the line l will be mapped to itself and the half-planes with respect to l will not swap. The point A_1 will map to some point A_2 on the same distance from A_1 as A_0. Iterating the same map we will get infinitely many points $A_3, A_4, \dots$ on the same line with the condition $d(A_i, A_{i+1}) = d(A_0, A_1)$.

Remark. Pairs of lines in the conformal models: see Fig. 83 for intersecting, parallel and divergent (or ultra-parallel) lines in the Poincaré disc and in the upper half-plane.

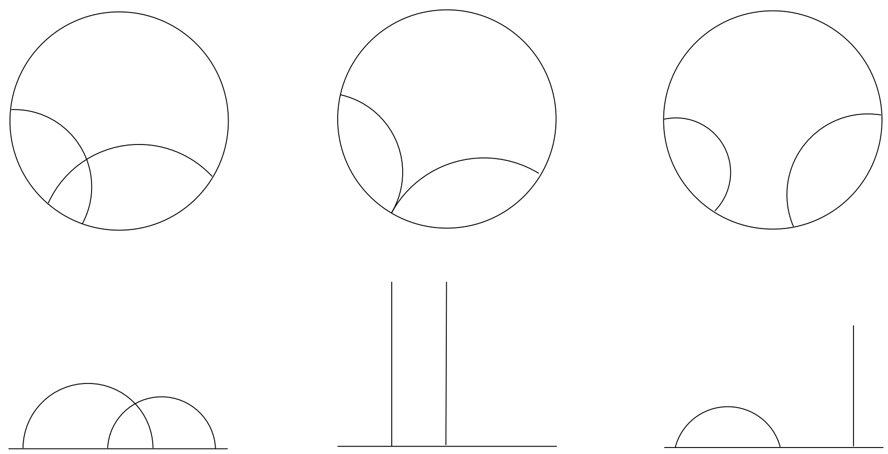


Figure 83: Pairs of line in $\mathbb{H}^2$: intersecting, parallel, divergent (ultra-parallel).

Example. Let l, l' be parallel lines. Then $d(l, l') = 0$ (where by distance between the sets α and β we mean $d(\alpha, \beta) = \inf\limits_{A\in\alpha, B\in\beta}(A, B)$).

To show this, we consider the lines in the upper half-plane model, and we map the point $l \cap l' \in \partial\mathbb{H}^2$ to ∞. Then l and l' are represented by vertical half-lines. By applying an isometry $z \to az + b$, $a, b \in \mathbb{R}$ we may also assume that l lies on the imaginary axis and l' on the line given by $Re(z) = 1$, see Fig. 84 Consider the points $ki \in l$ and $ki + 1 \in l'$. Then

$$\cosh d(ki, ki + 1) = 1 + \frac{1}{2k^2} \to 1$$

as $k \to \infty$. This implies that $d(k, ki) \to 0$ as $k \to \infty$.

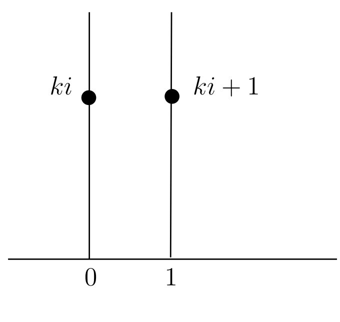


Figure 84: Parallel lines are on distance 0.

6.3 Elementary hyperbolic geometry

Remark. 1. Triangle inequality implies that $d(A, B)$ satisfies axioms of distance and that hyperbolic lines are shortest paths.
2. In hyperbolic geometry, all Euclid's Axioms, except Parallel Axiom hold, while the later clearly does not hold (see Fig. 85).
3. Recall that Parallel Axiom for hyperbolic geometry says:
There are infinitely many lines l' disjoint from a given line l and passing through a given point $A \notin l$.

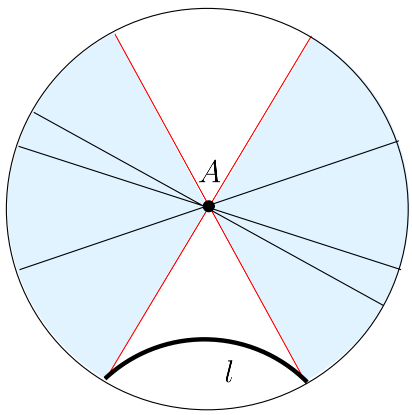


Figure 85: Parallel axiom does not hold in $\mathbb{H}^2$: given a line l and a point A there are infinitely many lines through A ultra-parallel with l, and there are two lines through A parallel to l (labelled red).

Definition 6.24. For a line l and a point $A \notin l$, an angle of parallelism $\varphi = \varphi(A, l)$ is the half-angle between the rays emanating from A and parallel to l, see Fig. 86, left.
Equivalently: drop a perpendicular AH to l, then $\varphi = \angle HAQ$, $Q \in l \cap \partial\mathbb{H}^2$;
Equivalently: a ray AX from A intersects l iff $\angle HAY \le \varphi$, see Fig. 86, right.

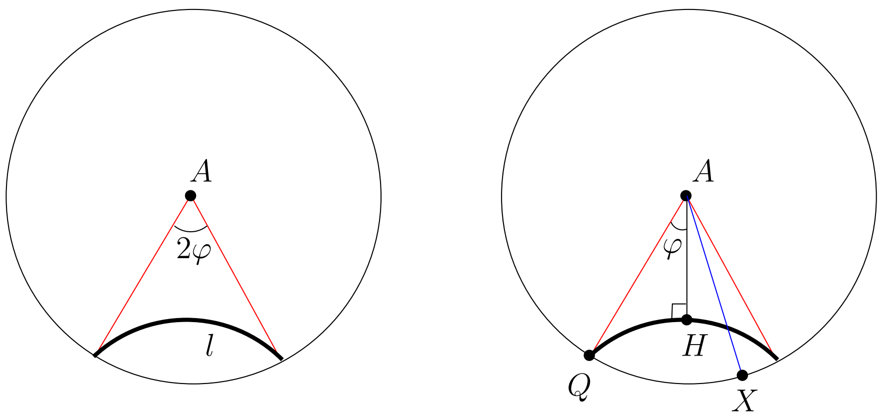


Figure 86: Definition of angle of parallelism.

Remark. The angle of parallelism φ only depends on distance $d(A,l) = \min_{B \in l}(A,B) = d(A,H)$. To see this, map A to the centre of the disc by isometry, so that AH will be mapped to a vertical ray. Then $d(A,H)$ completely determines l as $AG \perp l$.

Proposition 6.25. *For a line l and a point $A \notin l$, let $a = d(A,l)$ and φ be the angle of parallelism. Then* $\cosh a = \frac{1}{\sin\varphi}$.

Proof. We will compute in the upper half-plane model. We may assume that H is the point i, l be a vertical line through i, and $Q := l \cap \partial\mathbb{H}^2$ is the point ∞, see Fig. 87 The right angle at the point H will be formed by the imaginary axis and the circle $|z| = 1$. So, the points of triangle $\triangle HAQ$ are given by $H = i$, $Q = \infty$, $A = e^{i\psi}$ for some $\psi \in [0, \pi/2]$. Notice that if O is the origin then OA is orthogonal to the circle $|z| = 1$. So, ψ is the angle between the radius OA and the horizontal line, while φ is the angle between the tangent to the circle at A and the vertical line - which means that $\psi = \varphi$ and $A = e^{i\varphi}$.

Now, we can compute the distance $d(A,H)$ by

$$\begin{aligned}\cosh d(A,H) = 1 + \frac{|i - e^{i\varphi}|^2}{2Im(i)Im(e^{i\varphi})} &= 1 + \frac{|i - e^{i\varphi}|^2}{2\sin\varphi} \\ &= 1 + \frac{1 + 1 - 2\cos(\frac{\pi}{2} - \varphi)}{2\sin\varphi} = 1 + \frac{1 - \sin\varphi}{\sin\varphi} = \frac{1}{\sin\varphi}.\end{aligned}$$

□

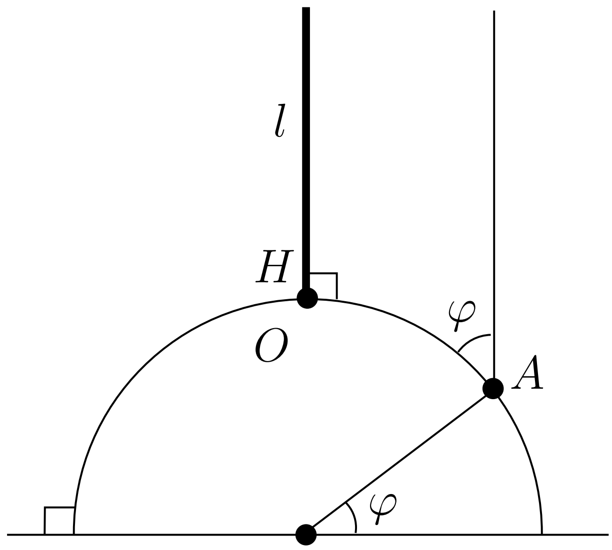


Figure 87: To the proof of Proposition 6.25.

Remark. When $a \to 0$ we have $\cosh a \to 1$ which implies that $\sin\varphi \to 1$, i.e. $\varphi \to \pi/2$, as in Euclidean geometry.

Theorem 6.26 (Hyperbolic Pythagorean theorem)**.** *In a triangle with a right angle γ, $\cosh c = \cosh a \cosh b$.*

Proof. Without loss of generality we may assume that $C = i$, $A = ki$, $B = e^{i\varphi}$, see Fig.88. Then using Proposition 6.25 we get

$$\cosh a = \frac{1}{\sin\varphi}.$$

We also compute using the distance formula:

$$\cosh b = 1 + \frac{(k-1)^2}{2k} = \frac{1+k^2}{2k}$$

and

$$\cosh c = 1 + \frac{\cos^2\varphi + (k-\sin\varphi)^2}{2k\sin\varphi} = \frac{1+k^2}{2k\sin\varphi},$$

which implies the theorem.

□

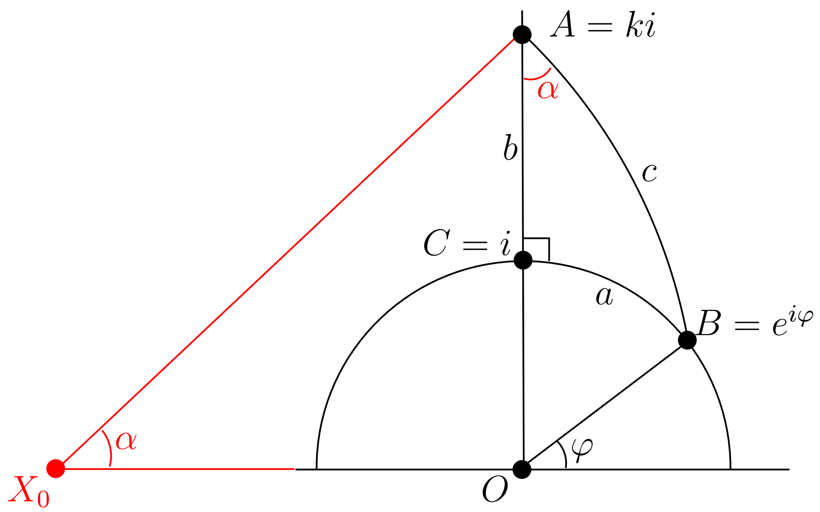


Figure 88: To the proof of Pythagorean Theorem and Lemma 6.6.

Remark. For small values of x we have

$$\cosh x = \frac{1}{2}(1 + x + \frac{x^2}{2} + \cdots + 1 - x + \frac{x^2}{2} + \ldots) = 1 + \frac{x^2}{2} + \ldots,$$

so the Pythagorean theorem can be written as

$$(1 + \frac{c^2}{2} + \ldots) \approx (1 + \frac{a^2}{2} + \ldots)(1 + \frac{b^2}{2} + \ldots),$$

which is equivalent to $c^2 \approx a^2 + b^2$, approaching the Euclidean version of Pythagorean theorem.

Lemma 6.27. *In a triangle with a right angle γ holds:*

$$\sinh a = \sinh c \sin\alpha.$$

Proof. We compute using the same Fig.88 as before:

$$\sinh^2 a = \cosh^2 a - 1 = \frac{1}{\sin^2\varphi} - 1 = \frac{\cos^2\varphi}{\sin^2\varphi}.$$

Also,

$$\sinh^2 c = \cosh^2 c - 1 = (\frac{1+k^2}{2k\sin\varphi})^2 - 1 = \frac{(k^2+1)^2 - 4k^2\sin^2\varphi}{4k^2\sin^2\varphi}.$$

Next, we need to compute $\angle\alpha = \angle BAC$. To do this, let x be the number such that $X = (x,0)$ is the centre of the Euclidean circle containing the (hyperbolic) segment AB. Notice that $\angle AXO = \alpha$ (as the XA is the radius of the circle representing the hyperbolic segment AB and hence is perpendicular to it at A). Notice that as $XA = XB$, we get $x^2 + k^2 = (\cos\varphi - x)^2 + \sin^2\varphi$, from where we get $k^2 = 1 - 2x\cos\varphi$, and hence $x = \frac{1-k^2}{2\cos\varphi}$.

Next, from the right-angled triangle $\triangle AXO$ we have

$$\sin^2\alpha = \frac{k^2}{k^2+x^2} = \frac{k^2}{k^2 + (\frac{k^2-1}{2\cos\varphi})^2} = \frac{4k^2\cos^2\varphi}{(k^2-1)^2 + 4k^2\cos^2\varphi} =$$

$$\frac{4k^2\cos^2\varphi}{(k^2+1)^2 - 4k^2 + 4k^2\cos^2\varphi} = \frac{4k^2\cos^2\varphi}{(k^2+1)^2 - 4k^2\sin^2\varphi}.$$

This implies that

$$\sinh c \sin\alpha = \frac{\cos\varphi}{\sin\varphi} = \sinh a.$$

□

Theorem 6.28 (Sine rule)**.** $\frac{\sinh a}{\sin\alpha} = \frac{\sinh b}{\sin\beta} = \frac{\sinh c}{\sin\gamma}$.

Proof. Let CH be the altitude in $\triangle ABC$, see Fig. 89. Denote $h = |CH|$ (where by $|CH|$ we mean the hyperbolic length of the segment CH). Then from the right-angled triangles AHC and BHC by Lemma 6.6 we get

$$\sinh h = \sinh b \cdot \sin\alpha = \sinh a \cdot \sin\beta,$$

which implies $\frac{\sinh b}{\sin\beta} = \frac{\sinh a}{\sin\alpha}$.

□

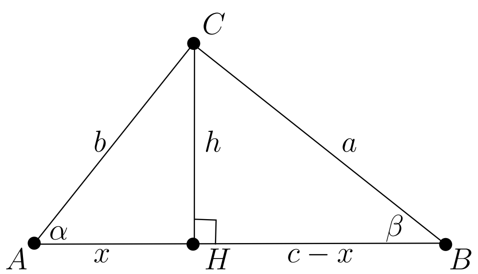


Figure 89: To the proof of Sine and Cosine Rules.

Notice that in the proof of the sine rule we did not use any model!

Exercise:

$$\cosh(a-b) = \cosh a \cosh b - \sinh a \sinh b \tag{6.1}$$

Hint: one can prove it from the definition $\cosh x = \frac{1}{2}(e^x + e^{-x})$.

Exercise 6.29. In a triangle with a right angle γ holds:

$$\tanh b = \tanh c \cos \alpha.$$

The proof is a part of Assignment 15-16, Question 15.2.

Theorem 6.30 (Cosine rule)**.** $\cosh a = \cosh b \cosh c - \sinh b \sinh c \cos \alpha.$

Proof. Let CH be the altitude dropped from C, let h be its length, let x be the length of AH, see Fig. 89. Then from Pythagorean Theorem we have

$$\cosh b = \cosh h \cosh x, \tag{6.2}$$

similarly

$$\begin{aligned}
\cosh a &= \cosh h \cosh(c - x) \overset{(6.1)}{=} \cosh h(\cosh c \cosh x - \sinh c \sinh x) \\
&\overset{(6.2)}{=} \cosh b \cosh c - \frac{\cosh b}{\cosh x} \sinh c \sinh x = \cosh b \cosh c - \cosh b \sinh c \tanh x \\
&\overset{(6.5)}{=} \cosh b \cosh c - \cosh b \tanh b \sinh c \cos \alpha \\
&= \cosh b \cosh c - \sinh b \sinh c \cos \alpha.
\end{aligned}$$

□

Remark. For small values of a, b, c we get Euclidean sine and cosine laws.

Theorem 6.31 (Second cosine rule)**.** $\cos \alpha = -\cos \beta \cos \gamma + \sin \beta \sin \gamma \cosh a.$

We omit the proof (one can find it in the book by Prasolov and Tikhomirov).

Exercise 6.32 (Congruence of triangles)**.**
Prove SSS, SAS, ASA and AAA rules of congruence of triangles in hyperbolic plane. (One can do it either as a corollary of sine/cosine laws or directly, see also HW 13.3). *Hint:* to prove AAA one can use the diagrams in Fig. 90.

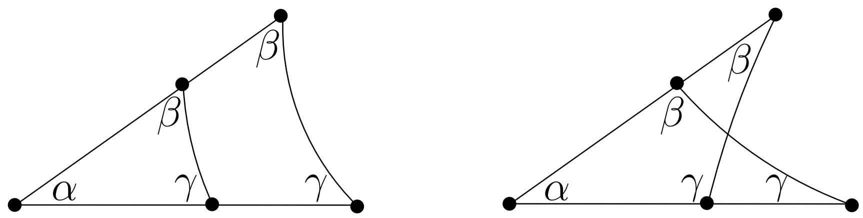


Figure 90: Proving AAA congruence.

Remark. Here you can find an applet to make hyperbolic tessellations of images, by Malin Christersson.

Example: One can use the sine law to compute length of circle of radius r:

$$l(r) = 2\pi \sinh r. \tag{6.3}$$

Proof. To show it, we inscribe a regular n-gon P_n into the circle (we can draw it in the disc model with the centre of P_n at the centre of disc, then the vertices of P_n will be represented by vertices of regular Euclidean n-gon, see Fig. 91). Then we subdivide P_n into $2n$ right-angled triangles, and compute perimeter of P_n as

$$\mathcal{P}(P_n) = 2nd,$$

where d is the half of the side of P_n. From the sine rule we have

$$\frac{\sinh d}{\sin \frac{2\pi}{2n}} = \frac{\sinh r}{\sin \frac{\pi}{2}},$$

which implies that $\sinh d = \sinh r \sin \frac{\pi}{n}$.

Now, we compute the length of the circle as a limit of the perimeter of P_n when $n \to \infty$:

$$l(r) = \lim_{n\to\infty} \mathcal{P}(P_n) = \lim_{n\to\infty} 2n \cdot arcsinh(\sinh r \sin \frac{\pi}{n}) \overset{x\sim\sinh x}{=}$$
$$\lim_{n\to\infty} 2n(\sinh r \sin \frac{\pi}{n}) = \lim_{n\to\infty} 2n \sinh r \cdot \frac{\pi}{n} = 2\pi \sinh r.$$

□

Remark. Strictly speaking, using the sequence of polygons P_n, $n \to \infty$ we only show that the length of circle is larger or equal to $2\pi \sinh r$ (as each arc is larger than the corresponding geodesic side of the polygon). To get also the lower bound one needs to consider regular polygons Q_n for which the circle is inscribed (and it is a bit harder to show that the perimeter of each polygon Q_n is indeed larger than $l(r)$).

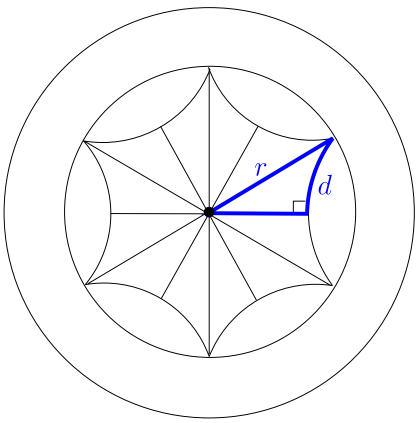


Figure 91: Computing the length of a circle of radius r.

So, the length $l(r)$ of the circle of radius r in spherical, Euclidean and hyperbolic geometry can be expressed by a similar formulae:

	S^2	$\mathbb{E}^2$	$\mathbb{H}^2$
$l(r)$	$2\pi \sin r$	$2\pi r$	$2\pi \sinh r$

Corollary. Uniform statement for sine law in S^2, $\mathbb{E}^2$ or $\mathbb{H}^2$: $\frac{l(a)}{\sin\alpha} = \frac{l(b)}{\sin\beta} = \frac{l(c)}{\sin\gamma}$, where $l(r)$ is the length of circle of radius r in the corresponding geometry.

Remark. As we can see from the formula 6.3, in the hyperbolic geometry the circle length $l(r)$ grows exponentially when $r \to \infty$. One can find examples of such structures in nature (salad leafs, sea weeds, etc.).

6.4 Area of hyperbolic triangle

We will show that area of a hyperbolic triangle with angles α, β, γ is given by

$$S_{\triangle ABC} = \pi - (\alpha + \beta + \gamma). \tag{6.4}$$

Notice that the formula makes sense by two reasons:

- As we know, $\alpha + \beta + \gamma < \pi$, so the area computed by (6.4) is positive;
- the angles α, β, γ completely determine hyperbolic triangle up to isometry (AAA congruence of triangles), so they should determine the area of the triangle.

Definition 6.33. A hyperbolic polygon with all vertices on the absolute is called ideal polygon.

Notice, that in view of Theorem 6.9, all ideal triangles are congruent. In particular, they have the same area.

Theorem 6.34. $S_{\triangle ABC} = \pi - (\alpha + \beta + \gamma)$.

Proof. First, we introduce some notation:

- Let λ the area of an ideal triangle.
- Let $f(\alpha)$ be area of a triangle with angles $(\alpha, 0, 0)$ (i.e. with a triangle with one vertex in $\mathbb{H}^2$ and two vertices at the absolute, see Fig. 92, left).

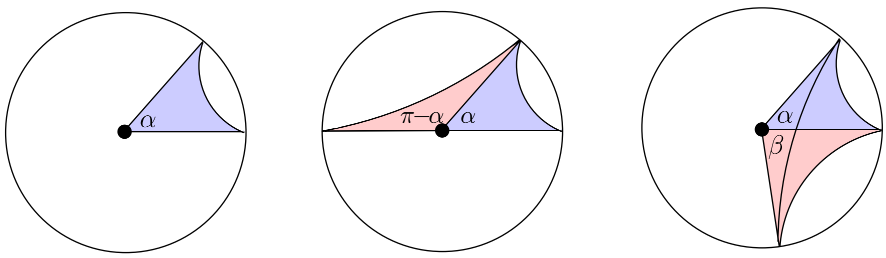


Figure 92: Properties of area $f(\alpha)$ of triangle with angles $(\alpha, 0, 0)$.

We will show several properties of the function $f(\alpha)$:

(1) $f(\alpha)+f(\pi-\alpha) = \lambda$ (an ideal triangle can be assembled from two smaller triangles as in Fig. 92, middle).

(2) $f(\alpha) + f(\beta) = f(\alpha + \beta) + \lambda$ (a quadrilateral in Fig. 92, right can be splitted in two triangles in two ways).

(3) By Property (1), $f(\frac{\pi}{2}) + f(\frac{\pi}{2}) = \lambda$, which implies $f(\frac{\pi}{2}) = \frac{\lambda}{2}$.
Next, by property (2), $f(\frac{\pi}{4}) + f(\frac{\pi}{4}) = f(\frac{\pi}{2}) + \lambda$ which implies $f(\frac{\pi}{4}) = \frac{3}{4}\lambda$.
By (1) again, $f(\frac{3\pi}{4}) = \lambda - \frac{3}{4}\lambda = \frac{\lambda}{4}$.
And by (2) again: $f(\frac{\pi}{8}) = \frac{1}{2}f(\frac{\pi}{4}) + \frac{\lambda}{2} = \frac{7}{8}\lambda$.
....
In the same way we obtain that $f(\frac{k\pi}{2^m}) = \lambda(1 - \frac{k}{2^m})$.

(4) $f(\alpha) > f(\alpha')$ if $\alpha < \alpha' < \pi$. Indeed, let $0 < \beta = \alpha' - \alpha$, then

$$f(\alpha) - f(\alpha + \beta) \overset{(2)}{=} \lambda - f(\beta) \overset{(1)}{=} f(\pi - \beta) > 0.$$

(5) From (3) and (4) we conclude that $f(\alpha) = \frac{\lambda}{\pi}(\pi - \alpha)$.

Using Property (5) in the same way as we used areas of spherical digons for proving the formula of area for the spherical case, we will show that $S_{\triangle ABC} = \frac{\lambda}{\pi}(\pi - (\alpha + \beta + \gamma))$ (see Lemma 6.35 below). Then, in Lemma 6.36 we will use small triangles to find out that $\lambda = \pi$, which will finish the proof.

□

Lemma 6.35. $S_{\triangle ABC} = \frac{\lambda}{\pi}(\pi - (\alpha + \beta + \gamma))$

Proof. An intersection of two lines at angle α produces two triangles with angles $(\alpha, 0, 0)$ ("hyperbolic α-digons") (if $\alpha = \pi/2$ there are four such triangles but we will be interested in one pair of non-adjacent ones). A triangle with angles α, β, γ produces three pairs of hyperbolic digons, which all together cover an ideal hyperbolic hexagon, see Fig. 93, left. Notice that the triangle $\triangle ABC$ itself is covered three times, while all other parts of the hexagon only covered once. So,

$$2f(\alpha) + 2f(\beta) + 2f(\gamma) - 2S_{\triangle ABC} = S_{ideal\ hexagon} = 4\lambda,$$

where the last equality holds since an ideal hexagon can be composed of 4 ideal triangles (see Fig. 93, right). From this we get $f(\alpha) + f(\beta) + f(\gamma) - S_{\triangle}ABC = 2\lambda$, i.e.

$$S_{\triangle ABC} = \lambda((1 - \frac{\alpha}{\pi}) + (1 - \frac{\beta}{\pi}) + (1 - \frac{\gamma}{\pi}) - 2\lambda = \lambda(1 - \frac{\alpha}{\pi} - \frac{\beta}{\pi} - \frac{\gamma}{\pi}).$$

□

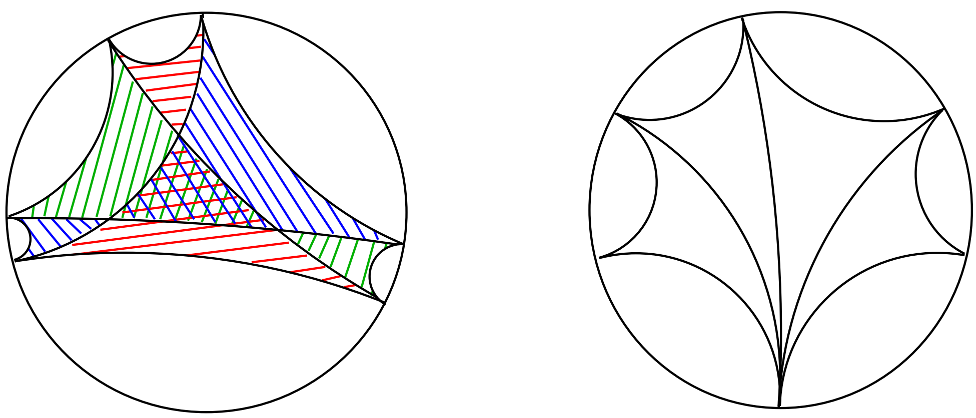


Figure 93: Area of hyperbolic triangle.

Lemma 6.36. $\lambda = \pi$.

Proof. Consider a small right-angled triangle with sides $a, b, c \to 0$, $\gamma = \pi/2$. Then the triangle is almost Euclidean, i.e. $\alpha + \beta + \gamma \to \pi$, $\pi - (\alpha + \beta + \gamma) \to 0$. This implies that

$$\pi-(\alpha+\beta+\gamma) \approx \sin(\pi-(\alpha+\beta+\gamma) \overset{\gamma=\frac{\pi}{2}}{=} \sin(\frac{\pi}{2}-(\alpha+\beta)) = \cos(\alpha+\beta) = \cos\alpha\cos\beta-\sin\alpha\sin\beta.$$

Next, we use two formulae concerning right-angled triangles and obtained in HW 15.2:

$$\tanh b = \tanh c \cos\alpha \tag{6.5}$$

and

$$\sinh a = \sinh c \sin\alpha \tag{6.6}$$

and continue the computation as follows:

$$\begin{aligned}\pi - (\alpha+\beta+\gamma) \approx \frac{\tanh b}{\tanh c}\cdot\frac{\tanh a}{\tanh c} - \frac{\sinh a}{\sinh c}\cdot\frac{\sinh b}{\sinh c} &= \\ \frac{\sinh a \sinh b}{\sinh^2 c}\left(\frac{\cosh^2 c}{\cosh a \cosh b} - 1\right) &\overset{\cosh c=\cosh a\cosh b}{=} \\ \frac{\sinh a\sinh b}{\sinh^2 c}(\cosh c - 1) \approx \frac{a\cdot b}{c^2}\cdot\frac{c^2}{2} = \frac{ab}{2} &= S_{\text{Euclidean triangle}}.\end{aligned}$$

As $S_{\triangle ABC} = \frac{\lambda}{\pi}(\pi - (\alpha+\beta+\gamma))$ and $\pi - (\alpha+\beta+\gamma)$ tend to the area of Euclidean triangle when a, b, c decrease, we conclude that $\lambda = \pi$.

□

Remark. The value $\delta = \pi - (\alpha+\beta+\gamma)$ is called a defect of the triangle.

Corollary 6.37. *Area of an n-gon: $S_n = (n-2)\pi - \sum_{i=1}^n \alpha_i$.*

Remark. (Refraction). Return for a minute to a real life. Recall, that the speed of light depends on properties of the material. At the same time, light choose the quickest path, which implies that a light ray bends when hitting the boundary of two transparent materials.

More precisely, consider the edge of two transparent materials (we will call them "air" and "water"). We assume that the speed of light in the air if v_1 (m/sec) and in the water is $v_2 < v_1$ (m/sec). We want to connect two given points (with coordinates $(0, h_1)$ and $(1, h_1)$, see Fig. 94), one in the air and one in the water, by the "shortest path" in sense that it is fastest for the light ray. It will be a broken line with two segment (as inside each material the speed of light is constant), but we do not know in advance at which point $X = (x, 0)$ will the ray hit the boundary of water. Depending on coordinate x, the time $t(x)$ needed for the ray to travel between the two points is

$$t(x) = \frac{\sqrt{h_1^2 + x^2}}{v_1} + \frac{\sqrt{h_2^2 + (1-x)^2}}{v_2},$$

and we need to find $\min_x t(x)$. We find where the derivative of t vanishes:

$$t'(x) = \frac{1}{2v_1}\frac{2x}{\sqrt{h_1^2+x^2}} - \frac{1}{2v_2}\frac{2(1-x)}{\sqrt{h_2^2+(1-x)^2}} = 0.$$

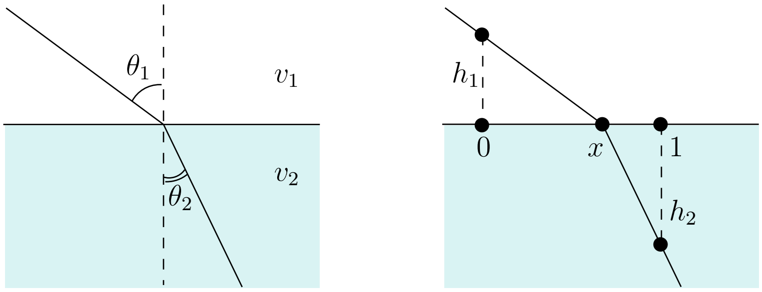


Figure 94: Refraction.

From this equation (and using notation as on Fig. 94, left) we obtain

$$\frac{\sin\theta_1}{v_1} = \frac{\sin\theta_2}{v_2}.$$

This implies that if $h_1 = h_2$ then the ray will travel more in the air than in the water.

If the density of the material changes several times we get a broken line as the shortest path, like the one in Fig. 95. And when the density grows continuously when approaching the real line, so that the speed at the point (x, y) is equal to $1/y$, one can get hyperbolic lines as geodesics (i.e. as trajectory of the light rays).

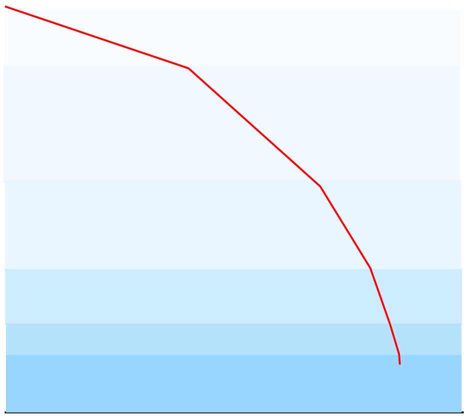


Figure 95: Broken lines as shortest paths.

6.5 References

- Section 6 is based on Lectures VI – VIII in Prasolov's book.
  Alternatively, see pp.95–104 in Section 5.2 of Prasolov, Tikhomirov.

- See also

  - *Hyperbolic Geometry*. Lecture notes by Caroline Series.
  - *Hyperbolic Geometry*. Lecture notes by Charles Walkden.

- A short paper to stimulate your non-Euclidean intuition:
  Non-Euclidean billiards in VR by Jeff Weeks.

- Webpages, pictures, videos:

  - Hyperbolic Geometry Artworks by Paul Nylander.
  - Webpage on hyperbolic geometry by the Institute for Figuring (with hyperbolic soccer ball and crocheted hyperbolic planes).
  - Tilings by (and after) Escher. Webpage on Mathematical Imagery by Jos Leys.
  - How to create repeating hyperbolic patterns, by Douglas Dunham (based on Escher's patterns). See also here.
  - Playing Sports in Hyperbolic Space - Dick Canary in a Numberphile video (by Brady Haran).

- Software:

  - Applet for creating hyperbolic drawings in Poincaré disc.
  - Applet to make hyperbolic tessellations of images, by Malin Christersson.

7 Other models of hyperbolic geometry

7.1 Klein disc, revised

Reminder:

- the model is inside the unit disc, lines are represented by chords.
- distance in Klein disc $d(A,B) = \frac{1}{2}\big|ln[A,B,X,Y]\big|$, where X, Y are the endpoints of the chord through A, B, see Fig. 52;
- isometries are projective maps preserving the disc.

Theorem 7.1. *Geometry of the Klein disc (D_K) coincides with geometry of the Poincaré disc (D_P).*

Proof. Idea of the proof: we will build a map $f : D_K \to D_P$ which takes the Klein disc to the Poincaré disc and will show that the distance between points in D_K will coincide with the distance between their images in D_P.

We will construct the map f as a composition of two projection. First, consider a unit sphere $S^2 \in \mathbb{R}^3$, and let $D_K = \{(x,y,z) \in \mathbb{R}^3 \mid z = 0,\ x^2 + y^2 + z^2 < 1\}$ be the horizontal unit disc inside the sphere. Consider the orthogonal (vertical) projection $p : D_k \to S^2$ of the disc to the lower hemisphere:

$$p : (x,y,0) \mapsto (x,y,-\sqrt{1-x^2-y^2}).$$

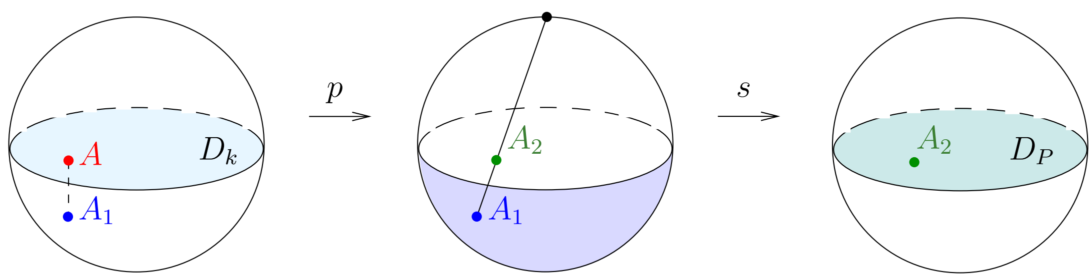


Figure 96: Projecting the Klein disc to the hemisphere and then to the Poincaré disc.

Next, we apply a stereographic projection $s : S^2 \to D_P$ from the North Pole $(0,0,1)$ which will take the lower hemisphere to the disc

$$D_P = \{(x,y,z) \in \mathbb{R}^3 \mid z = 0, x^2 + y^2 + z^2 < 1\}$$

(see Fig. 96). The composition

$$f := s \circ p$$

has the following properties:

- f maps D_k to D_P bijectively (as a map of sets).
- An image of a line in D_k is a line in D_P:
  Indeed, a chord is mapped by p to a vertical semicircle, i.e. to a semicircle lying on the sphere and orthogonal to the boundary of the horizontal disc. Then the stereographic projection maps it to a part of a circle orthogonal to the boundary of the disc (as s preserves the angles).

- Let $A, B \in D_k$, let $A_1 = p(A), B_1 = p(B)$ and let $A_2 = f(A)$, $B_2 = f(B)$. Let X, Y be the endpoints of the chord AB. Then

$$[A, B, X, Y] = [A_2, B_2, X, Y]^2. \tag{7.1}$$

Proof. As the stereographic projection s preserves cross-ratios, it is enough to prove that

$$[A, B, X, Y] = [A_1, B_1, X, Y]^2.$$

Note that since p maps XY to a semicircle, we have $\angle XA_1Y = \pi/2$, and moreover, triangles $\triangle AA_1X$, $\triangle AYA_1$ and $\triangle A_1YX$ are all similar. So, we get

$$\frac{|XA|}{|YA|} = \frac{|XA|}{|AA_1|} \cdot \frac{|AA_1|}{|YA|} = \left(\frac{|XA_1|}{|YA_1|}\right)^2.$$

So, we compute

$$[A, B, X, Y] = \frac{XA}{XB} \Big/ \frac{YA}{YB} = \frac{XA}{YA} \Big/ \frac{XB}{YB} = \left(\frac{XA_1}{YA_1} \Big/ \frac{XB_1}{YB_1}\right)^2 = [A_1, B_1, X, Y]^2.$$

End of proof of (7.1).

- Denote by $d_K(A, B)$ and $d_P(A_2, B_2)$ the distance measured between A and B in the Klein model and the distance between their images in the Poincaré model, as in Fig. 96. Then

$$d_K(A, B) = \frac{1}{2}|ln[A, B, X, Y]| \overset{(7.1)}{=}$$
$$\frac{1}{2}|ln[A_2, B_2, X, Y]^2| = |ln[A_2, B_2, X, Y]| = d_p(A_2, B_2).$$

$\square$

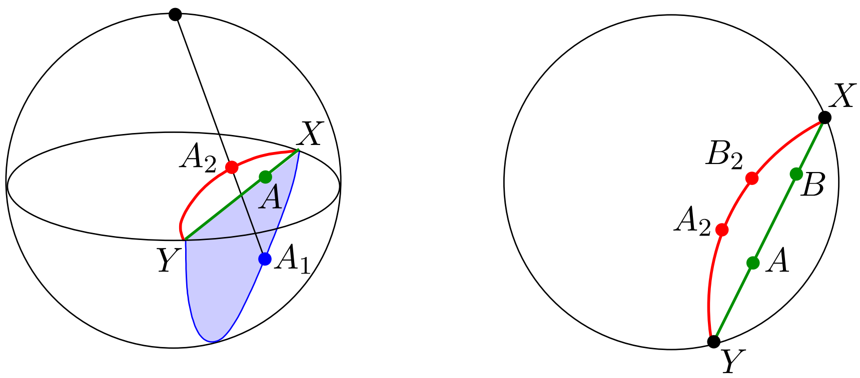


Figure 97: Comparing the cross-ratios in Klein and Poincaré disc models.

Remark: One can use light to project the hemisphere model to Klein disc, Poincare disc and upper half-plane. See the following video by Hynry Segerman and Saul Schleimer.

Remark: When to use the Klein disc model? It is useful in many cases when we need to work with lines and right angles.

Examples:

- Right angles are displayed nicely in the Klein model (see Proposition 4.33, see also Fig. 98, left).
- One can construct the common perpendicular to any two divergent lines (see Proposition 4.34, see Fig. 98, middle left).
- One can construct a midpoint for any segment (see Fig. 98, middle right).
- and an angle bisector (see also Fig. 98, right).
- This implies that one can construct centres of the inscribed and circumscribed circles (when exist).

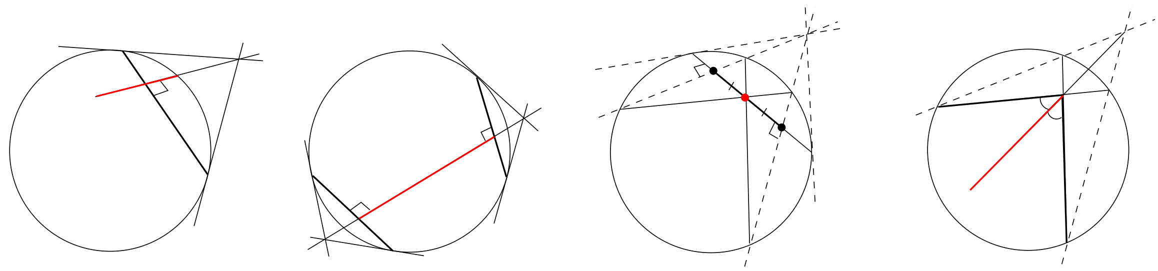


Figure 98: Constructions in Klein model: right angles, common perpendicular, midpoint, angle bisector.

Remark: circles in the Klein model are represented by ellipses. Indeed, a circle centred at the centre of the disc is clearly represented by a Euclidean circle. Projective transformations take a circle to an ellipse, or hyperbola, or parabola - however, out of them only ellipses fit inside the unit disc. So we conclude that all circles are represented by ellipses.

7.2 The model in two-sheet hyperboloid

Consider the hyperboloid $H \in \mathbb{R}^3$ given by the equation

$$x_1^2 + x_2^2 - x_3^2 = -1,$$

where $x_1, x_2, x_3 \in \mathbb{R}$. This is a two-sheet hyperboloid, it has two connected components (one with $z > 0$ and another with $z < 0$). We will projectivise it, i.e. identify the points $(x_1, x_2, x_3) \sim (-x_1, -x_2, -x_3)$.

<u>Model</u>:

- $\mathbb{H}^2$ = {points of the upper sheet}. Can be also understood as the set of lines through O intersecting H, see Fig. 99.
- the absolute $\partial\mathbb{H}^2\{$ are (projectivised) points of the cone $x_1^2 + x_2^2 - x_3^2 = 0$ }, i.e. lines spanning the cone.

- Lines in H^2: intersections of planes through O with the hyperboloid.

- Distance: $d(A,B) = \frac{1}{2}\left|ln[A,B,X,Y]\right|$
(cross-ratio of the four corresponding lines in $\mathbb{R}^3$, see Fig. 99, right).

Group: - Isometries: projective transformations preserving the cone.

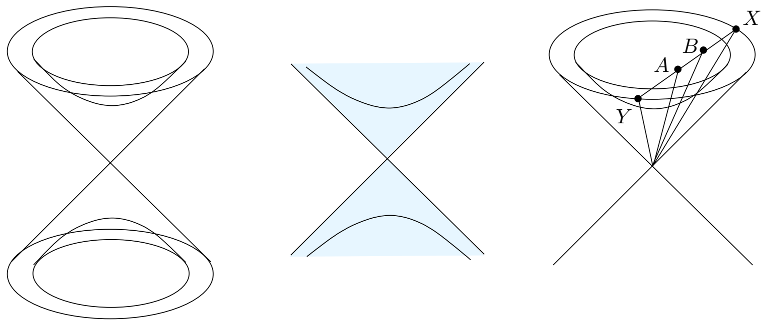


Figure 99: The model on two-sheet hyperboloid.

Theorem 7.2. *This determines the same hyperbolic geometry as the Klein model.*

Proof. We construct a bijective map $pr : H \to D_K$ from the hyperboloid model H to the Klein disc D_K and will show that this map is distance-preserving.

Let $D_K = \{x_3 = 1\} \cap \{x_1^2 + x_2^2 - x_3^2 < 0\}$ by the intersection of the horizontal plane $x_3 = 1$ with the inside of the cone. Let $pr : H \to D_K$ be the projection from the origin of the hyperboloid H to the disc D_K, see Fig. 100. Then

- Points of H are mapped bijectively to points of D_K.
- Lines in H (i.e. intersections of the hyperboloid with planes through the origin) are mapped to lines in D_K (i.e. intersection of the plane $x_3 = 1$ with the planes).
- Cross-ratio of points in H (i.e. cross-ratio of the corresponding lines in $\mathbb{R}^3$) coincides with the cross-ratio of the points of their intersection with the plane $x_3 = 1$.

From this we conclude, that pr is a distance-preserving bijection.

□

For $x = (x_1, x_2, x_3)$, $y = (y_1, y_2, y_3)$ define a pseudo-scalar product by

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 - x_3 y_3.$$

Then

- points of the $\mathbb{H}^2$: $\langle x, x \rangle = -1$;
- points of the $\partial \mathbb{H}^2$: $\langle x, x \rangle = 0$;
- hyperbolic line l_a: $a_1 x_1 + a_2 x_2 - a_3 x_3 = 0$, i.e. $\langle a, x \rangle = 0$.

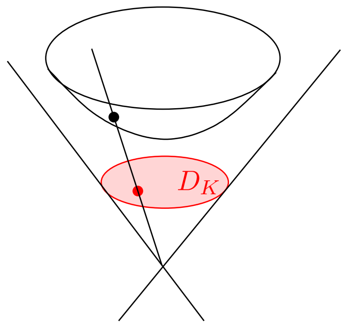


Figure 100: Projection from hyperboloid the the Klein disc. (Here the disc is drawn in the plane $z = 1/2$ for the clarity of the diagram).

Remark. Geometric meaning of $\langle a, a\rangle$:

- if $\langle a, a\rangle > 0$ then l_a intersects the cone producing a hyperbolic line;
- if $\langle a, a\rangle = 0$ then l_a is tangent to cone producing the point a on the absolute;
- if $\langle a, a\rangle < 0$ then l_a does not intersect the cone and gives no line (but a corresponds to a point of $\mathbb{H}^2$).

To explain this, notice that the pseudo-scalar product $\langle a, a\rangle$ does not change when a rotates around the axis Ox_3, in particular, without loss of generality we may assume that a lies in the plane $x_2 = 0$. So, we assume $a_2 = 0$. We also assume that $a_1 > 0$. We can also assume that $a_3 > 0$ as $\langle a, a\rangle = \langle -a, -a\rangle$.

So, in the assumption that $a_2 = 0$, $a_1, a_3 > 0$ we have:

- If $\langle a, a\rangle > 0$, then $a_1^2 - a_3^2 > 0$ and $a_1 > a_3$ which implies that the set l_a intersects the cone at $a_1x_1 - a_3x_3 = 0$ i.e. where $|x_1| < |x_3|$, which is impossible inside the light cone.
- If $\langle a, a\rangle = 0$, then we get points where $|x_1| = |x_3|$, i.e. the point on the cone.
- If $\langle a, a\rangle < 0$, we get $|x_1| > |x_3|$ which gives a point inside the cone, so an inner point of the model.

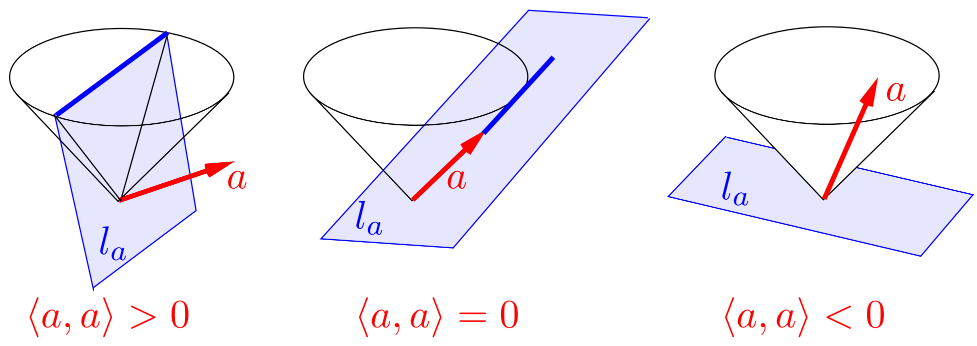


Figure 101: Geometric meaning of pseudo-scalar square.

Remark 7.3 (Reflections in the hyperboloid model)**.** One can show that a reflection r_a with respect to the line l_a in the hyperboloid model can be written as $r_a : x \mapsto x - 2\frac{\langle x,a\rangle}{\langle a,a\rangle} a$ (see Theorem 8.1 below), and that it preserves the pseudo-scalar product $\langle u, v\rangle$ for any vectors u, v (see the computation in HW 2.7).

Theorem 7.4. $\cosh^2 d(u,v) = \frac{\langle u,v\rangle^2}{\langle u,u\rangle\langle v,v\rangle}$ *for* $u, v \in \mathbb{H}^2$*, i.e. for* u, v *satisfying* $\langle u,u\rangle < 0$*,* $\langle v,v\rangle < 0$*.*

Proof. It is sufficient to prove the theorem for $u = (0,0,1)$ and $v = (x,0,z)$ (to see this we first apply a reflection inside the Klein disc which takes any given point to the centre of the disc, and then we project it to the hyperboloid; we can also apply a rotation about the centre (a composition of two reflections) to ensure that one coordinate is zero). See Fig. 102. Notice, that the isometries applied above do not affect the left-hand side in the theorem (being isometries), and they do not affect the right-hand sides in view of Remark 7.3 (being a composition of reflections).

For $u = (0,0,1)$ and $v = (x,0,z)$, the right-hand side is as follows:

$$\frac{\langle u,v\rangle^2}{\langle u,u\rangle\langle v,v\rangle} = \frac{(-z)^2}{-1\cdot(x^2 - z^2)} \overset{x^2-z^2=-1}{=} z^2.$$

To compute the left-hand side, we first compute the distance $d(u,v)$ by the definition:

$$d(u,v) = \frac{1}{2} ln \left| \frac{1-0}{1-\frac{x}{z}} \Big/ \frac{-1-0}{-1-\frac{x}{z}} \right| = \frac{1}{2} ln \left| \frac{z+x}{z-x} \right| =$$

$$\frac{1}{2} ln \left| \frac{(z+x)^2}{z^2 - x^2} \right| \overset{z^2-x^2=1}{=} \frac{1}{2} ln (x+z)^2 = ln(x+z).$$

From this we conclude $e^{d(u,v)} = x + z$, which implies

$$\cosh^2 d(u,v) = \left(\frac{e^{d(u,v)} + e^{-d(u,v)}}{2}\right)^2 = \left(\frac{x+z+\frac{1}{x+z}}{2}\right)^2 = \left(\frac{z + \frac{x^2+xz+1}{x+z}}{2}\right)^2$$

$$\overset{x^2+1=z^2}{=} \frac{1}{4}\left(z + \frac{z(x+z)}{x+z}\right)^2 = \frac{1}{4}(2z)^2 = z^2.$$

□

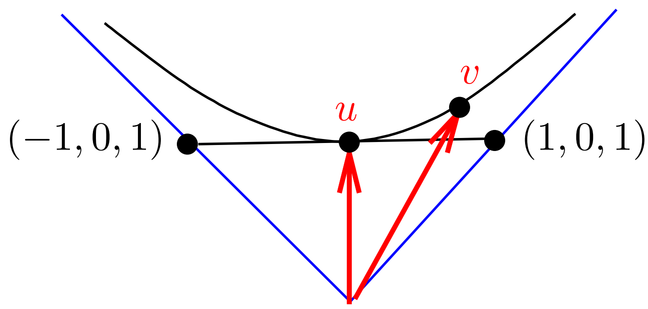


Figure 102: To the proof of Theorem 7.4 .

By a similar computation to the one in Theorem 7.4 one can prove the following theorem.

Theorem 7.5. *Denote by* $Q = Q(u,v) := |\frac{\langle u,v\rangle^2}{\langle u,u\rangle\langle v,v\rangle}|$. *Then*

(1) if $\langle u,u\rangle < 0$, $\langle v,v\rangle > 0$, *then* u *gives a point and* v *give a line* l_v *on* $\mathbb{H}^2$, *and* $\sinh^2 d(u, l_v) = Q$*;*

(2) if $\langle u,u\rangle > 0$, $\langle v,v\rangle > 0$ *then* u *and* v *define two lines* l_u *and* l_v *on* $\mathbb{H}^2$ *and*

- *if* $Q < 1$, *then* l_u *intersects* l_v *forming angle* φ *satisfying* $Q = \cos^2\varphi$*;*
- *if* $Q = 1$, *then* l_u *is parallel to* l_v*;*
- *if* $Q > 1$, *then* l_u *and* l_v *are ultra-parallel lines satisfying* $Q = \cosh^2 d(l_u, l_v)$.

Remark 7.6. One can ask whether the upper sheet of the hyperboloid together with the metric unduced from $\mathbb{R}^3$ is isometric to the hyperbolic plane, $\mathbb{H}^2$. The short answer is "NO". The reason is as follows:

> **Theorem (Hilbert).** *There is no isometric embedding of* $\mathbb{H}^2$ *to* $\mathbb{R}^3$. One can find the proof in Section 5.11 of the book by Do Carmo.

Notice, that some parts of the hyperbolic plane can be embedded to $\mathbb{R}^3$, and an example of such embedding is given by a pseudosphere (notice that this surface has infinite diameter, it has a finite volume).

7.3 References

- The exposition of Section 7 is partially based on parts of Lectures VI and XIII of Prasolov's book.
  Alternatively, see Section 5.2 of Prasolov, Tikhomirov.

- The relation between the four models (even five, including the hemisphere!) of hyperbolic geometry is described in Section 7 of
  J. W. Cannon, W. J. Floyd, R. Kenyon, W. R. Parry, *Hyperbolic Geometry.*

- There are also other models of hyperbolic plane, which we do not consider in this course. For some of them see *Hyperbolic Spaces* by John R. Parker.

- For the proof of Hilbert's theorem (on absence of isometric embeddings of $\mathbb{H}^2$ to $\mathbb{R}^3$) see Section 5.11 of

  M. do Carmo, *Differential Geometry of Curves and Surfaces*, Imprint Englewood Cliffs: Prentice-Hall (1976).

- Webpages, videos:

  - *Stereographic projection and models for hyperbolic geometry*, 3D toys to illustrate by Henry Segerman.
  - *Illuminating hyperbolic geometry*, Short video (4:25 min) by Henry Segerman and Saul Schleimer on projecting the hemisphere model to Klein disc, Poincare disc and upper half-plane.
  - Even games on hyperbolic field on Zeno Rogue webpage
    (this will be more related to the later parts of the course) .

8 Classification of isometries of $\mathbb{H}^2$

8.1 Reflections

Definition. A reflection r_l with respect to a hyperbolic line l is an isometry preserving the line l pointwise and swapping the half-planes.

Notice that as in Euclidean and spherical case, if $A' = r_l(A)$ then the line AA' is perpendicular to the line l (we can see this from the pair of congruent triangles $\triangle AMB \cong \triangle A'MB$ where $M = AA' \cap l$ and $B \in l$ any point $B \neq M$).

Example. In the Poincaré disc and upper half-plane models: reflections are represented by Euclidean reflections and inversions.
More precisely,

- In the Poincaré disc, the (Euclidean) reflection with respect to any diameter and an inversion with respect to any circle orthogonal to the absolute is hyperbolic reflection (as it is an isometry which preserves the corresponding hyperbolic line pointwise and swaps the half-planes).
- Similarly, in the upper half-plane, (Euclidean) reflections with respect to vertical lines and inversions with respect to the circles orthogonal to the absolute are hyperbolic reflections.

Example-exercise. In the Klein disc model: given A and l, one can construct $r_l(A)$. *Hint:* we know how to construct a midpoint of a segment and a line orthogonal to the given line and crossing it in a given point, see also Fig. 103.

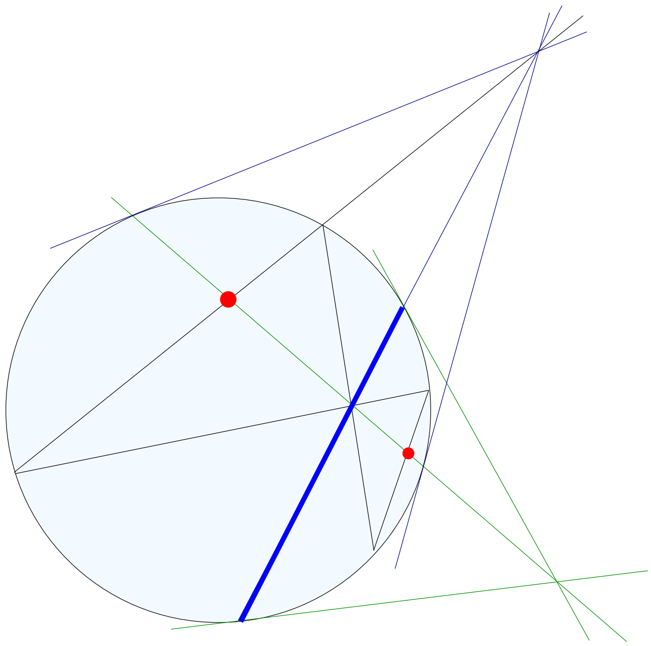


Figure 103: Reflecting a point with respect to a line in the Klein model.

Next, we consider a reflection in the hyperboloid model. Let $l = l_a = \{x \mid \langle x, a\rangle = 0\}$, where $\langle a, a\rangle > 0$.

Theorem 8.1. *In hyperboloid model: given a s.t.* $\langle a,a\rangle > 0$ *(i.e.* $\langle x,a\rangle = 0$ *defines a line* l_a*), the map* $r_a : x \mapsto x - 2\frac{\langle x,a\rangle}{\langle a,a\rangle}a$ *is the reflection with respect to the line* l_a*.*

Proof. First, notice that the map r_a preserves the pseudo-scalar product $\langle x,y\rangle$ (check this by a direct computation similar to one in HW 2.7). This implies that the hyperboloid is mapped by r_a to itself. Also, r_a is a linear transformation: indeed, if $a = (a_1,a_2,a_3)$ and $x = (x_1,x_2,x_3)$ then $r_a(x) = Ax$ where $A = I - \frac{2}{\langle a,a\rangle}B$ for the following matrix

$$B = \begin{pmatrix} a_1^2 & a_1a_2 & -a_1a_3 \\ a_1a_2 & a_2^2 & -a_2a_3 \\ a_1a_3 & a_2a_3 & -a_3^2 \end{pmatrix}.$$

Hence, r_a preserves the cross-ratio, and therefore preserves the distance. So, r_a is an isometry of the hyperboloid model. Furthermore, if $\langle a,x\rangle = 0$ then $r_a(x) = x$, which means that r_a preserves the line l_a pointwise. As $r_a(x) \neq x$ for $x \notin l_a$, we conclude that $r_a \neq id$. Hence, r_a is the reflection with respect to l_a.

□

8.2 Classification

Theorem 8.2. *Any isometry of* $\mathbb{H}^2$ *is a composition of at most 3 reflections.*

Proof. The proof of the theorem is very similar to the proof of its Euclidean analogue, Theorem 1.10.

Let $f \in Isom(\mathbb{H}^2)$ be an isometry. In view of Theorem 6.14 it is determined by an image of a flag. Consider a flag $F = (A,l,h^+)$ where $A \in \mathbb{H}^2$ is a point, l is a ray from A and h^+ is a choice of half-plane with respect to l. Let $f(F) = (A',l',(h')^+)$. Then f can be obtained as a composition of the following three reflections:

- Let r_1 be the hyperbolic reflection which takes A to A' (i.e. the reflection with respect to the perpendicular bisector of AA').
- Let r_2 be the reflection which preserves A' and takes $r_1(l)$ to l' (i.e. the reflection with respect to the angle bisector of the angle at A' formed by $r_1(l)$ and l').
- If $r_2 \circ r_1 \neq f$, let $r_3 = r_{l'}$ be the reflection with respect to l'.

Then the composition $r_3 \circ r_2 \circ r_1$ (or just $r_2 \circ r_1$) takes the flag F to $f(F)$ and hence, coincides with f.

□

Example 8.3. Let l_1 and l_2 be two lines in hyperbolic plane and r_1 and r_2 be the reflections with respect to them. What can we say about the composition $r_2 \circ r_1$?

- If $l_1 \cap l_2 = A \in \mathbb{H}^2$, then we can take the intersection point A to the centre of the Poincaré disc model, so that the hyperbolic reflections with respect to l_1 and l_2 will be represented by Euclidean reflections in the model, see Fig. 104, left. We conclude that $r_2 \circ r_1$ is a rotation about the intersection point A by the angle 2θ where θ is the angle between the lines.
- If $l_1 \cap l_2 = A \in \partial\mathbb{H}^2$, we can map the point A to the point ∞ of the upper half-plane model, see Fig. 104, middle. Then the reflections r_1 and r_2 are represented by Euclidean reflections with respect to vertical lines. Applying the a suitable isometry $az+b$, $a,b \in \mathbb{R}$ we can assume that the vertical lines are $Re(x) = 0$ and $Re(x) = 1$, and $r_2 \circ r_1$ is a translation $z \to z+2$.

- If $l_1 \cap l_2 = \emptyset$, i,e, the lines are divergent, then they have a common perpendicular. Let h be the common perpendicular, we can map it to the imaginary axis in the upper half-plane model. Then the lines l_1 and l_2 are represented by half-circles of centred at the origin, and after applying isometry $z \to az$ we may assume that these are half-circles given by $|z| = 1$ and $|z| = k$ (see Fig. 104, right). The reflections r_1, r_2 in this case are inversions with respect to the circles, i.e. are given by $z \to 1/\bar{z}$ and $z \to k^2/\bar{z}$. So, the composition can be written as $r_2 \circ r_1 : z \to k^2/\overline{(1/\bar{z})} = k^2 z$.

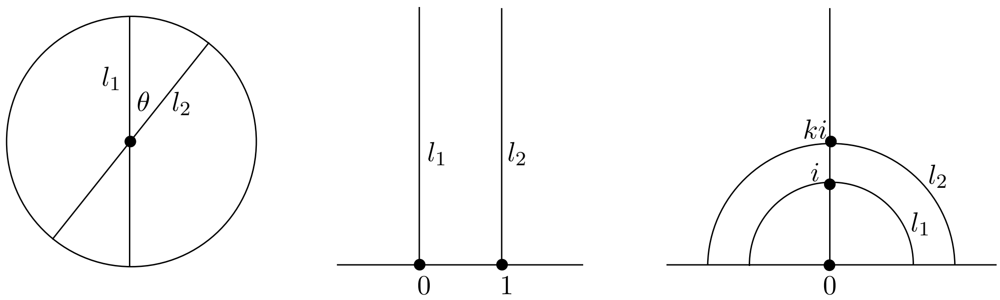


Figure 104: Compositions of two reflections.

Corollary 8.4. *A non-trivial orientation-preserving isometry of $\mathbb{H}^2$ has either 1 fixed point in $\mathbb{H}^2$, or 1 fixed point on the absolute, or two fixed points on the absolute.*

Proof. By Theorem 8.2 any hyperbolic isometry is a composition of at most 3 reflections. Hence, an orientation-preserving isometry is either identity or a composition of 2 reflections. Compositions of two reflections are considered in Example 8.3, so we see that the result depends on mutual position of the corresponding lines l_1, l_2 and

- if the lines intersect, $r_2 \circ r_1$ is a rotation and has a unique fixed point inside $\mathbb{H}^2$;
- if the lines are parallel, then $r_2 \circ r_1$ has a unique fixed point at the boundary (and no fix points inside $\mathbb{H}^2$);
- if the lines are divergent, then $r_2 \circ r_1$ has two fixed points at the boundary (and no fixed points inside $\mathbb{H}^2$).

$\square$

Definition 8.5. A non-trivial orientation-preserving isometry of $\mathbb{H}^2$ is called

- elliptic if it has 1 fixed point in $\mathbb{H}^2$,
- parabolic if it has 1 fixed point at $\partial\mathbb{H}^2$,
- hyperbolic if it has 2 fixed points at $\partial\mathbb{H}^2$.

Exercise. An orientation-reversing isometry of $\mathbb{H}^2$ is either a reflection or a glide reflection (see also HW 17.3).

Example 8.6. In the upper half-plane model, an orientation-preserving isometry is represented by the transformation $z \mapsto \frac{az+b}{cz+d}$ with $a, b, c, d \in \mathbb{R}$, $ad - bc = 1$. When is it elliptic? parabolic? hyperbolic?
We consider the fixed points, i.e. the solutions for

$$z = \frac{az+b}{cz+d},$$

this is equivalent to $cz^2 + z(d-a) - b = 0$, which has the solutions

$$z_{1,2} = \frac{a - d \pm \sqrt{(d-a)^2 + 4bc}}{2c}.$$

Taking in account that $ad - bc = 1$, we can rewrite the discriminant $D = (d-a)^2 + 4bc$ as follows:

$$D = (d-a)^2 + 4bc = (d-a)^2 + 4(ad-1) = (d+a)^2 - 4.$$

So, we conclude that

- when $|d+a| < 2$, we have $D < 0$ and the equation has two complex conjugate roots (exactly one in the upper half-plane);
- when $|d+a| = 2$, we have $D = 0$ and the equation has a unique (double) real root;
- when $|d+a| > 2$, we have $D > 0$ and the equation has two distinct real roots.

See Fig. 105. We conclude that the type of the transformation $z \to \frac{az+b}{cz+d}$ depends on the trace $a+d$ of the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$: it is elliptic if $|d+a| < 2$, parabolic if $|d+a| = 2$ and hyperbolic if $|d+a| > 2$.

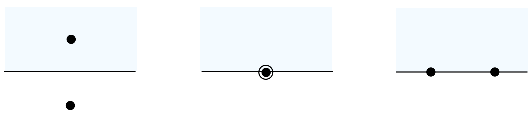


Figure 105: Fixed points of elliptic, parabolic and hyperbolic transformations.

Warning: To apply the trace criterion one first need to check that $ad - bc = 1$!

Remark 8.7 (Invariant sets for isometries)**.** The following sets preserved by elliptic, parabolic and hyperbolic isometries respectively (but not pointwise), see Fig. 106:

- elliptic isometries preserve concentric circles;
- parabolic isometries preserve curves called horocycles;
- hyperbolic isometries preserve curves called equidistant curves.

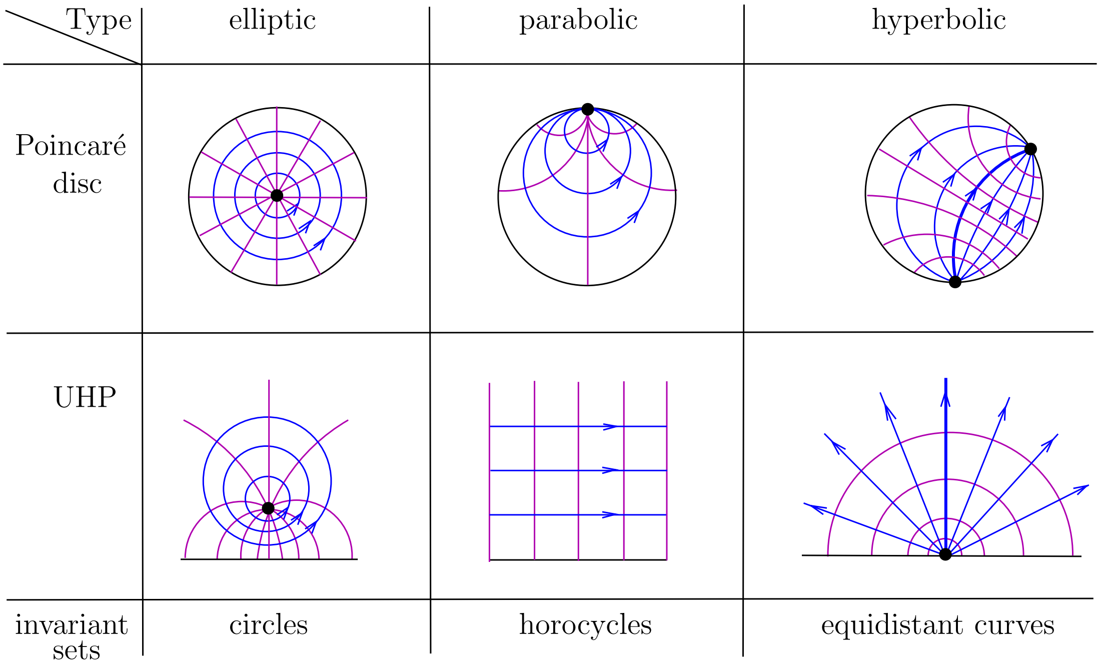


Figure 106: Elliptic, parabolic and hyperbolic transformations in Poincare disc and upper half-plane. Invariant sets (in blue) and families of lines orthogonal to them (magenta).

8.3 Horocycles and Equidistant curves.

Motivation. A *circle* is a set of points on the same distance from a given point.

Properties:

1. All lines through the centre are orthogonal to the circle.
2. The distance between two concentric circles γ and γ' is constant (i.e. given a point $A \in \gamma$ and a closest to A point $A' \in \gamma'$, the distance $d(A, A')$ does not depend on the choice of A).

Definition 8.8. A <u>horocycle</u> h is the following limit of circles:

- let $P \in \mathbb{H}^2$ be a point, and l be a ray from P;
- for $t > 0$ let $O_t \in l$ be a point s.t. $d(P, O_t) = t$;
- let $\gamma(t)$ be a circle centred at O_t of radius t;
- a horocycle is $h = \lim\limits_{t\to\infty} \gamma(t)$.
- The point $X = \lim\limits_{t\to\infty} O(t) \in \partial\mathbb{H}^2$ is called the <u>centre</u> of the horocycle h.

See Fig. 107.

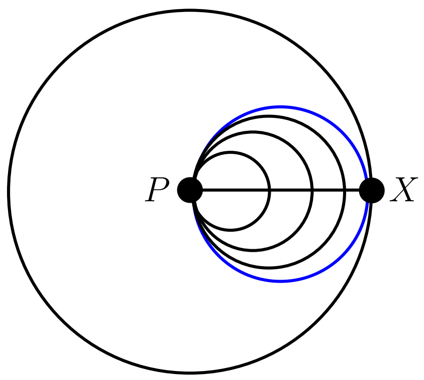


Figure 107: Horocycle as a limit of circles.

Remark 8.9 (Horocycles in the models)**.**

- In the Poincaré disc, every circle tangent to the absolute represents some horocycle (and every horocycle is a circle tangent to the absolute).
  To see that a given circle $\mathcal{C}$ tangent to the absolute is a horocycle, consider a point $P \in \mathcal{C}$. We can map P to the centre of the disc. Then $\mathcal{C}$ is mapped to the circle in Definition 107
- In the upper half-plane, a horocycle can be represented by a circle tangent to the absolute and by a line parallel to the absolute (the latter is tangent to the absolute at ∞).
- Isometries act transitively on the horocycles (in the upper half-plane, take the centre of the horocycle to ∞, then scale by $z \to az$ if needed).

Properties of horocycles:

1. All lines through the centre of the horocycle are orthogonal to the horocycle.
2. The distance between two concentric horocycles h and h' is constant.
   (i.e. given a point $A \in h$ and a closest to A point $A' \in h'$, the distance $d(A, A')$ does not depend on the choice of A).

Both of these properties are clear for the horocycle centred at ∞ in the upper half-plane, and hence, hold for any other horocycle.

Definition 8.10. An equidistant curve e to a line l is a locus of points on a given distance from l.

How to find the equidistant curve?

- Let l be the line 0∞ through 0 and ∞ in the upper half-plane, see Fig. 108, left;
- take a point A on the distance $d = d(A, l)$ from l;
- then the map $f : z \to kz$, $k \in R_+$ is an isometry preserving the line 0∞, hence $d(l, f(A)) = d(l, A) = d$;

- by varying k get a (Euclidean) ray OA which represents a part of the horocycle;

- reflection r_l with respect to l is also an isometry, so we get another ray;

- points inside the cone are closer to l than A, and points outside are further away.

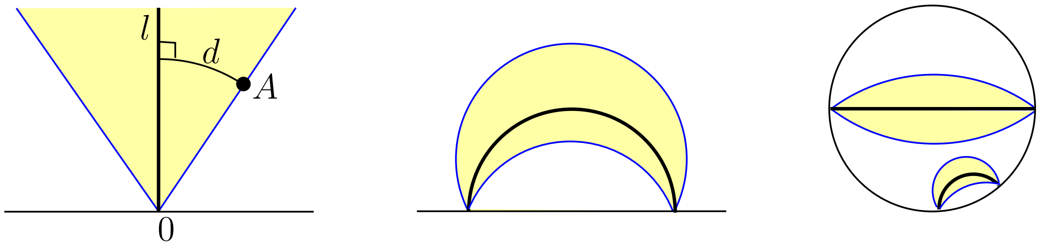


Figure 108: Examples of equidistant curves.

Examples. In UHP, if l is a vertical ray 0∞, then e is a union of to (Euclidean) rays from 0 making the same angle with l. If l is represented a half of Euclidean circle, then e is a "banana". In the Poincaré disc: also get banana, see Fig. 108.

Properties of equidistant curves:

1. All lines orthogonal to l are orthogonal to all equidistant curves for l.

2. Two equidistant curves to the same line stay on the same distance.

3. Distance between two equidistant curves e, e' is achieved along an orthogonal line.

To see these properties, we draw l as a vertical half-line through 0 in the upper half-plane, see Fig. 109. Then any equidistant curve to l is a union of two (Euclidean) rays from the origin O; any line orthogonal to l is represented by a semicircle centred at O, so it orthogonal to the equidistant curve.

Furthermore, if e and e' are two distinct equidistant curves, $A \in e$ is a point, and $A' \in e'$ is a point on the same line $m \perp l$, then $d(A, A')$ does not depend on the choice of $A \in e$: any other choice $A_1 \in e$ may be obtained from the initial by the isometry $z \to kz$ for some $k \in \mathbb{R}_+$, this isometry will take l to itself, so $d(A, l) = d(A', l)$.

Now, let e and e' be two equidistant curves for l lying on distances $d > d'$ respectively. Let $A \in e$ and let m be a line through A orthogonal to l, let $M = l \cap m$. Suppose that $d(A, e') = d(A, B)$ for some $B \in e'$, $B \notin m$. Let n be a line through B orthogonal to l and $N = n \cap l$. Then

$$d(A, l) = d(A, B) + d(B, l) = d(A, B) + d(B, N) > d(A, N),$$

where the last inequality holds in view of triangle inequality. We obtain $d(A, l) > d(A, N)$ which is impossible as $N \in l$.

Remark 8.11. 1. Given an elliptic, parabolic or hyperbolic isometry f of $\mathbb{H}^2$, for every point $A \in \mathbb{H}^2$ there exists a unique invariant curve of f passing through A (circle, horocycle or equidistant curve respectively). There is also a unique line orthogonal to all invariant curves of f and passing through A.

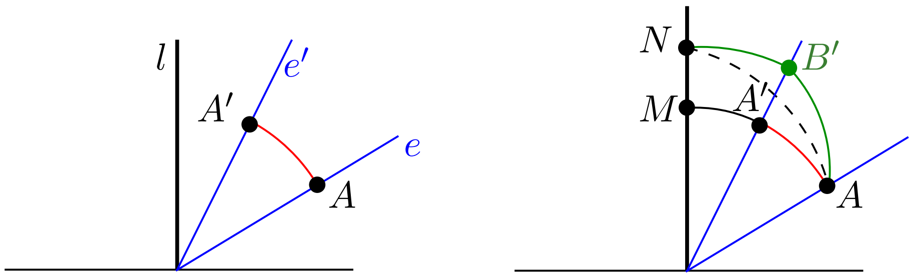


Figure 109: Properties of equidistant curves.

2. Representation of elliptic, parabolic and hyperbolic isometries as $r_2 \circ r_1$ is not unique: r_1 is a reflection with respect to any line from the orthogonal family, then there is a unique choice for r_2.

Remark 8.12. How do the horocycles and equidistant curves look like in the Klein model? A horocycle is by definition a limit of circles, i.e. sets represented by ellipses in the Klein model. So a horocycle is an ellipse tangent to the absolute. One can show (using a projection) that equidistant curves are pieces of ellipses intersecting the absolute.

8.4 Discrete reflection groups in S^2, $\mathbb{E}^2$ and $\mathbb{H}^2$ (NE)

(Non-examinable section!)

Recall from Definition 1.31 that an action $G : X$ of a group G on a metric space X is discrete if none of its orbits has accumulation points.

Example. A group generated by two reflections with respect to two intersecting lines (in S^2, $\mathbb{E}^2$ or $\mathbb{H}^2$) is discrete if and only if the angle α between the lines satisfies $\alpha = k\pi/m$, where $k, m \in \mathbb{Z}_+$ (see also Problems Class 2, Question 2.1). The rays lying on these lines bound a fundamental domain if and only if $\alpha = \pi/m$ for some $m \in \mathbb{Z}_{\geq 2}$, integer greater than 1 (see also Problems Class 2, Question 2.2).

In general, discrete group actions in hyperbolic case are not so easy to construct, and the aim of the current section is to describe one construction which produces infinite series of examples.

Definition. A reflection group is a group generated by (finitely many) reflections. In other words, we chose (finitely many) lines $l_1, \dots, l_n$ in $X = S^2$, $\mathbb{E}^2$ or $\mathbb{H}^2$ and consider a minimal group $G = \langle r_1, \dots, r_n \rangle$ containing the reflections $r_1, \dots, r_n$ with respect to these lines.

A reflection group $G = \langle r_1, \dots, r_n \rangle$ may contain infinitely many reflections: for any generating reflection r_j, $j \in 1, \dots, n$ and any element $g \in G$ (i.e. $g = r_{i_k} \circ \dots \circ r_{i_2} \circ r_{i_1}$) the element $h = g r_j g^{-1}$ is a reflection with respect to the line $g(l_j)$. We will call all these lines mirrors of G.

Question. When a reflection group G is discrete?

Suppose $G : X$ is a discrete action. Then there is a point $x \in X$ not lying on any mirror of G (otherwise, all orbits have accumulation points). Moreover, since the orbit of x should not have accumulation points, there is a disc centred at x containing no other elements from the same orbit. Hence, there is a (possibly smaller) disc around x not intersected by any mirror of G.

Let P be the largest connected set containing x and not intersected by any mirror of G. It is clear, that P is bounded by mirrors of G (otherwise, it could be larger). If P is also a bounded set in the metric space X, then P is bounded by finitely many mirrors (otherwise one can find an accumulation point of these mirrors and the action of G will not be discrete). So, if P is a bounded set, it is a polygon with finitely many sides. In view of the example above, we conclude that all angles of P are of the form π/k_i, $k_i \in \mathbb{Z}_+$ (otherwise the group is not discrete or P is crossed by some mirror).

Definition. Polygons with angles π/k_i, $k_i \in \mathbb{Z}_+$ are called Coxeter polygons. (And more generally, in higher dimensions, a Coxeter polyhedron is a polyhedron whose hyperfaces meet each other at angles π/k_i, $k_i \in \mathbb{Z}_+$).

Theorem. Every discrete reflection group in $X = S^2$, $\mathbb{E}^2$ or $\mathbb{H}^2$ has a fundamental domain, and if the fundamental domain is a bounded set then it is a Coxeter polygon. Moreover, every Coxeter polygon is a fundamental domain for some discrete reflection group.

. We have (partially) justified the first part of the theorem. The second part will follow from a more general construction and Poincare's theorem (which we will discuss but will not prove later).

The theorem implies that studying Coxeter polygons is (more or less) equivalent to studying reflection groups. Below we list what is known about them in S^2, $\mathbb{E}^2$ or $\mathbb{H}^2$.

S^2: A spherical Coxeter digon exists for any angle π/k, $k \in \mathbb{Z}_{\geq 2}$.

A spherical Coxeter triangle should have angles $(\pi/k, \pi/l, \pi/m)$, $k, l, m \in \mathbb{Z}_{\geq 2}$, integers larger than 1. As the sum of angles of a spherical triangle is larger than π, we conclude that $\frac{1}{k} + \frac{1}{l} + \frac{1}{m} > 1$. Assuming without loss of generality that $k \leq l \leq m$ we obtain $k = 2$. Then, if $l = 2$ we may have any value of $m \geq 2$. If $l = 3$ then $m = 3$, or 4, or 5. And if $l \geq 4$ there are no solutions.
We conclude that a spherical Coxeter triangle should have angles $(\pi/k, \pi/l, \pi/m)$ with (k, l, m) one of the following triples: $(2, 2, m)$, $(2, 3, 3)$, $(2, 3, 4)$, $(2, 3, 5)$.
(Notice that the corresponding reflection groups are symmetry groups of dihedron, tetrahedron, cube and octahedron, icosahedron and dodecahedron respectively. See also Section 2.9).

A spherical Coxeter quadrilateral or any other n-gon with $n \geq 3$ does not exist as the condition on the angles of a Coxeter polygon is not compatible with the angle sum inequality for spherical n-gons, i.e. $\sum \alpha_i \geq (n-2)\pi$.

$\mathbb{E}^2$: A Euclidean Coxeter triangle should have angles $(\pi/k, \pi/l, \pi/m)$ satisfying $\frac{1}{k} + \frac{1}{l} + \frac{1}{m} = 1$. As before, assuming $k \leq l \leq m$ one can easily check that this is only possible for the following values of (k, l, m): $(2, 3, 6)$, $(2, 4, 4)$, $(3, 3, 3)$.

A Euclidean Coxeter quadrilateral only can satisfy the angle sum condition when all its angles are right angles (i.e. it is a rectangle).

From the same angle sum condition we conclude that no Euclidean Coxeter polygon has more than 4 sides.

Remark: we can also have some unbounded domains serving as fundamental domains of discrete reflection groups, i.e. any domain bounded by one or two parallel lines, or two lines intersecting at an angle π/k, $k \in \mathbb{Z}_{\geq 2}$, or a domain bounded by two parallel lines and one line perpendicular to them.

$\mathbb{H}^2$: In hyperbolic plane, we have infinitely many Coxeter polygons.

For every triple (k, l, m) with $k, l, m \in \mathbb{Z}_{\geq 2}$ and $\frac{1}{k} + \frac{1}{l} + \frac{1}{m} < 1$ (i.e. any triple not mentioned above) there exists a hyperbolic Coxeter triangle.

For four or more integer numbers (different from $(2, 2, 2, 2)$) there exists a hyperbolic Coxeter polygon with the corresponding angles (and more over, one can show that there is an $(n-3)$-parametric family of n-gons with given n angles!)

And each of these Coxeter polygons is a fundamental domain for the corresponding discrete reflection group acting on $\mathbb{H}^2$!

So, there are many more Coxeter polygons (and hence discrete reflection groups) on hyperbolic plane than on the sphere or Euclidean plane.

Remark. The situation in higher dimension is even more different for the three types of spaces.

- S^d: Spherical Coxeter polytopes were classified by H. S. M Coxeter in 1934. There are finitely many of them in each dimension $d \geq 2$ (non-empty set of them); some of the corresponding reflection groups are symmetry groups of regular polytopes.
- $\mathbb{E}^d$: Euclidean were also classified by H. S. M. Coxeter. There are finitely many of them in each dimension (up to affine transformations, like the ones changing the proportions of sides in a rectangle).
- $\mathbb{H}^d$: The question of classification of hyperbolic Coxeter polytopes is still a widely open question, probably very far from a solution. The main things known to the moment are:
  - There are infinitely many examples of hyperbolic Coxeter polytopes in dimensions $d = 2, 3, 4, 5, 6$.
  - **Theorem** (Vinberg, 1984).
    Bounded hyperbolic Coxeter polytopes do not exist in $\mathbb{H}^d$ for any $d \geq 29$.
  - In dimensions $d = 7$ and $d = 8$, only 2 and 1 bounded examples are known respectively.
  - In dimensions $d = 9, ..., 29$ – nothing known at all!

8.5 References

- Material on types of isometries in hyperbolic geometry, and on horocycles and equidistant curves is based on Lecture IX of Prasolov's book.
  Alternatively, see Section 5.3 in Prasolov and Tikhomirov (p. 113–116).
- Tilings of S^2, $\mathbb{E}^2$ and $\mathbb{H}^2$ by triangles are described in Lecture X of Prasolov's book.
  Alternatively, see Section 5 of the Addendum in Prasolov and Tikhomirov (p. 185–187).
- One can find a detailed exposition concerning discrete reflection groups in
  E. B. Vinberg (Ed.), *Geometry II* , Encyclopaedia of Mathematical Sciences, Vol. 29, Springer-Verlag.
- Webpages:
  - Hyperbolic tessellations, including some printable hyperbolic tilings, webpage by David E. Joyce.

9 Geometry in modern maths - some topics (NE)

(Non-examinable section!!!)

The aim of this section is provide an overview of the zoo of different geometries we have seen before, to unify them and to put in the context of modern mathematics.

Note, that the aim as stated is very broad and any sense of completeness here would not be achievable within several lectures. So, it is just a brief discussion of a selection of topics, which depends on and reflects my personal preferences.

9.1 Taming infinity via horocycles

Martin Hairer (2014 Fields medallist) gave a Collingwood Lecture 2015 in Durham and spoke about "Renormalisation"; in particular, he said something close to: "If you have a diverging integral, subtract infinity (in a coherent way) and work then with finite values".

We illustrate this idea with horocycles:

- any point of a horocycle h is on infinite distance from the centre X of the horocycle;
- two concentric horocycles are on a finite distance from each other;
- choose "level zero" horocycle, and measure the (signed) distance to it, see Fig. 110.

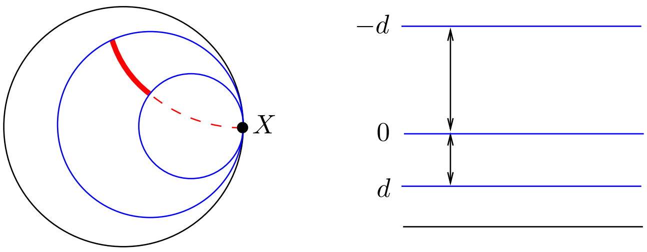


Figure 110: Distance between horocycles is finite.

Lambda-length of an infinite segment:

- Given $X, Y \in \partial H^2$, choose horocycles h_X and h_Y centred at these points.
- Let d_{XY} be the finite portion of the line XY lying outside of h_X and h_Y, see Fig. 111, left. It is a signed length, may be zero or negative if h_x intersects h_y, see Fig. 111, middle.
- Define $\lambda_{XY} = exp(d_{XY}/2)$, the lambda-length of XY.

Recall from Euclidean geometry Ptolemy Theorem (see Problems Class 5 for the proof):

Ptolemy Theorem. *In $\mathbb{E}^2$, a cyclic quadrilateral $ABCD$ satisfies*

$$|AC| \cdot |BD| = |AB| \cdot |CD| + |AD| \cdot |BC|.$$

There is a hyperbolic analogue of this theorem (see Fig. 111, right):

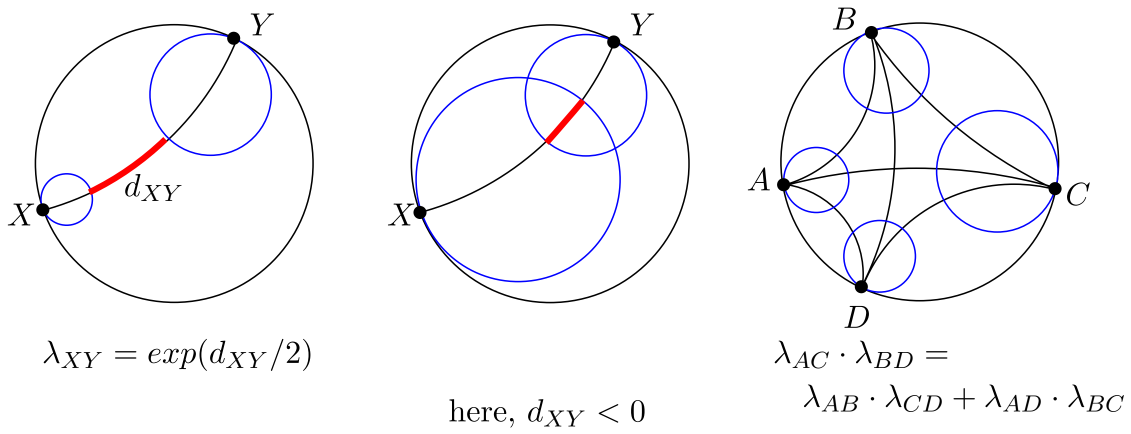


Figure 111: Lambda-length and hyperbolic Ptolemy Theorem.

Theorem 9.1 (Hyperbolic Ptolemy Theorem)**.** *For a hyperbolic ideal quadrilateral $ABCD$, choose any horocycles centred at A, B, C, D. Then*

$$\lambda_{AC} \cdot \lambda_{BD} = \lambda_{AB} \cdot \lambda_{CD} + \lambda_{AD} \cdot \lambda_{BC}.$$

Remark.

1. The proof of the hyperbolic Ptolemy Theorem can be done by a computation in the upper half-plane (omitted). See also [35].

2. The identity does not depend on the choice of the horocycles: if we change one horocycle taking another horocycles on distance d, then all lengths of arcs with endpoint at the centre of that horocycle will change by the same value d, and hence all lambda-length of the corresponding arcs will multiply by $exp(d/2)$. So, in the Ptolemy relation all summands of the identity will be multiplied by $exp(d/2)$, and the relation will be preserved.

3. This (together with applying an isometry to the quadrilateral) implies that it is sufficient to check the identity for the configuration shown in Fig. 112, where the identity rewrites as $\lambda_{AC} = 1 + \lambda_{BD}$.

4. Why do we care?

   - Given an ideal triangle A_1, A_2, A_3 and $c_{12}, c_{23}, c_{31} \in \mathbb{R}_+$ there exists a unique choice of horocycles such that $\lambda_{A_iA_j} = c_{ij}$ (we leave this as an exercise).
   - So, given a triangulated polygon $A_1A_2 \dots A_n$ and numbers $\lambda_{ij} > 0$ associated to the diagonals and sides of the polygon in the triangulation, one can find a unique hyperbolic metric on the corresponding surface (and a unique choice of horocycles) so that lambda lengths of the arcs of triangulation coincide with the given numbers λ_{ij}. In other words, the set of numbers $\{\lambda_{ij}\}$ associated to the arcs of the triangulation provides the coordinates on the space of all (decorated) hyperbolic structures on this polygon.

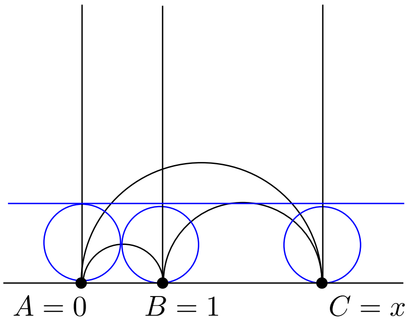


Figure 112: Diagram for the proof of hyperbolic Ptolemy Theorem.

- One can retell the same story about any surface which one can triangulate into ideal polygons.
- Using Ptolemy relation (several times if needed) we can compute a lambda length of any diagonal in the polygon as a function of initial values λ_{ij} associated to the arcs of the triangulation.
- This describes one of the connections of hyperbolic geometry to a recent theory of Cluster Algebras, which was introduced by Fomin and Zelevinsky in 2002 and turned out to be connected to numerous fields in mathematics and mathematical physics (including combinatorics of polytopes, representation theory, integrable systems). See here for a short introduction.

9.2 Three metric geometries: S^2, $\mathbb{E}^2$, $\mathbb{H}^2$, unified

One can consider a sphere of growing radius r, which eventually, when it is very big, is approaching a plane. Notice that the distance on the sphere of radius r is given by

$$d(A, B) = r \cdot \angle AOB.$$

Similarly, we can introduce a parameter into the distance function on the hyperbolic plane (say in the Klein disc model):

$$d(A, B) = R \cdot \frac{1}{2}|ln[A, B, X, Y]|,$$

so that when R is big the geometry is very similar locally to Euclidean one, while when R is smaller it gets "more and more hyperbolic", see Fig. 113.

Remark 9.2 (About curvature)**.** More precisely, this is quantified by a notion of curvature - which is very large and positive for a sphere of small radius, then decreases to 0 when the radius of the sphere grows, then is zero for a plane, and small negative for hyperbolic plane defined with large parameter R, and very large negative for a hyperbolic plane defined with small "radius" R. The standard hyperbolic plane, which we considered in the course, has constant curvature -1. We will not focus on the notion of curvature in this course, but one can find more about it in any course of Differential Geometry, see for example [7].

Below, we will show that spherical geometry (of any radius) can be unified with hyperbolic geometry (again of any radius). This will explain why we have obtained similar formulae in different geometries.

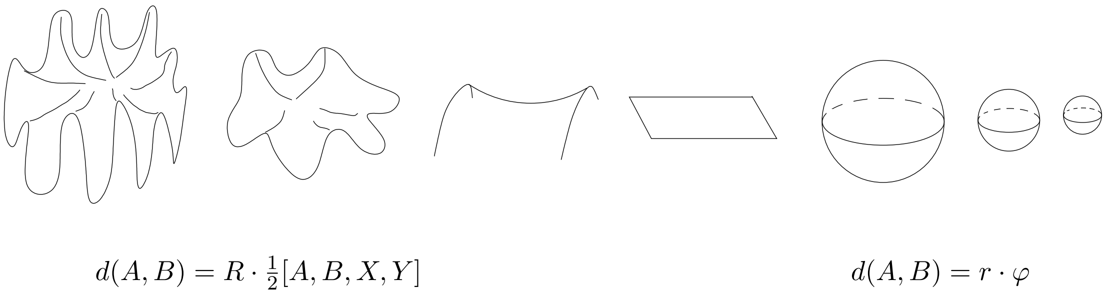


Figure 113: Spheres of growing radius r approach a plane – and similarly, hyperbolic planes of "growing radius R" approach the same plane.

We will use complex projective geometry to show that the distance on S^2 can be written as

$$d(A,B) = \pm\frac{r}{2i}|ln[A,B,X,Y]|.$$

We will take two points A and B on a sphere of radius r, assuming that $\angle AOB = \varphi$.

- The sphere of radius r is given by an equation $x_1^2 + x_2^2 + x_3^2 = r^2$. For this computation, we will assume that $x_i \in \mathbb{C}$, so $x = (x_1, x_2, x_3) \in \mathbb{C}^3$, or more precisely $x \in \mathbb{C}P^2$, i.e. a triple (x_1, x_2, x_3) is considered up to multiplication by a (non-zero) complex number.
- In case of $\mathbb{H}^2$ we consider the hyperboloid as a sphere $x_1^2 + x_2^2 + x_3^2 = -R^2$ of imaginary radius iR, rewriting this for $x_3' = ix_3$ we get exactly the hyperboloid model considered before.)
- In the hyperboloid model of $\mathbb{H}^2$, the distance is expressed through the cross-ratio $[A,B,X,Y]$ (where X and Y correspond to the endpoints of the lines through the points A and B).
- For the sphere, it is clear what are A and B (the points), but what would be the objects corresponding to X and Y?

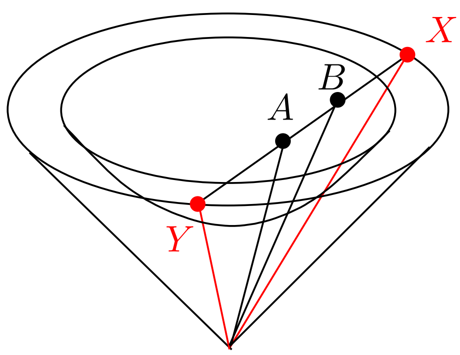


Figure 114: $\{X,Y\} = \Pi_{AB} \cap \{\langle x,x \rangle = 0\}$.

- To find the points X, Y we use the same rule as in the hyperboloid model:

$$\{X, Y\} = \Pi_{AB} \cap \{\langle x, x\rangle = 0\},$$

i.e. the intersection of the plane spanned by AB with the cone, see Fig. 114.

- Both in spherical and hyperbolic case, the plane through the points $a = (a_1, a_2, a_3)$ and $b = (b_1, b_2, b_3)$ is given by

$$(a_1 + \lambda b_1, a_2 + \lambda b_2, a_3 + \lambda b_3),$$

where $\lambda \in \mathbb{C} \cup \infty$ (the points a and b correspond to $\lambda = 0$ and $\lambda = \infty$ respectively).

- Intersection of this plane with the cone $\langle x, x\rangle = 0$ gives

$$(a_1 + \lambda b_1)^2 + (a_2 + \lambda b_2)^2 + (a_3 + \lambda b_3)^2 = 0.$$

- Taking in account $\langle a, a\rangle = r^2 = \langle b, b\rangle$ and $\langle a, b\rangle = r^2 \cos\varphi$ this gives

$$1 + 2\lambda cos\varphi + \lambda^2 = 0.$$

- Solving the equation for λ we get $x = \lambda_1$ and $y = \lambda_2$:

$$\lambda_{1,2} = cos\varphi \pm i \sin\varphi,$$

which allows to compute the points x and y.

- $[a, b, x, y] = [0, \infty, \lambda_1, \lambda_2] = \frac{\lambda_1 - 0}{\lambda_1 - \infty} \Big/ \frac{\lambda_2 - 0}{\lambda_2 - \infty} = \frac{\lambda_1}{\lambda_2} = exp(\pm 2i\varphi)$, from this we conclude that $ln[a, b, x, y] = \pm 2i\varphi$ and

$$\varphi = \pm \frac{r}{2i} |ln[a, b, x, y]|.$$

- Finally, we get for the spherical case:

$$d(A, B) = \pm \frac{r}{2i} |ln[A, B, X, Y]|.$$

Notice that by considering a sphere of imaginary radius $r = iR$ we recover the distance formula from the Klein model of hyperbolic geometry, $d(A, B) = \pm \frac{R}{2} |ln[A, B, X, Y]|$.

Remark 9.3. This explains appearance of similar formulae in spherical and hyperbolic geometries, in particular, this gives a proof of the second cosine law in the hyperbolic case.

Different geometries of constant curvature can be explicitly compared as in the following theorem:

Theorem 9.4 (Comparison Theorem, Aleksandrov-Toponogov)**.** *Given $a, b, c \in \mathbb{R}_{\geq 0}$ such that $a + b < c$, $a + c < b$ and $b + c < a$, consider triangles in $\mathbb{H}^2, \mathbb{E}^2$ and S^2 with sides a, b, c. Let $m_{\mathbb{H}^2}$, $m_{\mathbb{E}^2}$ and m_{S^2} be the medians connecting C with the midpoint of AB in each of the three triangles (see Fig. 115). Then $m_{\mathbb{H}^2} < m_{\mathbb{E}^2} < m_{S^2}$.*

The proof uses technique of Jacobi fields (which you can learn in Riemannian Geometry module). See *Toponogov's Theorem and Applications* by Wolfgang Meyer for the proof and applications.

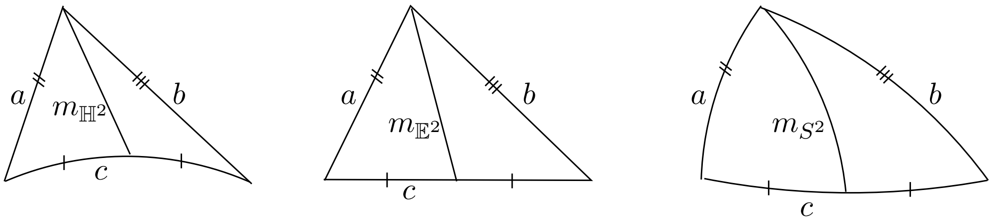


Figure 115: Comparison Theorem: $m_{\mathbb{H}^2} < m_{\mathbb{E}^2} < m_{S^2}$

9.3 Discrete groups of isometries of $\mathbb{H}^2$: Examples

Hyperbolic plane is rich with examples of discrete groups of isometries. At the same time it is not obvious even how to start to construct examples. We will describe one construction with the rough idea like:

- take a good polygon;
- find a good pairing of its sides;
- this will provide a tessellation of the hyperbolic plane, and hence a discrete group acting on it.

Assumption: Let $G \in Isom^+(\mathbb{H}^2)$ be a discrete group. We will assume that it has a fundamental domain F s.t. F is a polygon with finitely many sides.

Then we will get a tiling of $\mathbb{H}^2$ (in a similar manner to how copies of squares can tile a Euclidean plane).

By this we have described what do we mean by a *good polygon*, now we need to formulate what is a *good side pairing*. For this we will need several steps:

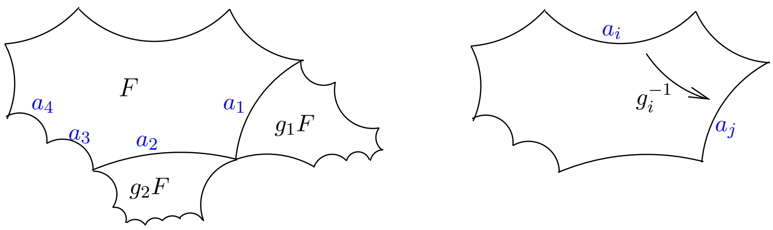


Figure 116: Side pairing.

Constructing a side pairing:

- Denote $a_1, \dots, a_n$ the sides of the fundamental domain F (there are finitely many by assumption).
- We have a tiling of $\mathbb{H}^2$ by copies of the fundamental domain F. For every side a_i of F consider the adjacent fundamental domain g_iF, $g_i \in G$, $g_i \neq id$ such that $a_i = F \cap g_iF$
- Note that $g_i^{-1}(a_i) \in F$, i.e. $g_i^{-1}(a_i)$ is another side of F, some a_j, so that we get a map $g_i^{-1} : a_i \to a_j$. As $g_j^{-1} = g_i : a_j \to a_i$, we get a pairing for all sides of F.

Example: Let $G : \mathbb{E}^2$ be the group generated by two shifts $g_1 : z \to z + 1$ and $g_2 : z \to z + i$. Then the fundamental domain for this group is a square, and the side pairing it realised by g_1 and g_2 (g_1 is pairing vertical sides of the square while g_2 does it for horizontal ones).

Constructing a graph Γ:

- Let $A_1, \dots, A_n$ be the vertices of F. Every g_i takes two adjacent vertices A_iA_{i+1} to other two vertices.

- Consider an oriented graph Γ:
  vertices of Γ are vertices A_i of F,
  edges of Γ are side pairings: $A_i \to A_j$ if $g_i(A_i) = A_j$

Example: For the group $G : \mathbb{E}^2$ generated by $g_1 : z \to z + 1$ and $g_2 : z \to z + i$ the graph Γ has 4 vertices, connected as in Fig. 117, right.

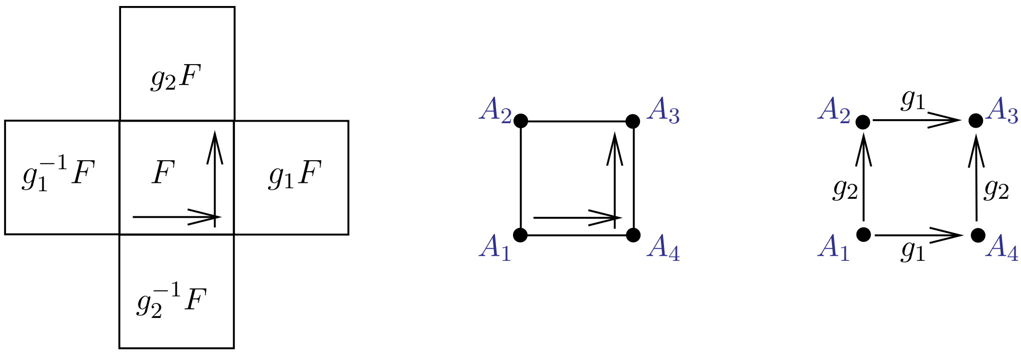


Figure 117: Side pairing for the group $G : \mathbb{E}^2$ together with the graph Γ.

Properties of Γ and equivalent vertices:

- Notice that every vertex A_i is incident to exactly two edges of Γ (one coming from the side pairing of A_iA_{i+1} and another from $A_{i-1}A_i$).

- We conclude that Γ consists of finitely many cycles.

- Vertices in one cycles are called equivalent.
  (These vertices lie in one orbit of $G = \langle g_1, \dots g_n \rangle$).

- Consider one cycle. Relabel the vertices and the maps so that $A_1, \dots, A_K$ be consecutive vertices in one cycle and $g_i(A_i) = A_{i+1}$.

Lemma. Let $A_1, \dots, A_k$ make one cycle, so that $g_i(A_i) = A_{i+1}$, $g_k(A_k) = A_1$, where g_i are side pairings of F and $A_1, \dots, A_k \in \mathbb{H}^2$ (but not $\partial\mathbb{H}^2$). Then $g = g_k g_{k_1} \dots g_1$ is a rotation about A_1 by the angle $\alpha_1 + \dots + \alpha_k$, where α_i is the angle of F at A_i.

Proof. We will prove the Lemma in additional assumption that $g_1, \dots, g_k$ are orientation-preserving.

- Since $g(A_1) = g_k g_{k_1} \dots g_1(A_1) = A_1$ and each of $g \in G \subset Isom^+(\mathbb{H}^2)$ is orientation-preserving, g is a rotation about A_1.

- Consider
$$f = R_{A_1}^{-\alpha_1} g_k \dots R_{A_3}^{-\alpha_3} g_2 R_{A_2}^{-\alpha_2} g_1,$$
where $R_{A_i}^{\alpha_i}$ is a rotation around A_i by α_i.

- Notice that $f(A_1) = A_1$ (g_1 takes A_1 to A_2, then $R_{A_2}^{\alpha_2}$ preserves it, then g_2 takes to A_3, etc. till g_k takes it back to A_1).

- Denote e_i' (resp. e_i) the edge of F preceding (resp. succeeding) the vertex A_i in the clockwise order. We will assume that e_1' is paired to e_2 by g_1, see Fig. 118, left.

- Then, $f(e_1) = e_1$. Indeed, $g_1(e_1)$ makes angle α_1 with e_2, so, $R_{A_2}^{\alpha_2} g_1(e_1)$ makes angle α_1 with e_2'. Then $g_2 R_{A_2}^{\alpha_2} g_1(e_1)$ makes angle α_1 with e_3 and $R_{A_3}^{\alpha_3} g_2 R_{A_2}^{\alpha_2} g_1(e_1)$ makes angle α_1 with e_3'. Continuing the same way till the end we get that $f(e_1)$ makes angle α_1 with e_1', so $f(e_1) = e_1$.

- Since $f(A) = A$ and $f(e_1) = e_1$ we conclude that $f = id$ (by uniqueness of isometry taking a flag to itself).

- **Exercise:** for every $h \in Isom^+(\mathbb{H}^2)$ holds $hR_A^\alpha = R_{h(A)}^\alpha h$.
(*Hint:* recall and use properties of fixed points, see Proposition 1.18).

- From the exercise above we have
$$\begin{aligned} id = f = R_{A_1}^{-\alpha_1} g_k \dots R_{A_3}^{-\alpha_3} g_2 R_{A_2}^{-\alpha_2} g_1 = R_{A_1}^{-\alpha_1} g_k \dots R_{A_3}^{-\alpha_3} g_2 g_1 R_{A_1}^{-\alpha_2} = \dots \\ = g_k \dots g_1 R_{A_1}^{-\alpha_1} \circ R_{A_1}^{-\alpha_k} \circ \dots \circ R_{A_3}^{-\alpha_3} \circ R_{A_1}^{-\alpha_2} = g_k \dots g_1 R_{A_1}^{-(\alpha_1+\alpha_2+\dots+\alpha_k)}, \end{aligned}$$
which implies that $g_k \dots g_1 = R_{A_1}^{\alpha_1+\alpha_2+\dots+\alpha_k}$.

$\square$

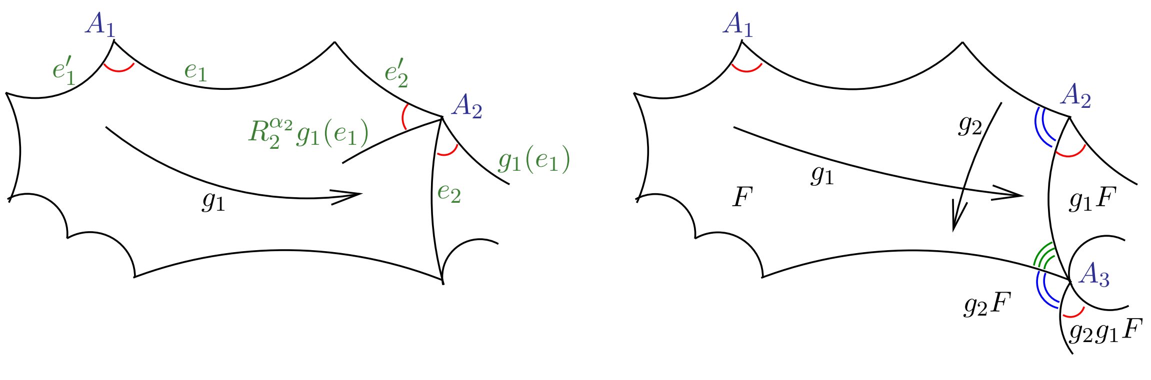


Figure 118: To the proof of the lemma (left) and the claim (right).

Claim. Polygons g_kF, $g_kg_{k-1}F$,…,$g_kg_{k-1}\dots g_1F$ have a common vertex A_1, with angles $\alpha_k, \alpha_{k-1}, \dots, \alpha_1$ at A_1.

The proof can be done inductively by considering the sets g_iF, $g_ig_{i-1}F$,. . . ,$g_ig_{i-1}\dots g_1F$ which share the vertex A_{i+1}. We skip the proof but illustrate the first steps in Fig. 118, right.

Corollary. Elements of the group $\langle g_1,\dots,g_n\rangle$ generated by side pairings tile the neighbourhood of A_1 if and only if $\alpha_1+\dots\alpha_k=2\pi/m$ for $m\in\mathbb{N}$.

The corollary provides us with a necessary condition, saying what we need to require from a side pairing if we want it to define a discrete action. The following theorem shows that this necessary condition is also sufficient:

Theorem 9.5 (Poincaré's Theorem.)**.** *Let $F\subset\mathbb{H}^2$ be a convex polygon, with finitely many sides, without vertices at the absolute, s.t.*

a) its sides are paired by isometries $\{g_1,\dots,g_n\}$;

b) angle sum in equivalent vertices is $2\pi/m_i$ for $m_i\in\mathbb{N}$.

Then

1) the group $G=\langle g_1,\dots,g_n\rangle$ is discrete;

2) F is its fundamental domain.

3) if all $g_i\in Isom^+(\mathbb{H}^2)$ then defining relations in G are vertex relations like $(g_kg_{k-1}\dots g_1)^{m_i}=e$.

We omit the proof of the theorem, see the lecture notes *Hyperbolic geometry* by Caroline Series for the details.

Remark.

1. Similar statement holds in $\mathbb{E}^2$ and S^2.
2. In $\mathbb{H}^2$ we can allow ideal vertices (with an extra condition that for the corresponding cycle the isometry $g_kg_{k-1}\dots g_1$ should be parabolic).

Examples: the following groups are discrete (more examples than we had in the lecture):

1. P= regular hexagon in $\mathbb{E}^2$, the group G generated by translations pairing the opposite sides of P, see Fig. 119, left.
   There are two classes of equivalent vertices, each containing 3 vertices. As $\alpha_i=2\pi/3$ we have that the total angle for every cycle is 2π, as required.
2. P= regular hexagon in $\mathbb{E}^2$, G generated by rotations by $2\pi/3$ about three non-adjacent vertices, see Fig. 119, middle.
   (This time, there is one classes of 3 vertices, with total angle $3\cdot 2\pi/3=2\pi$ and 3 classes of single vertex, with angles $2\pi/3$).
3. Regular hyperbolic octagon with angles $\pi/4m$, G generated by hyperbolic translations pairing the opposite sides of P, see Fig. 119, right.
   There is one cycle of vertices, with total angle $8\cdot\pi/4=2\pi$.

4. Let P be a polygon all whose angles are integer submultiples of π, i.e. the angle at the vertex A_i is π/m_i, $m_i \in \mathbb{N}$ (called Coxeter polygon), G generated by reflections with respect to the sides of P.
   In this case, every side is paired with itself, every vertex forms a separate class of equivalent vertices, and the cycle is given by $g_{i+1}g_i$ where g_i and g_{i+1} are reflections with respect to sides incident to the vertex A_i. Then $g_{i+1}g_i$ is a rotation around A_i by angle $2\alpha_i = 2\pi/m_i$. By Poincaré Theorem, the group $G = \langle g_1, \dots, g_n \rangle$ is discrete and$P$ is its fundamental domain.

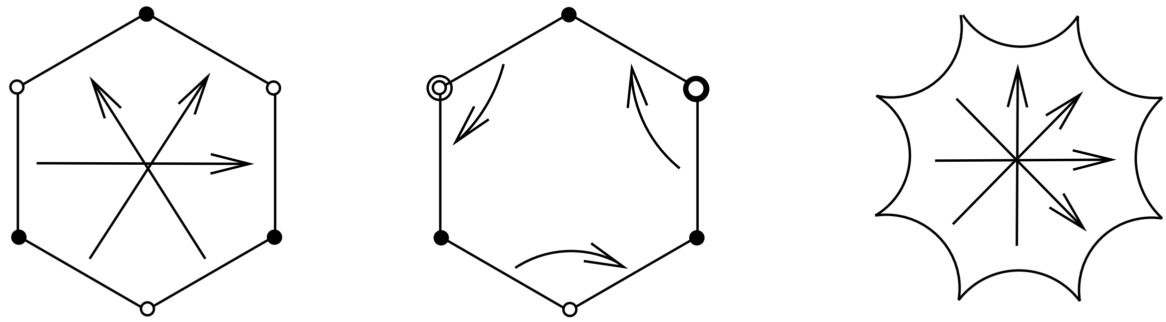


Figure 119: Examples of discrete groups defined by side parings of regular hexagon in $\mathbb{E}^2$ and regular octagon with angles $\pi/4$ in $\mathbb{H}^2$.

9.4 Hyperbolic surfaces

Definition 9.6. A surface S is called hyperbolic if every point $p \in S$ has a neighbourhood isometric to a disc on $\mathbb{H}^2$, see Fig. 120.

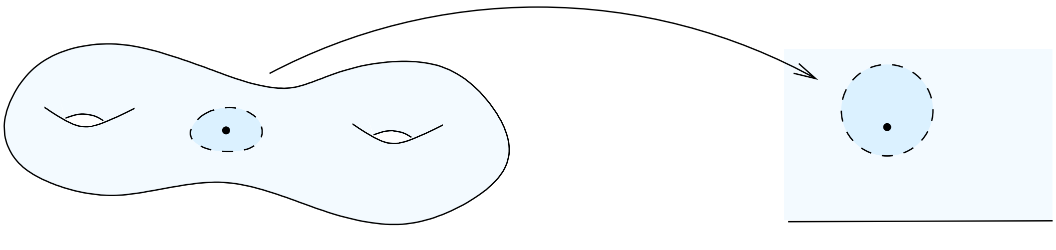


Figure 120: Hyperbolic surface.

How to construct such a surface? We will describe several ideas for that.

1. Glue from hyperbolic polygons.

Example: Euclidean torus can be glued from a square with identified opposite sides (see Example 1.35).

Example: Hyperbolic surface of genus 2 ("two holed torus") glued of a regular octagon with angles $\pi/4$ (opposite sides identified by hyperbolic translations, as in Fig. 119, right).

First of all, such a regular octagon exists: it is best viewed as a regular octagon with the centre at the centre of the Poincaré disc. Then a very small regular octagon is obtuse-angled, as a Euclidean regular octagon, then we increase the size of the octagon (distance from the centre), and can eventually turn it into ideal octagon, with zero angles. By continuity, somewhere in between there is a position where the regular octagon has angles $\pi/4$.

One can check, that when identifying the sides as in Fig. 119, right, we identify all 8 vertices, so all 8 angles of size $\pi/4$ each will be glued together to form a neighbourhood with a total angle 2π, as required. It is also easy to see, that every point on a side of the octagon will have a complete hyperbolic neighbourhood (again of angle 2π).

The sequence of identifications as in Fig. 121 shows that after gluing we obtain a sphere with two handles.

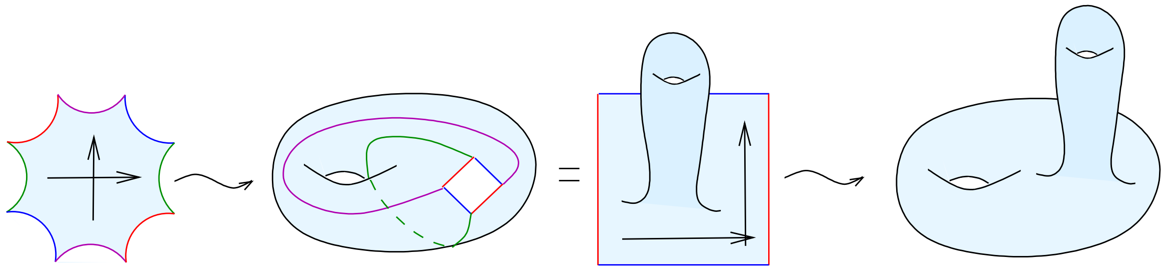


Figure 121: Gluing a hyperbolic surface with two handles from a hyperbolic octagon.

Remark. Alternatively, we can glue the sides of the octagon as shown in Fig. 122. This will result in the same topological surface (sphere with two handles), but the hyperbolic structure obtained on it will be different.

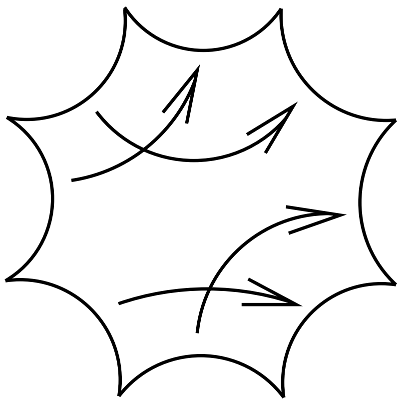


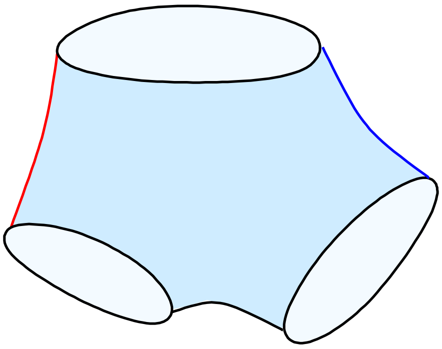
Figure 122: Another side pairing for the octagon.

One can play *H2Snake*, a game of snake on a hyperbolic surface of genus 2, to experience the geometry of this surface.

2. Pants decompositions.

A pair of pants is a sphere with three holes, see Fig. 123, left. A hyperbolic pair of pants may be glued from two right-angled hyperbolic hexagons, as in Fig. 123, right.

Gluing several pairs of pants by the boundaries, one can get (almost) every compact orientable topological surface (see Fig. 124) Exceptions are a sphere and a torus, which naturally carry spherical and Euclidean geometry, but not hyperbolic.



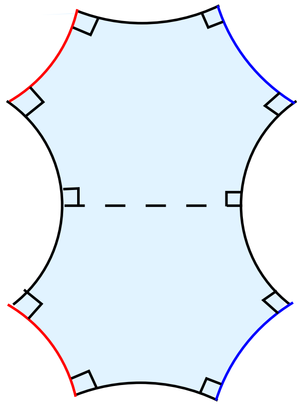


Figure 123: Hyperbolic pair of pants from two right-angled hexagons.

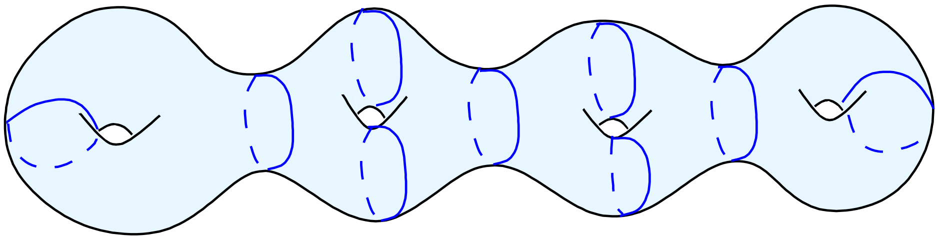


Figure 124: Pants decomposition of a surface.

3. Quotient of $\mathbb{H}^2$ by a discrete group.

Let $G : \mathbb{H}^2$ be a discrete action. Consider an orbit space $\mathbb{H}^2/G$.
Sometimes we get a hyperbolic surface, but not always (in particular, one needs to require that G contains no elliptic elements and no reflections).

Example. Consider a regular octagon with angles $\pi/4$ and one of two side pairing described above. Let the sides be paired by isometries g_1, g_2, g_3, g_4. Consider the group $G = \langle g_1, g_2, g_3, g_4 \rangle$ generated by g_1, g_2, g_3, g_4. Then $\mathbb{H}^2/G$ is a hyperbolic surface (topologically, sphere with two handles).

What will happen if G contains elliptic elements? One will obtain a surface with cone singularities.

Example. Consider a regular hyperbolic quadrilateral with angles $\pi/4$ and opposite sides identified. All four vertices will be identified. but the total angle at the resulting point will be $4 \cdot \pi/4 = \pi$ rather than 2π. Let $G = \langle g_1, g_2 \rangle$ be the group generated by the side pairing isometries. Then, $\mathbb{H}^2/G$ will be a torus with a cone point (of angle π) (this structure is called an orbifold rather than manifold).

4. Developing map.

Chose a point P on a hyperbolic surface. For each loop based at P (i.e. a path starting from P and ending at P) we construct an isometry on $\mathbb{H}^2$ in the following way:

- We can cover every point of the loop by a neighbourhood isometric to a disc in $\mathbb{H}^2$.

- Since the loop is a compact set, we can choose a finite subcover in any open cover, i.e. finitely many such neighbourhoods $U_1, \dots, U_n$ covering the whole loop.

- Chose a disc $\bar{U}_1 \in \mathbb{H}^2$ isometric to U_1.

- Then attach to $\bar{U}_1$ another disc $\bar{U}_2 \in \mathbb{H}^2$ in the same way as U_2 is attached to U_1.

- Continuing in the same way we attach discs $\bar{U}_3, \dots \bar{U}_n$ - and finally, since on the surface we started with a closed loop we attach one more disc $\bar{U}_{n+1}$ which is obtained from attaching U_1 to U_n. The map from the surface to $\mathbb{H}^2$ constructed in this way is called a developing map (or more precisely, the map is from the paths on the surface starting from the point P).

- Notice that in general $\bar{U}_{n+1}$ does not coincide with $\bar{U}_1$, but they are always congruent.

- Consider the isometry which takes $\bar{U}_1$ to $\bar{U}_{n+1}$ ("the isometry" since from the chain of gluing we know which point of the the boundary of $\bar{U}_1$ is mapped to which point of the boundary of $\bar{U}_{n+1}$).

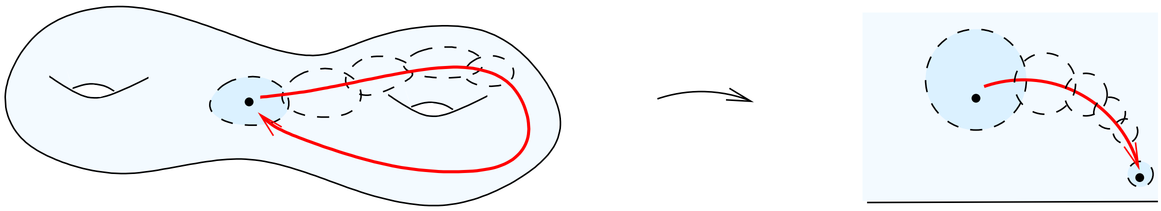


Figure 125: Developing map.

So, each loop on S gives rise to an isometry of $\mathbb{H}^2$. Consider a group G generated by all these isometries for $2g$ generating loops. Then one can show that G acts on $\mathbb{H}^2$ discretely, and $S = \mathbb{H}^2/G$ is its orbit space.

One can also show the following properties:

- $\forall g \in G, x \in \mathbb{H}^2$ if $gx = x$ then $g = id$ (such an action $G : X$ is called free).

- Every point $x \in \mathbb{H}^2$ has a neighbourhood containing no points of the orbit Gx.

So, we started with a hyperbolic surface and constructed a discrete group acting on $\mathbb{H}^2$.

5. Uniformisation theorem.

Theorem 9.7. *Any closed oriented hyperbolic (or Euclidean,or spherical) surface is a quotient of $\mathbb{H}^2$ (or $\mathbb{E}^2$, or S^2) by a free action of a discrete group.*

The proof can be derived from the Poincaré Theorem, see the lecture notes *Hyperbolic geometry* by Caroline Series.

9.5 Review via 3D hyperbolic space

I. Four models of $\mathbb{H}^3$

Ia. Upper half-space.

- Space: $\mathbb{H}^3 = \{(x, y, t) \in \mathbb{R}^3 \mid t > 0\}$.
- Absolute: $\partial\mathbb{H}^3 = \{(x, y, t) \in \mathbb{R}^3 \mid t = 0\}$.
- Hyperbolic lines: vertical rays and half-circles orthogonal to the absolute.
- Hyperbolic planes: vertical (Euclidean) half-planes and half-spheres centred at the absolute, see Fig. 126.
- Distance: $d(A, B) = |ln[A, B, X, Y]|$ (where X, Y are the ends of the line, and cross-ratio is computed in a vertical plane through A and B).
- Distance formula: $\cosh d(u, v) = 1 + \frac{|u-v|^2}{2u_3v_3}$,
  (here $|u - v|^2 = (u_1 - v_1)^2 + (u_2 - v_2)^2 + (u_3 - v_3)^2$).

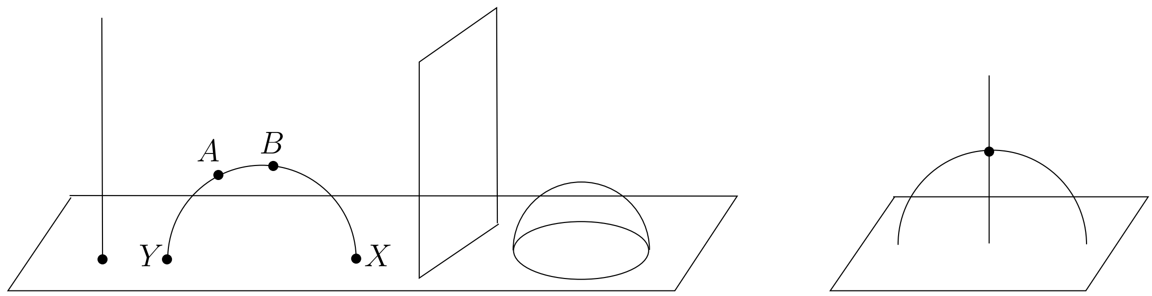


Figure 126: Upper half-space model of $\mathbb{H}^3$.

Isometries.

- Example: Hyperbolic reflections = (Euclidean) reflections with respect to the vertical planes and inversions with respect to the spheres centred at the absolute.
- $f \in Isom(\mathbb{H}^3)$ is determined by its restriction to the absolute:
  Through every point in $\mathbb{H}^3$ we can draw two distinct lines intersecting in that point, , see Fig. 126, right. So, if we know the images of the endpoints of the lines then we know the images of the lines and their intersection.
- $Isom(\mathbb{H}^3)$ is generated by reflections (every isometry is a composition of at most 4 reflections). Restrictions to $\partial\mathbb{H}^3$ are compositions of (Euclidean) reflections and inversions.
- $Isom^+\mathbb{H}^3 \cong M\ddot{o}b$.

Spheres, horospheres, equidistant surfaces.

- Spheres: Euclidean spheres (with another centre than in $\mathbb{E}^2$).

- Horospheres (limits of spheres): horizontal planes and spheres tangent to the absolute, see Fig. 127, middle.
  Horospheres are submanifolds in $\mathbb{H}^3$ isometric to $\mathbb{E}^2$ (with the same isometry group, same geodesics i.e. intersections of the horosphere h with the planes through its centre at the absolute).

- Equidistant surface (to a line): vertical cone (or banana for "half-circle" lines, see Fig. 127, right).

- Equidistant (to a plane Π represented by vertical half-plane): two (Euclidean) planes at the same angle to a vertical plane (at $\Pi \cap \partial\mathbb{H}^3$)

- Equidistant (to a plane Π represented by a hemisphere) or two pieces of spheres at the same angle to the sphere representing Π.

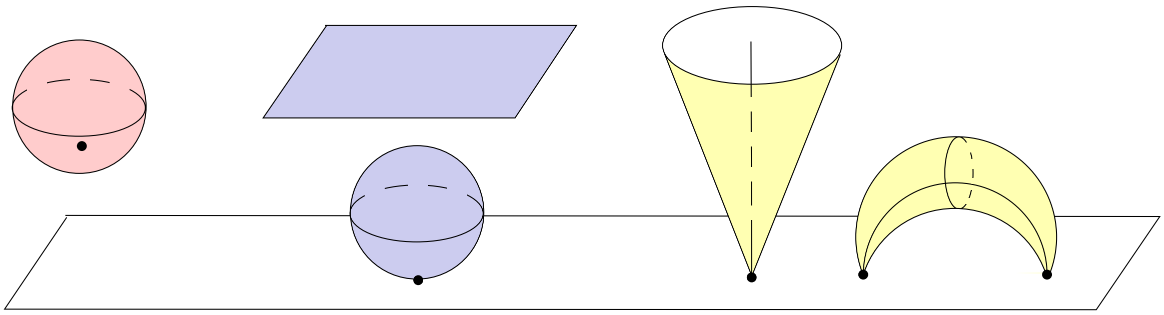


Figure 127: Spheres, horospheres and equidistant surfaces in the upper half-space model of $\mathbb{H}^3$.

Ib. Poincaré ball.

- Obtained by inversion from the upper half-space model.

- Space: $\mathbb{H}^3 = \{(x, y, t) \in \mathbb{R}^3 \mid x^2 + y^2 + t^2 < 1\}$.

- Absolute: $\partial\mathbb{H}^3 = \{(x, y, t) \in \mathbb{R}^3 \mid x^2 + y^2 + t^2 = 1\}$.

- Hyperbolic lines: parts of lines and circles orthogonal to $\partial\mathbb{H}^3$.

- Hyperbolic planes: parts of planes and spheres orthogonal to $\partial\mathbb{H}^3$.

- $d(A, B) = |ln[A, B, X, Y]|$
  (X, Y the ends of the line, cross-ratio computed in a plane).

Both Poincaré models are conformal: hyperbolic angles are represented by Euclidean angles of the same size.

Ic. Klein model.

- Space: $\mathbb{H}^3 = \{(x, y, t) \in \mathbb{R}^3 \mid x^2 + y^2 + t^2 < 1\}$.

- Absolute: $\partial\mathbb{H}^3 = \{(x, y, t) \in \mathbb{R}^3 \mid x^2 + y^2 + t^2 = 1\}$.

- Hyperbolic lines: chords. See Fig. 128, left.

- Hyperbolic planes: intersections with Euclidean planes.
- $d(A,B) = \frac{1}{2}|ln[A,B,X,Y]|$ (X,Y the ends of the line).
- Angles are distorted (except ones at the centre).
- Right angles are easy to control, see Fig. 128, middle and right.

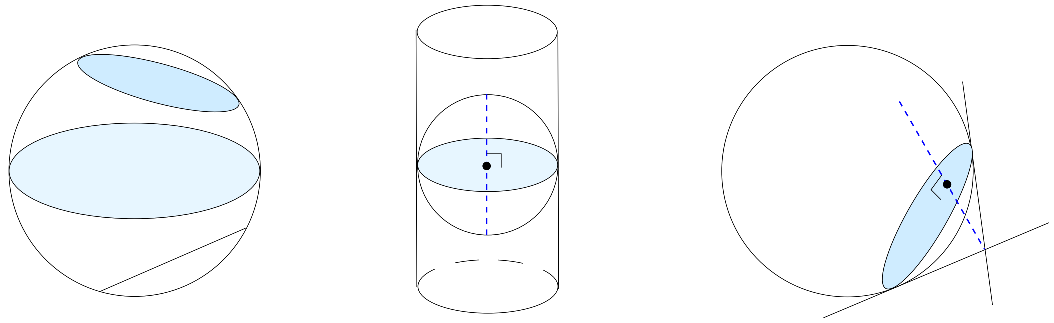


Figure 128: Klein model of $\mathbb{H}^3$: lines, planes and orthogonality.

Id. Hyperboloid model.

- Hyperboloid: $x_1^2 + x_2^2 + x_3^2 - x_4^2 = -1$, $x \in \mathbb{R}^4$.
- Pseudo-scalar product: $\langle x, y\rangle = x_1y_1 + x_2y_2 + x_3y_3 - x_4y_4$.
- Space: $\langle x, x\rangle = -1$.
- Absolute: $\langle x, x\rangle = 0$.
- Hyperbolic planes: $\langle x, a\rangle = 0$ for a s.t. $\langle a, a\rangle > 0$.
- $d(A,B) = \frac{1}{2}|ln[A,B,X,Y]|$ (cross-ratio of four lines).
- Formula: $\cosh^2(d(pt_1, pt_2)) = Q(pt_1, pt_2)$ where $Q\langle u, v\rangle = \frac{\langle u,v\rangle^2}{\langle u,u\rangle\langle v,v\rangle}$.

II. Orientation-preserving isometries of $\mathbb{H}^3$

- In the upper half-space, orientation-preserving isometries take planes to planes, i.e. circles lying on the absolute to circles on the absolute or lines on the absolute. Hence, they correspond to Möbius transformation of the absolute $\partial\mathbb{H}^3$:

$$\frac{az+b}{cz+d} \text{ with } z \in \partial\mathbb{H}^3, a, b, c, d \in \mathbb{C}, ad - bc \neq 0.$$

There are the following types of them:

  – Parabolic: 1 fixed point on $\partial\mathbb{H}^3$, the isometry is conjugate to $z \mapsto z + a$.
  – Non-parabolic: 2 fixed points on $\partial\mathbb{H}^3$, isometry is conjugate to $z \mapsto az$.
    - elliptic, $|a| = 1$, rotation about a vertical line.
    - hyperbolic, $a \in \mathbb{R}$, (Euclidean) dilation.
    - loxodromic, (otherwise), "spiral trajectory" =composition of rotation and dilation.

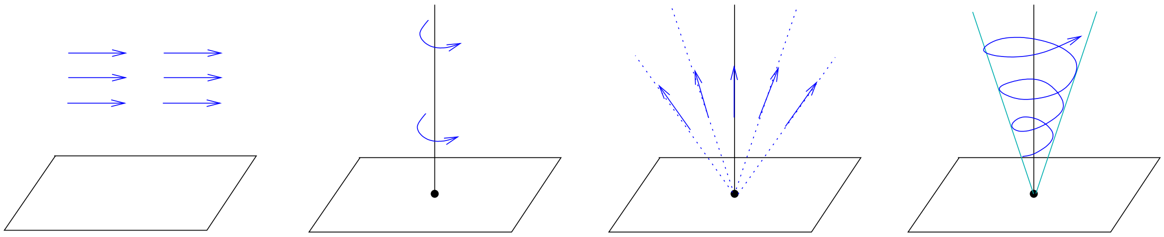


Figure 129: Isometries of $\mathbb{H}^3$: parabolic, elliptic, hyperbolic and loxodromic.

III. Some polytopes in $\mathbb{H}^3$

- Ideal tetrahedron. It is not unique up to an isometry! See Fig. 130, left.
  Exercise: There are 3 pairs of dihedral angles α, β, γ satisfying $\alpha + \beta + \gamma = \pi$).

- Regular right-angled dodecahedron.
  Indeed a regular dodecahedron of a very small size is almost Euclidean, so has obtuse angles. A very large dodecahedron (i.e. an ideal regular dodecahedron with all vertices on the absolute) has angles $\pi/3$. So, somewhere in between there is a right-angled dodecahedron.

- Right-angled ideal octahedron.
  Placing a vertex of an ideal octahedron to the infinity of the upper half-space model, one can see that it can only have equal angles when all angles are right, see Fig. 130, right.

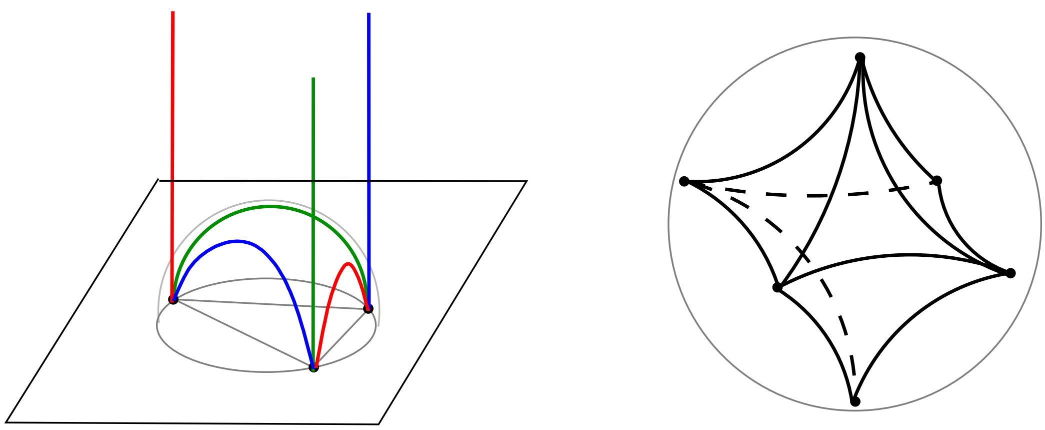


Figure 130: Hyperbolic polytopes: a tetrahedron and a right-angled ideal octahedron.

IV. Geometric structures on 3-manifolds

- We can glue 3-manifolds from polytopes...

- But we need to check that we obtain complete neighbourhoods around around edges and vertices.

- There is also a version of Poincaré theorem (which allows to construct discrete group actions in $\mathbb{H}^3$ given a suitable side pairing of a hyperbolic polyhedron).

V. Geometrisation conjecture

- William Thurston (1982): *all topological 3-manifolds are geometric manifolds*, i.e. every oriented compact 3-manifold without boundary can be cut into pieces having one of the following 8 geometries:
  S^3, $\mathbb{E}^3$, $\mathbb{H}^3$, $S^2 \times \mathbb{R}$, $\mathbb{H}^2 \times \mathbb{R}$, Nil, Sol and universal cover of $SL(2,\mathbb{R})$.
  (where Nil, Sol and $SL(2,\mathbb{R})$ are some exotic structures).

- (1982) William Thurston, (Fields medal, 1982)
  proved geometrisation conjecture for some manifolds called "Haken manifolds". In particular, all closed atoroidal Haken manifolds are hyperbolic (this statement is also called a hyperbolisation conjecture).

- (2003) Grigori Perelman, (Fields medal, 2006): general proof of the geometrisation conjecture.
  This also proves Poincaré conjecture:
  Every simply-connected closed 3-manifold is a 3-sphere. (Clay Millennium Prize).

Remark. In the *Not knot* 16-minute film (produced in 1991 by mathematicians at the Geometry Center at the University of Minnesota, directed by Charlie Gunn and Delle Maxwell) one can experience, how the hyperbolic 3-dimensional space feels from inside and how 3-dimensional hyperbolic manifolds arise as knot compliments.

9.6 References

Section 9 touches several vast fields in mathematics. The aim of this list of references is to give some possibility to start in each of these directions.

- Concerning lambda-lengths and hyperbolic version of Ptolemy theorem see

  - Anna Felikson, Ptolemy Relation and Friends, an AMR review.
  - Boris Springborn, *Ideal hyperbolic polyhedra and discrete uniformization*, Discrete Comput. Geom., 64(1):63–108, 2020. https://arxiv.org/pdf/1707.06848.pdf.
  - A similar idea based on horocycles (together with interplay between Euclidean geometry and the two disc models of hyperbolic geometry) is beautifully used in the following short paper:
    K. Drach, R. E. Schwartz, *A Hyperbolic View of the Seven Circles Theorem*.

- Concerning Aleksandrov-Toponogov comparison theorems, see

  - Wolfgang Meyer, *Toponogov's Theorem and Applications*.
  - Mikhael Gromov, *Sign and Geometric meaning of curvature*.
  - Jeff Cheeger, David G. Ebin, *Comparison theorems in Riemannian geometry*, North-Holland Mathematical Library, vol. 9, North-Holland, Amsterdam; American Elsevier, New York, 1975.

- To find more about fundamental domains of group actions, Poincaré theorem, and hyperbolic structure on surfaces, start with

  - Caroline Series, *Hyperbolic geometry*.
  - William Thurston, *The geometry and topology of three-manifolds*, Princeton University Mathematics Department (1979), lecture notes. here you will find separate pdf files of chapters.
  - Roice Nelson, Henry Segerman Visualising Hyperbolic Honeycombs. Describes the way to visualise 3-dimensional hyperbolic tilings. Includes many beautiful pictures (of tilings, indexed by their Schläfli symbols).

- Concerning Thurston's hyperbolization theorem, see the following books:

  - Mark Lackenby, *Hyperbolic Manifolds*. Lecture notes (2000). Continues here.
  - Bruno Martelli, *An Introduction to Geometric Topology*, Independently published, 488 pages, 3rd Edition, 2023.
  - Michael Kapovich, *Hyperbolic Manifolds and Discrete Groups. Lectures on Thurston's hyperbolization*.
  - Danny Calegary, *Chapter 2: Hyperbolic Geometry*, of a forthcoming book on 3-manifolds.

- And some general phylosophy:

  - William Thurston, On Proof and Progress in Mathematics, Bulletin of the American Math. Soc., Vol. 30, Number 2, April 1994, pp. 161–177.

- Videos, webpages, etc...
  - Isometry Classes of hyperbolic 3-space, webpage by Roice Nelson, includes animations of elliptic, hyperbolic, loxodromic and parabolic isometries.
  - And other hyperbolic things by Roice Nelson (including hyperbolic Rubik's cube, honeycombs, etc...)
  - H2Snake, a game of snake on a hyperbolic surface of genus 2.
  - Video comparing spherical, Euclidean and hyperbolic geometry (I don't know the author). Also, contains a link to videogames in hyperbolic plane.
  - 3-dimensional space webpage with describtion, illustration and animation of Thurston's eight geometries, project by Rémi Coulon, Sabetta Matsumoto, Henry Segerman, Steve Trettel.
  - Not knot, a short film produced by the Geometry Centre in U. of Minnesota, popularising Thurston's Geometrization Conjecture.
- Further reading on the topics touched in this section (several books and papers not necessarily accessible online...)

  - A. F. Beardon, *The Geometry of Discrete Groups*, Graduate Texts in Mathematics, 91. Springer Verlag (1983).
  - F. Bonahon, *Low-Dimensional Geometry: From Euclidean Surfaces to Hyperbolic Knots*. Student Mathematical Library, Volume: 49 (2009).
  - D.B.A.Epstein, *Geometric structures on Manifolds*, The Mathematical Intelligencer 14, Number1, (1982), p. 5–16.
  - S.Katok, *Fuchsian groups*, University of Chicago Press (1992).
  - W. P. Thurston, edited by Silvio Levy, *Three-dimensional geometry and topology*, Princeton Univ. Press (1997).

10 Bibliography and links

References

[1] J.W. Anderson, *Hyperbolic Geometry*, Springer (1999).

[2] A. Barvinok, *Combinatorics of Polytopes.* Lecture notes, University of Michigan (2010).

[3] A. F. Beardon, *The Geometry of Discrete Groups*, Graduate Texts in Mathematics,91. Springer Verlag (1983).

[4] F. Bonahon, *Low-Dimensional Geometry: From Euclidean Surfaces to Hyperbolic Knots.* Student Mathematical Library, Volume: 49 (2009).

[5] D. Calegary, *Chapter 2: Hyperbolic Geometry*, of a forthcoming book on 3-manifolds. University of Chicago (2015).

[6] J. W. Cannon, W. J. Floyd, R. Kenyon, W. R. Parry, *Hyperbolic Geometry.* Flavors of Geometry, MSRI Publications, Volume 31 (1997).

[7] M. do Carmo, *Differential geometry of curves and surfaces*, Imprint Englewood Cliffs: Prentice-Hall (1976).

[8] T. K. Carne, *Geometry and groups.* Lecture notes, Cambridge University.

[9] J. Cheeger, D. G. Ebin, *Comparison theorems in Riemannian geometry*, North-Holland Mathematical Library, vol. 9, North-Holland, Amsterdam; American Elsevier, New York (1975).

[10] H. S. M. Coxeter, Introduction to Geometry, Wiley (1963).

[11] T. Davis, *Inversion in a circle.*

[12] K. Drach, R. E. Schwartz, *A Hyperbolic View of the Seven Circles Theorem.*

[13] D.B.A.Eppstein, *Geometric structures on Manifolds*, The Mathematical Intelligencer 14, Number 1 (1982), p 5–16.

[14] A. Felikson, Ptolemy Relation and Friends, an AMR review.

[15] A. D. Gardiner, C. J. Bradley, *Plane Euclidean Geometry*, UKMT, Leeds (2012).

[16] M. J. Greenberg, *Euclidean and Non-Euclidean Geometries*, San Francisco: W. H. Freeman (2008).

[17] M. Gromov, *Sign and Geometric meaning of curvature*, IHES (1990).

[18] N. Hitchin, *Projective Geometry*, Lecture notes. Chapters 1, 2, 3, 4. University of Oxford.

[19] G. Jones, *Algebra and Geometry*, Lecture notes, University of Southampton.

[20] M. Kapovich, *Hyperbolic Manifolds and Discrete Groups. Lectures on Thurston's hyperbolization.* University of Utah (1994).

[21] S. Katok, *Fuchsian groups*, University of Chicago Press (1992).

[22] Kiselev's Geometry, *Book I. Planimetry* and *Book II. Stereometry.* (Adopted from Russian by Alexander Givental). Publisher's page of the book.

[23] Mark Lackenby, *Hyperbolic Manifolds.* Lecture notes (2000). Continues here.

[24] B. Martelli, *An Introduction to Geometric Topology*, Independently published, 488 pages, 3rd Edition, 2023.

[25] W. Meyer, *Toponogov's Theorem and Applications.* Lecture Notes, College on Differential Geometry, Trieste (1989).

[26] J. R. Parker, *Hyperbolic Spaces.* The Jyväskylä Notes, University of Durham (2007).

[27] A. Petrunin, *Euclidean plane and its relatives. A minimalist introduction.* Lecture notes, Pennsylvania State University (2020).

[28] N. Peyerimhoff, *Geometry III/IV*, Lecture notes (Section 1).

[29] V. V. Prasolov, *Non-Euclidean Geometry.* Lecture notes, Math in Moscow, Publishing Company of Moscow Centre for Continuous Mathematical Education, Moscow (2006).

[30] V. V. Prasolov, V. M. Tikhomirov, *Geometry.* American Mathematical Society, Providence, Rhode Island, 2001.

[31] E. Rees, *Notes on Geometry*, Universitext, Springer (2004).

[32] R. Schwartz, S. Tabachnikov, *Elementary Surprises in Projective Geometry*, arXiv:0910.1952v2.

[33] C. Series, *Hyperbolic geometry*, Lecture notes, University of Warwick, 2008.

[34] A. B. Sossinsky, *Geometries*, Providence, RI : American Mathematical Soc. (2012).

[35] B. Springborn, *Ideal hyperbolic polyhedra and discrete uniformization*, Discrete Comput. Geom., 64(1):63–108, (2020). https://arxiv.org/pdf/1707.06848.pdf]

[36] W. P. Thurston, *The geometry and topology of three-manifolds*, Princeton University Mathematics Department (1979), lecture notes.

[37] W. P. Thurston, edited by Silvio Levy, *Three-dimensional geometry and topology*, Princeton Univ. Press (1997).

[38] W. P. Thurston, On Proof and Progress in Mathematics, Bulletin of the American Math. Soc., Vol. 30, Number 2, April 1994, pp. 161–177.

[39] E. B. Vinberg (Ed.), *Geometry II* , Encyclopaedia of Mathematical Sciences, Vol. 29, Springer-Verlag.

[40] O. Viro, *Defining relations for reflections.* I, arXiv:1405.1460v1

[41] O. Viro, *Biflippers and head to tail composition rules* arXiv:1412.2397v2.

[42] C. P. Walkden, *Hyperbolic Geometry*. Lecture notes, The University of Manchester (2019).

[43] J. Weeks, *Non-Euclidean billiards in VR*, Bridges 2020 Conference Proceedings.

Webpages, videos, software, artworks, etc. :

- Webpages, websites, portals:
  - Cut-the-knot portal by Alexander Bogomolny.
  - Euclid's "Elements", website by David E. Joyce.
  - Drawing a Circle with a Framing Square and 2 Nails.
  - *Circle inversion*, webpage by Malin Christersson (Illustrated introduction with proofs).
  - *Geometry*, webpage by Vladimir V. Kisil (Supporting materials to a course in Euclidean, projective and inversive geometries).
  - Hyperbolic tessellations, including some printable hyperbolic tilings, webpage by David E. Joyce.
  - Isometry Classes of hyperbolic 3-space, webpage by Roice Nelson, includes animations of elliptic, hyperbolic, loxodromic and parabolic isometries.
  - More hyperbolic things by Roice Nelson (including hyperbolic Rubik's cube, honeycombs, etc...)
  - Illustrating Mathematics by Rémi Coulon (including Thurston's eight geometries, fractals, Penrose tilings, pantograph, etc...)
  - Hyperbolic tessellations by Don Hantch.
  - Tiling page and Hyperbolic geometry page on **The Geometry Junkyard** by David Eppstein.
- Videos, animations, etc:
  - *Why slicing cone gives an ellipse* - video on Grant Sanderson's YouTube channel 3Blue1Brown.
  - Loxodromic transformation in the page by Paul Nylander.
  - Dynamics of Möbius transformations is illustrated in the 2-minute video by D. Arnold and J. Rogness.
  - Animation demonstrating Inversion in circles, by M. Christersson.
  - 1-minute video illustrating stereographic projection by Henry Segerman.
  - Playing Sports in Hyperbolic Space - Dick Canary in a Numberphile video (by Brady Haran).
  - Illuminating hyperbolic geometry, Short video (4:25 min) by Henry Segerman and Saul Schleimer on projecting the hemisphere model to Klein disc, Poincare disc and upper half-plane.
  - Video comparing spherical, Euclidean and hyperbolic geometry (I don't know the author). Also, contains a link to videogames in hyperbolic plane.
- Films:
  - Not knot, a short film produced by the Geometry Centre in U. of Minnesota, popularising Thurston's Geometrization Conjecture.
  - Dimensions. 9 films for wide audience, 13 min each, produced by: Jos Leys, Étienne Ghys, Aurélien Alvarez.

- Software:
  - Applet for creating hyperbolic drawings in Poincaré disc.
  - Applet to make hyperbolic tessellations of images, by Malin Christersson.
  - Inversion Tool, hands-on demonstration of inversion on cut-the-knot partal.
  - **??**Cayley Graph Generator, Online group visualiser by Jean-Baptiste Bellynck. ¡/li¿
- Artwork:
  - M. C. Escher, official website.
  - Tilings by (and after) Escher. Webpage on Mathematical Imagery by Jos Leys.
  - Polyhedra and Art webpage by George W. Hart.
  - Polyhedral sculptures by George W. Hart.
  - Hyperbolic Geometry Artworks by Paul Nylander.
  - Webpage on hyperbolic geometry by the Institute for Figuring (with hyperbolic soccer ball and crocheted hyperbolic planes).
  - How to create repeating hyperbolic patterns, by Douglas Dunham (based on Escher's patterns). See also here.
  - *Stereographic projection and models for hyperbolic geometry*, 3D toys to illustrate by Henry Segerman.
  - D. Taimina, "Crocheting Adventures with Hyperbolic Planes". Published by A K Peters (2009).
  - R. Nelson, H. Segerman, *Visualising Hyperbolic Honeycombs*, arXiv:1511.02851.
- Games:
  - Games on hyperbolic field on Zeno Rogue webpage.
  - Hyperbolica, Non-Euclidean adventure games.
  - H2Snake, a game of snake on a hyperbolic surface of genus 2.