Geometry III/V

Anna Felikson

Durham University, 2024–2025

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).

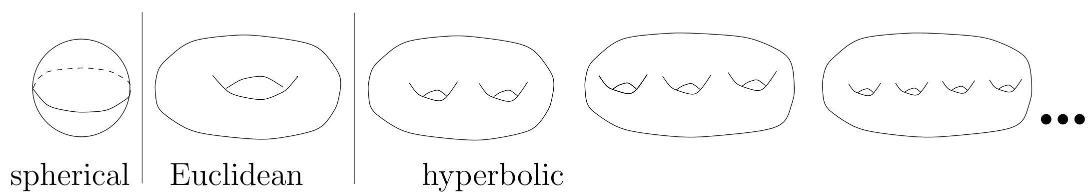


Figure 1: Geometric structures on surfaces.

- 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"

- <u>"Synthetic" way:</u>
  - 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).

- <u>"Analytic" way:</u>
  - 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 $r^2 = e$).
<u>Model 1:</u> Let r be a reflection on $\mathbb{R}^2$ (and e an identity map).
<u>Model 2:</u> $\{1, -1\} \in \mathbb{Z}$ 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 $AC = x$.

□

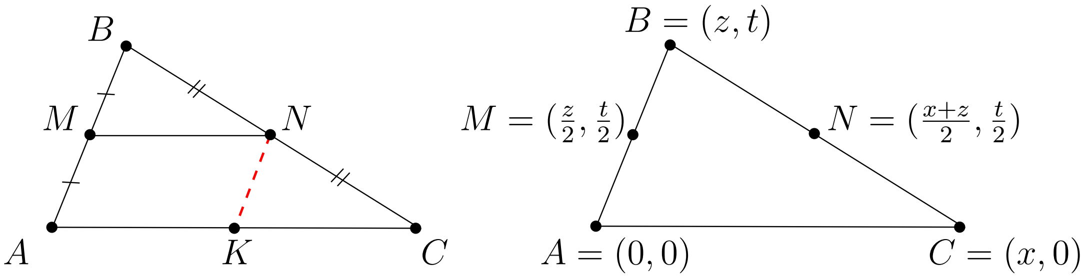


Figure 2: Two proofs of the theorem about midlines.

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 <u>without reproving</u>.
- 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 <u>many diagrams</u>:
- 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. <u>Problem solving</u> 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 ≈ 3000 BC
  (measuring land, building pyramids, astronomy).
- First records: Mesopotamia, Egypt ≈ 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 (≈ 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?

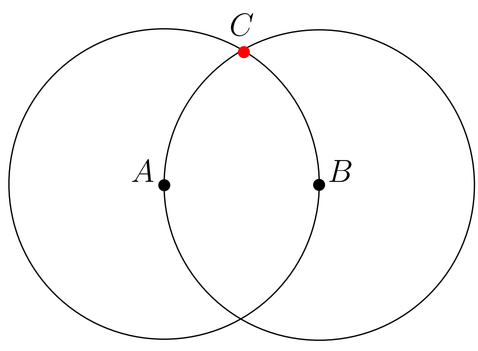


Figure 3: Euclid's proof of existence of equilateral triangle.

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.

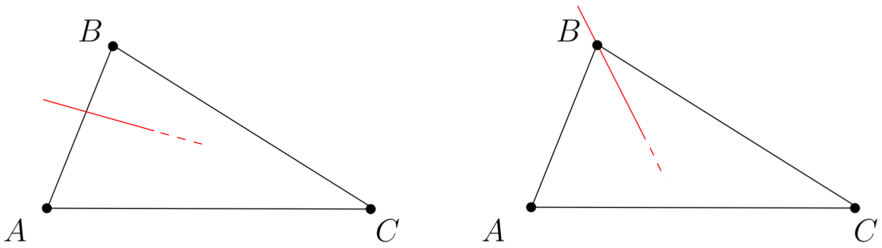


Figure 4: Pasch's Theorem and Crossbar Theorem.

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.

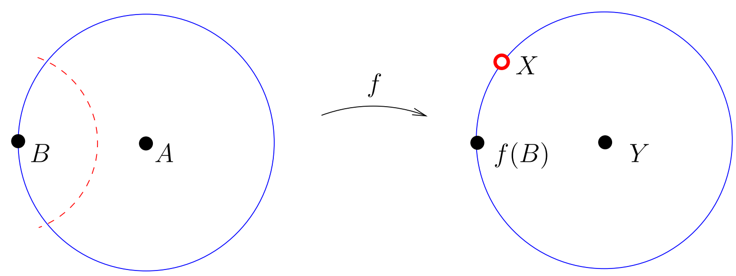


Figure 5: To the proof of surjectivity of isometry.

(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 congruent (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 a unique isometry sending A to A', B to B' and C to C'.*

Proof. Existence:

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 $D \in \mathbb{E}^2 : \quad \varphi(D) \neq D$ (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.

□

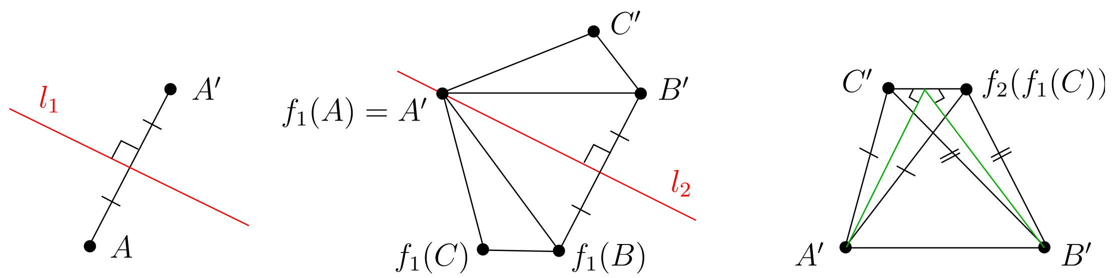


Figure 6: Isometry as a composition of reflections.

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 <u>fixed points</u> 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

□

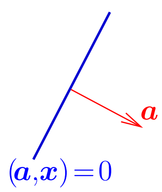
1.2 Isometries and orthogonal transformations

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

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

□

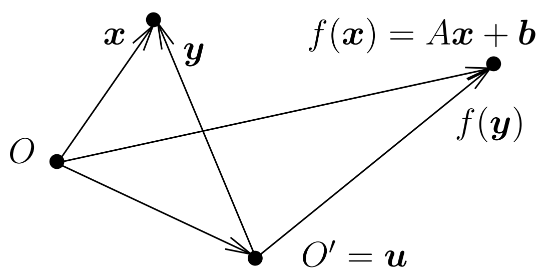


Figure 7: Linear part of isometry: independence of the origin.

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 $AA_1A_2A_3 \dots A_{n-1}A_nC \ne AC$.

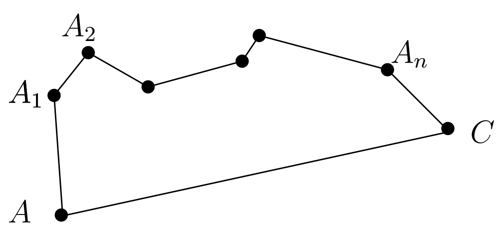


Figure 8: A broken line approximating a path.

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

$$l(\gamma\Big|_A^C) = \int_0^1 \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} dt \ge \int_0^1 \sqrt{\left(\frac{dx}{dt}\right)^2} dt$$

$$= \int_0^1 |\frac{dx}{dt}| dt \ge \int_0^1 \frac{dx}{dt} dt = x(t)\Big|_0^1 = x(1) - x(0) = b - 0 = d(A, B).$$

□

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 $w_1, \dots, w_n$ 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 <u>acts</u> 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 <u>transitive</u> 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).

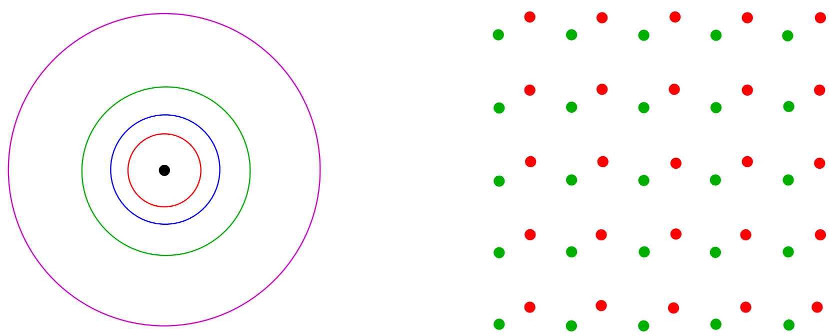


Figure 9: Orbits of O_2 (left) and $\mathbb{Z}^2 = \mathbb{Z} \times \mathbb{Z}$ (right)

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.

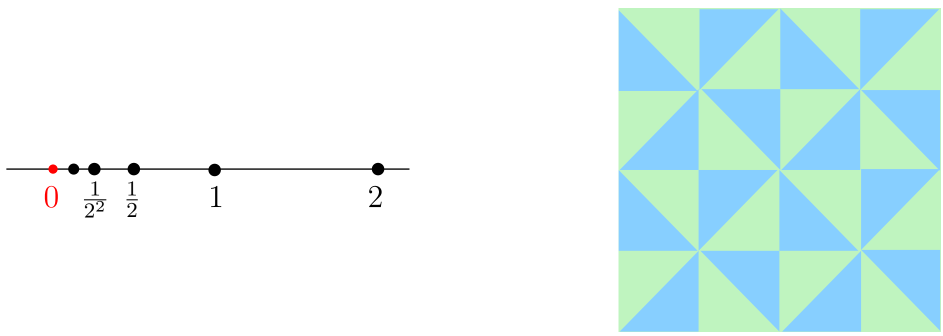


Figure 10: A non-discrete action (left) and a discrete action (right).

Definition 1.33. An open connected set $F \subset X$ is a <u>fundamental domain</u> 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 <u>orbit space</u> 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.

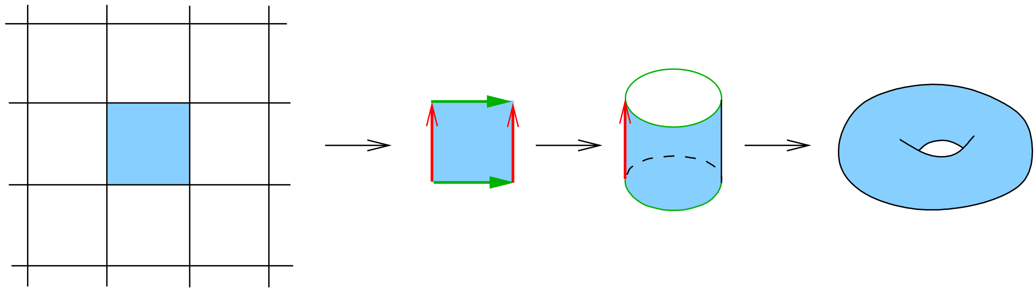


Figure 11: Fundamental domain for $\mathbb{Z}^2 : \mathbb{E}^2$ and a torus as an orbit space.

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 $d(A, B) := \inf_{a\in A, b\in B}(d(a, b))$.

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

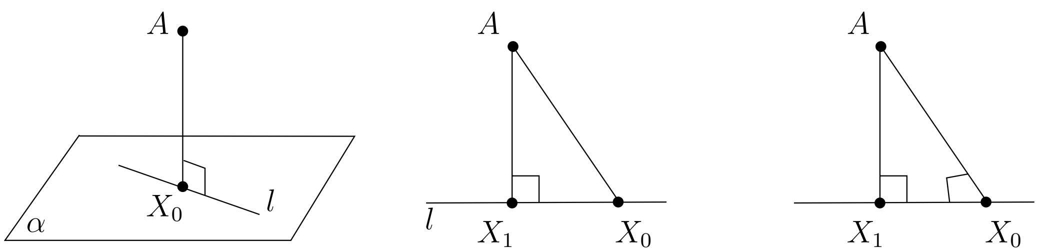


Figure 12: Distance between a point and a plane (see Proposition 1.38).

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.

□

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

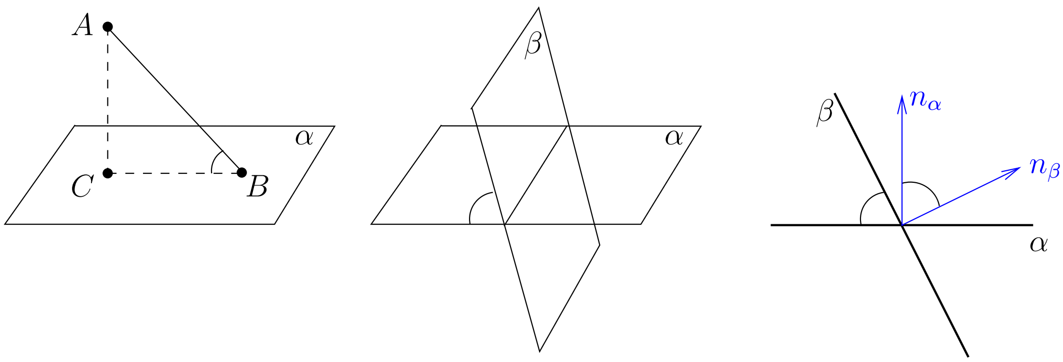


Figure 13: Angle between a line and a plane (left) and between two planes (right).

Definition 1.39. (a) The point $X_0 \in \alpha$ s.t. $d(A, \alpha) = AX_0$ is called an orthogonal projection 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 angle between the line AB and the plane α is $\angle(AB, \alpha) = \angle ABC$, where $C = proj_\alpha(A)$, (see Fig. 13, left).
Equivalently, $\angle(AB, \alpha) = \min_{X \in \alpha}(\angle ABX)$.

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 angle $\angle(\alpha, \beta)$ between two intersecting planes α and β is the angle between their normals (see Fig. 13 middle and right).
Equivalently, 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 $B \in \beta$.

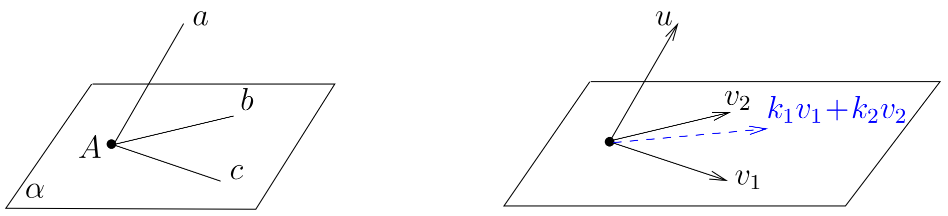


Figure 14: To Proposition 1.41.

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 $k_1, k_2 \in \mathbb{R}$.

□

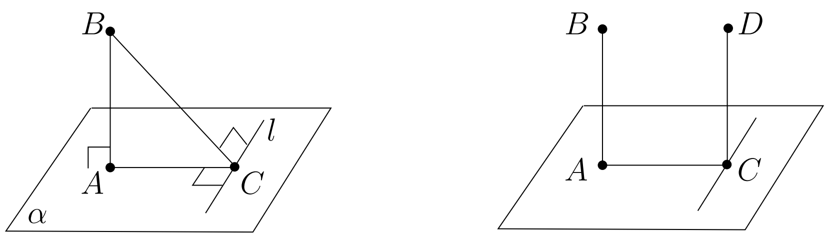


Figure 15: Theorem of three perpendiculars (and it's proof).

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).

□

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 $O = (0, 0, 0)$)

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

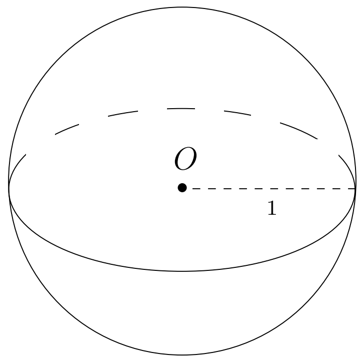


Figure 16: Sphere.

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).

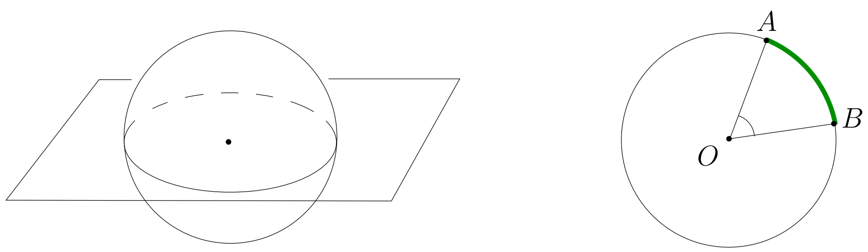


Figure 17: Great circles and distance on the sphere.

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

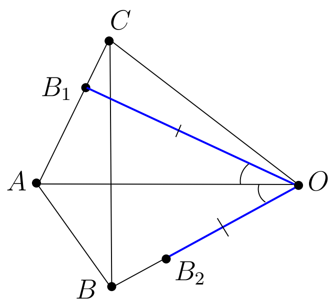


Figure 18: To the proof of triangle inequality for S^2.

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).

□

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. □

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

□

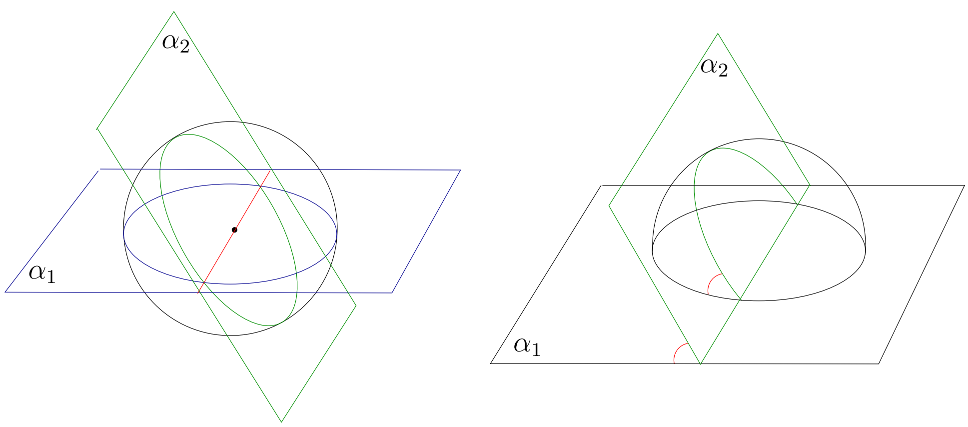


Figure 19: Intersection and angle between two lines on the sphere.

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.

□

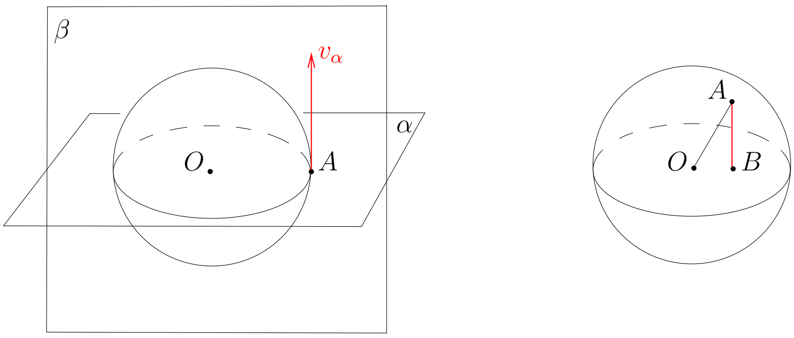


Figure 20: Existence and uniqueness of a perpendicular line on the sphere.

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.

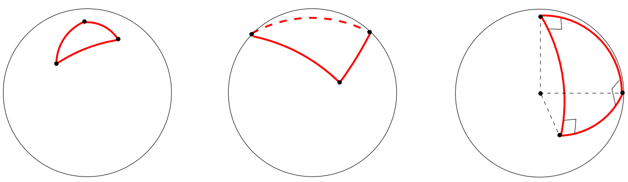


Figure 21: Spherical triangles

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. $Pol(D) = Pol(D') = l$.

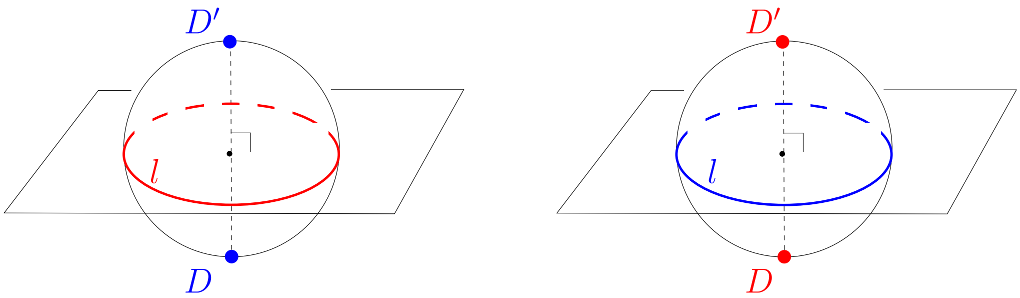


Figure 22: Polarity: $Pol(l) = \{D, D'\}$ (left) and $Pol(D) = Pol(D') = l$ (right).

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, $D' \subset Pol(A)$.

□

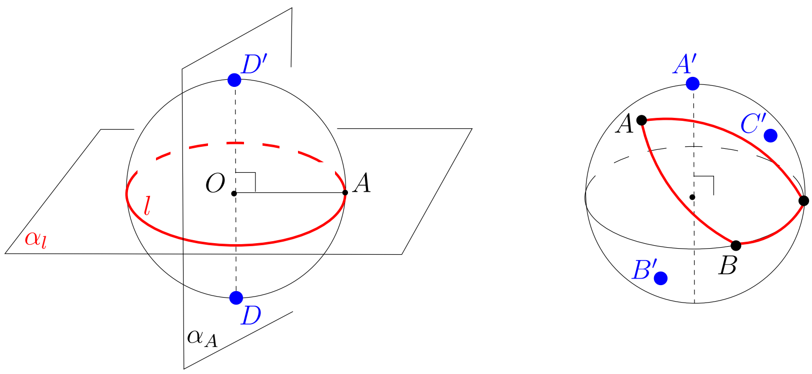


Figure 23: Left: $A \in l \Rightarrow Pol(l) \in Pol(A)$. Right: polar triangle.

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 $\angle AOA' \leq \pi/2$, and similar conditions hold for B' and C', see Fig. 23, right.

Remark. If $A' \in Pol(BC)$, then to say "$\angle AOA' \leq \pi/2$" 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.

□

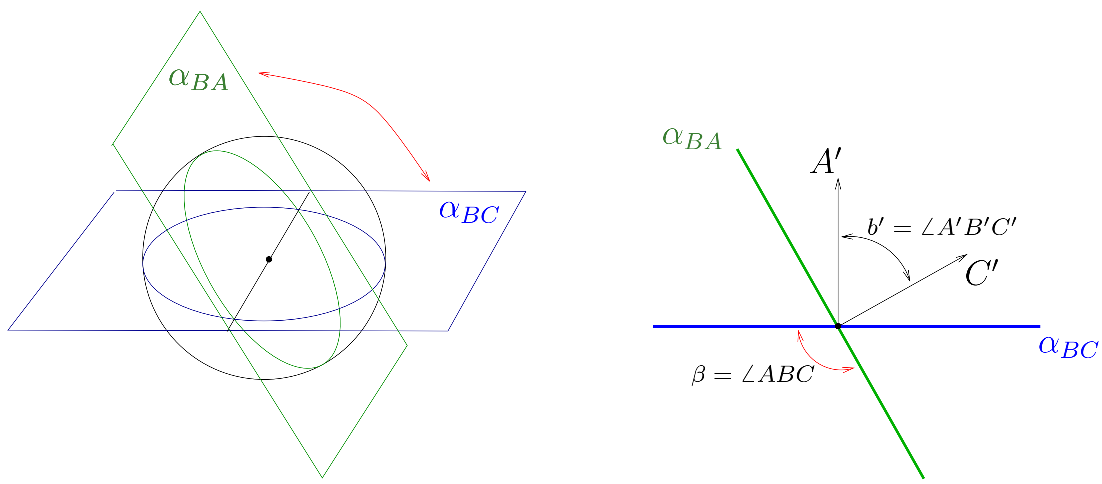


Figure 24: Proof of Bipolar Theorem.

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.

□

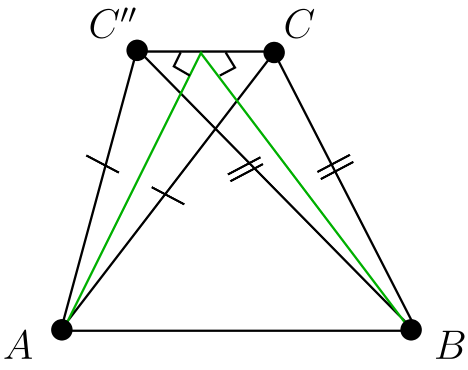


Figure 25: Proof of SSS.

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:

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

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

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

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

□

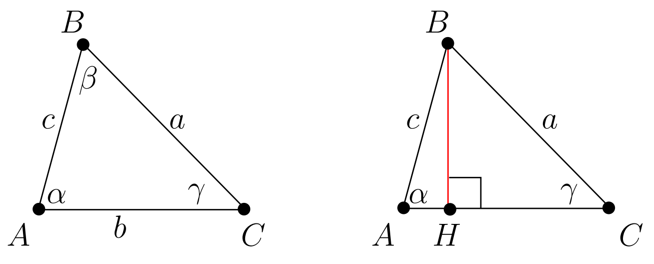


Figure 26: Triangle $\triangle ABC$.

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 $\frac{\sin b}{\sin \beta} = \frac{\sin c}{\sin \gamma}$.

□

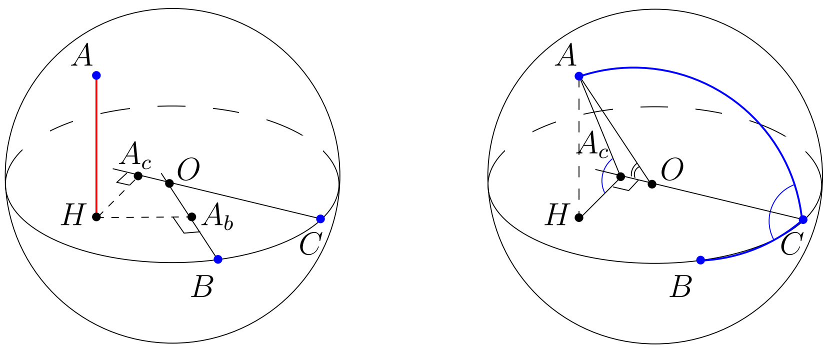


Figure 27: Proof of the sine rule on the sphere.

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 $S(2\pi) = S(sphere) = 4\pi R^2$, 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 $S_{ABC} = R^2(\alpha + \beta + \gamma - \pi)$.

□

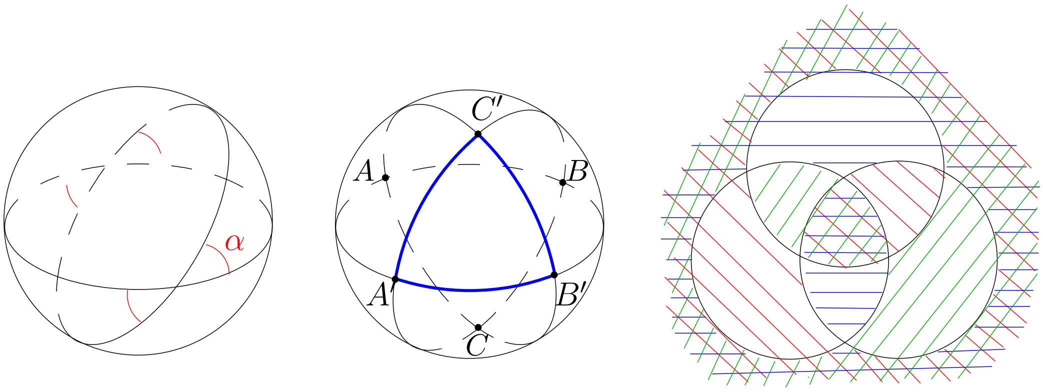


Figure 28: Computing the area of a triangle using digons

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). □

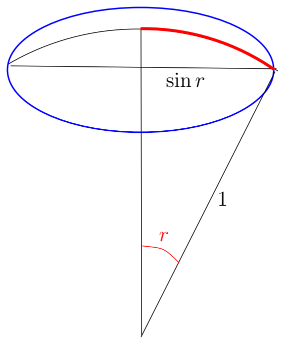


Figure 29: Computing length of spherical circle

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.

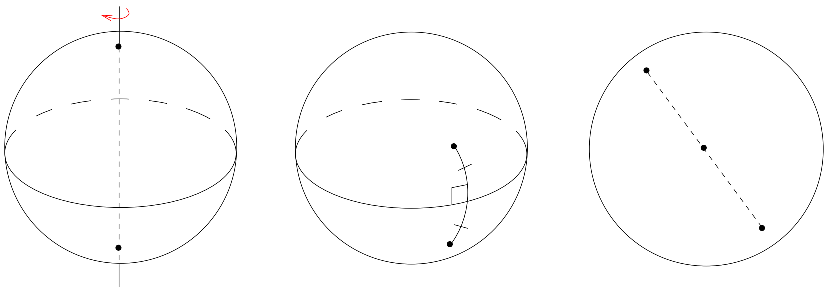


Figure 30: Examples of isometries on $\mathbb{S}^2$: rotation, reflection and antipodal map.

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?

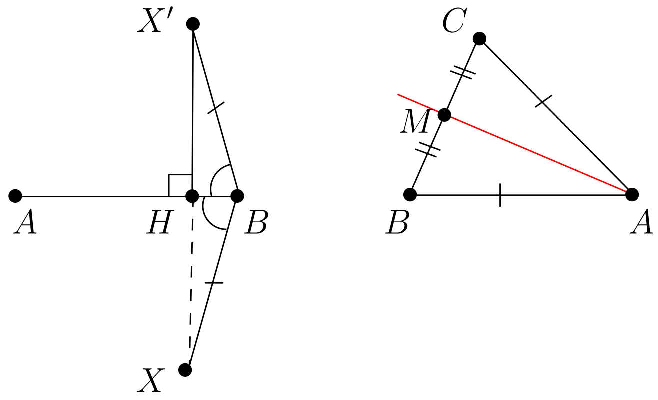


Figure 31: To the proofs of Propositions 2.30 and 2.31

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.

□

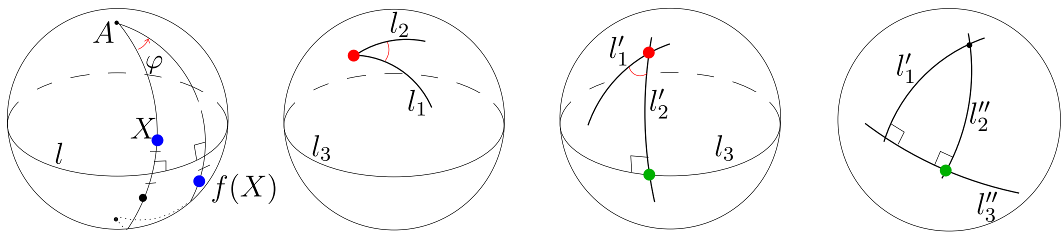


Figure 32: Glide reflection and a composition of three reflections

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 $R_{A,\varphi}^{-1} = r_l^{-1} \circ R_{B,\varphi} \circ r_l$. Also, $R_{A,\varphi}^{-1}$ is conjugate to $R_{A,\varphi}^{-1}$ 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 $R_{A,\varphi}^{-1} = r_1 \circ r_2 = r_1^{-1} \circ (r_2 \circ r_1) \circ r_1$.

□

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.

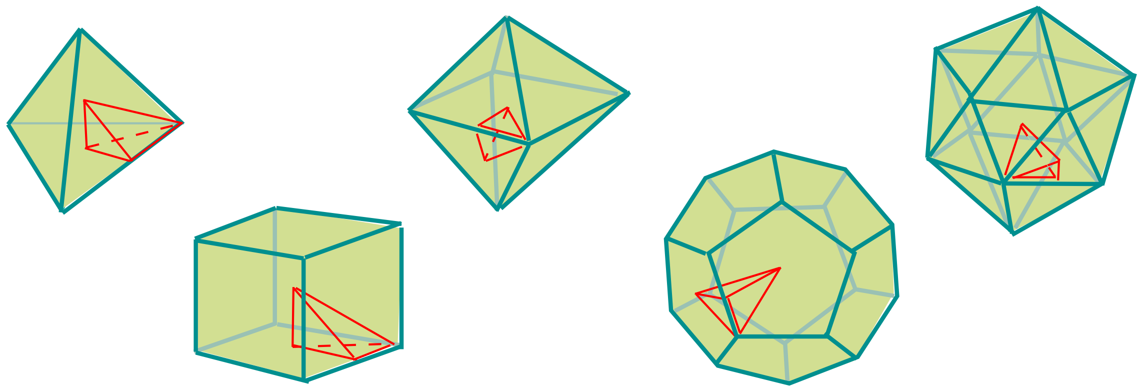


Figure 33: Regular polyhedra (from left to right): tetrahedron, cube, octahedron, dodecahedron and icosohedron.

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 dual 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 $|AC| = 2|MN|$.

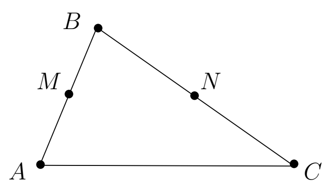


Figure 34: Length of midline using similarity

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 $f \notin Sim(\mathbb{R}^2$.

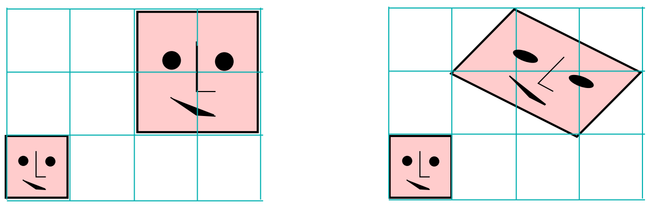


Figure 35: Examples of affine maps (see Example 3.3).

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 want 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).

□

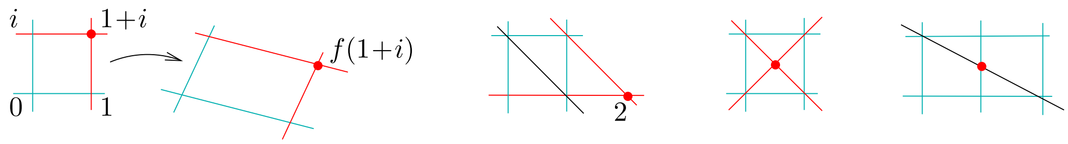


Figure 36: To the proof of Theorem 3.7.

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. □

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_∞.

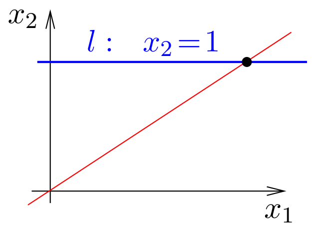


Figure 37: Projective line: set of lines through O in $\mathbb{R}^2$.

- **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

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

□

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.

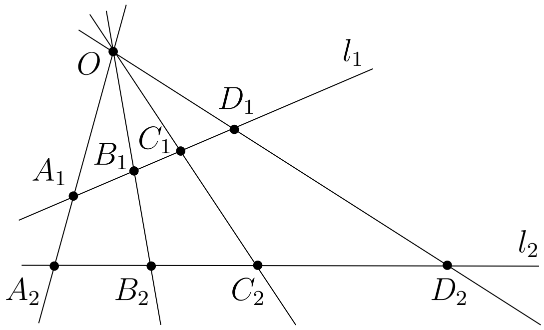


Figure 38: Projection preserves cross-ratio.

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 $A'B'C'$.

□

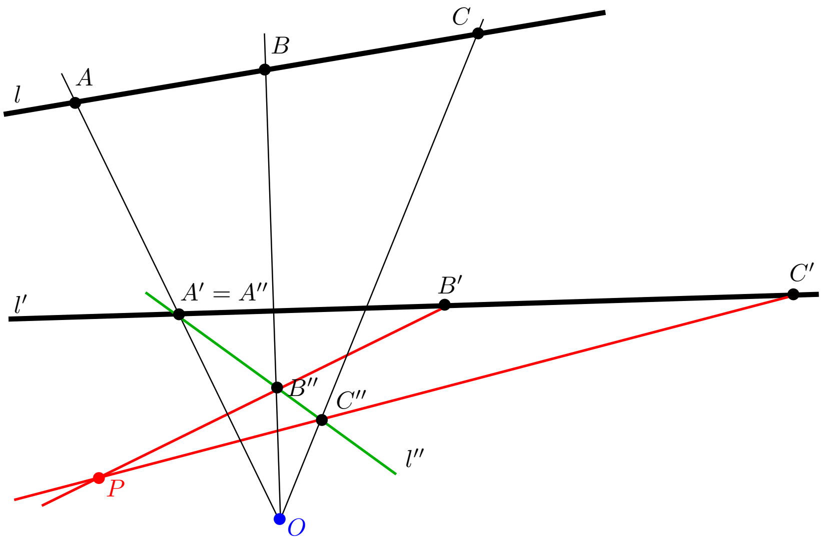


Figure 39: To the proof of Lemma 4.6.

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, $\mathbb{RP}^2$

Model:

- **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 "point at infinity".

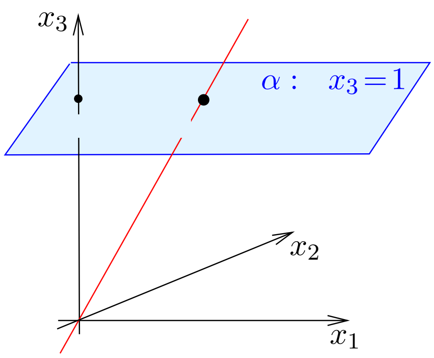


Figure 40: Projective plane: set of lines through O in $\mathbb{R}^3$.

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

    $$c_1 x_1 + c_2 x_2 + c_3 x_3 = 0, \tag{4.2}$$

    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

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

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

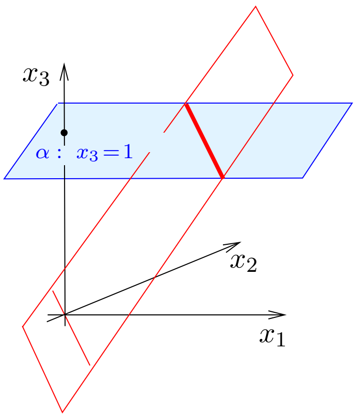


Figure 41: Projective plane: lines are planes through O in $\mathbb{R}^3$.

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.

$\square$

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.

$\square$

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.

$\square$

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 $P_2 = P_2'$.

□

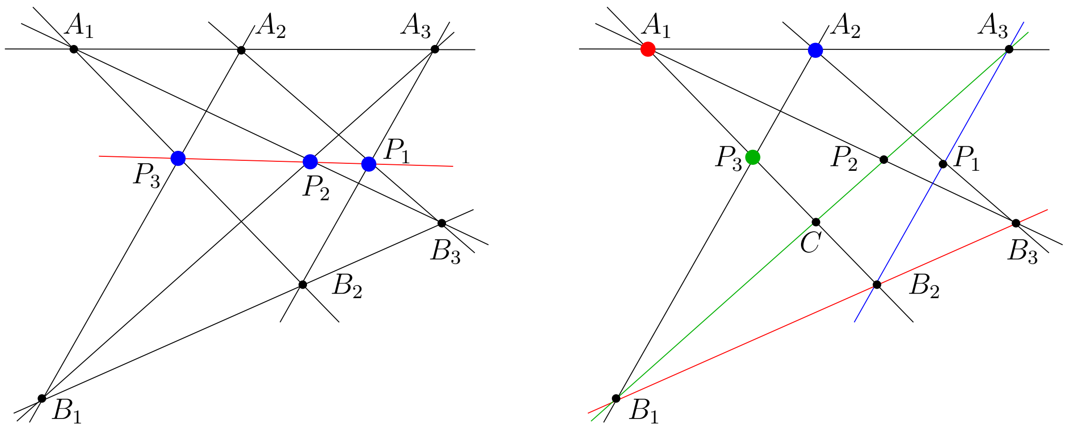


Figure 42: Pappus' Theorem and its proof by composition of 3 projections.

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.

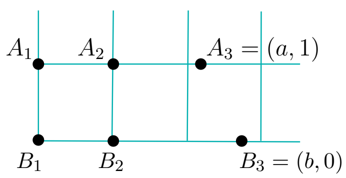


Figure 43: Another proof of Pappus' Theorem.

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!)

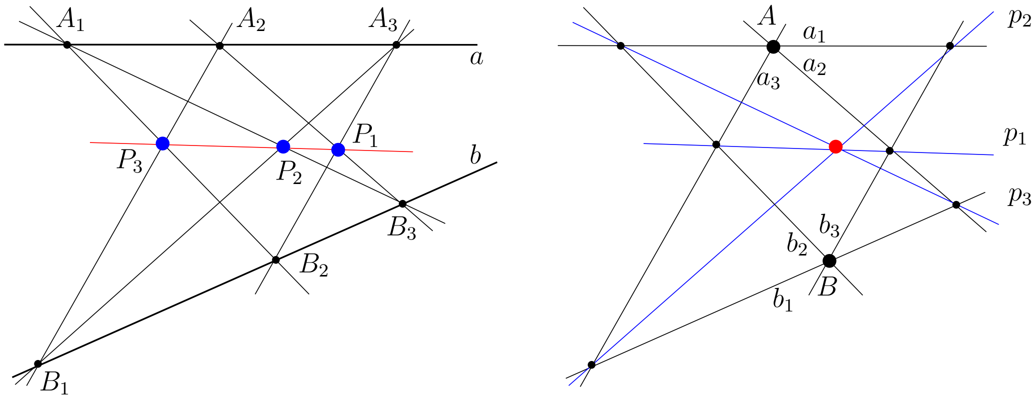


Figure 44: Pappus' Theorem and the dual statement.

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.

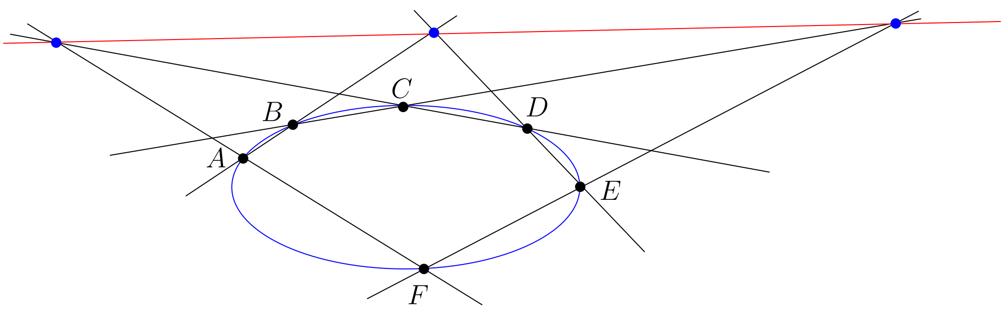


Figure 45: Pascal's Theorem.

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.

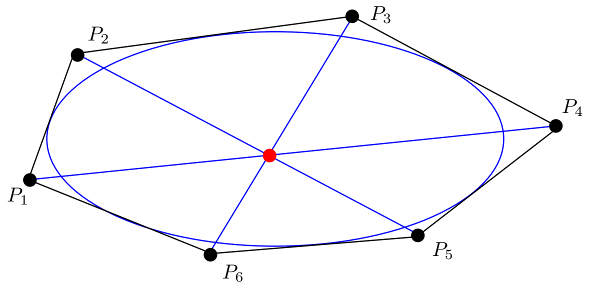


Figure 46: Brianchon's Theorem.

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.

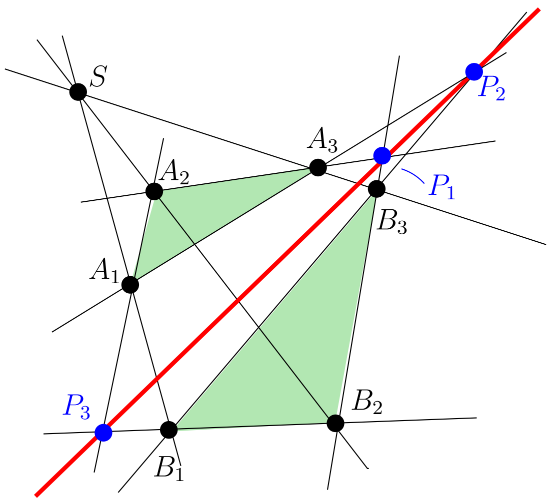


Figure 47: Desargues' Theorem.

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 $P_1, P_2, P_3 \in \gamma$.

□

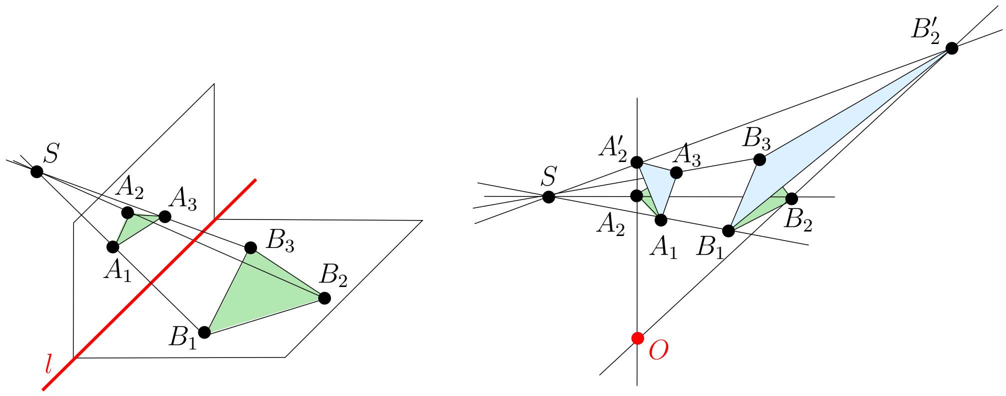


Figure 48: Proof of Desargues' Theorem.

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.

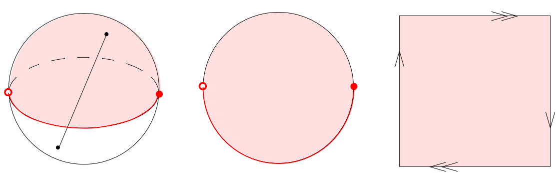


Figure 49: Topology of $\mathbb{RP}^2$.

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).

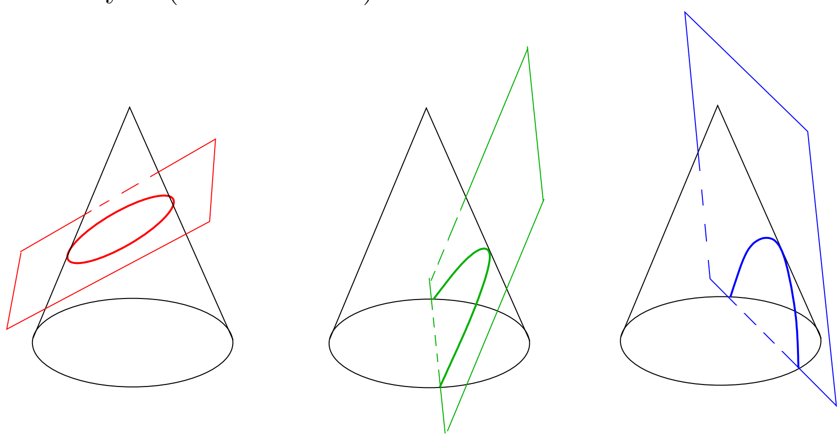


Figure 50: Conic sections: ellipse, parabola and hyperbola.

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 $a_1b_1 + a_2b_2 = a_3b_3$.

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 $a_1x_1 + a_2x_2 - a_3x_3 = 0$, 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 $a_1x_1 + a_2x_2 = 0$ 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* $x_1b_1 + x_2b_2 = x_3b_3$.

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 $a_1p_1 + a_2p_2 = a_3p_3$, so P is polar to A.
As $A \in l_Q$, we have $a_1q_1 + a_2q_2 = a_3q_3$, so Q is polar to A.
Therefore, PQ is the line polar to A, see Fig. 51, left.

□

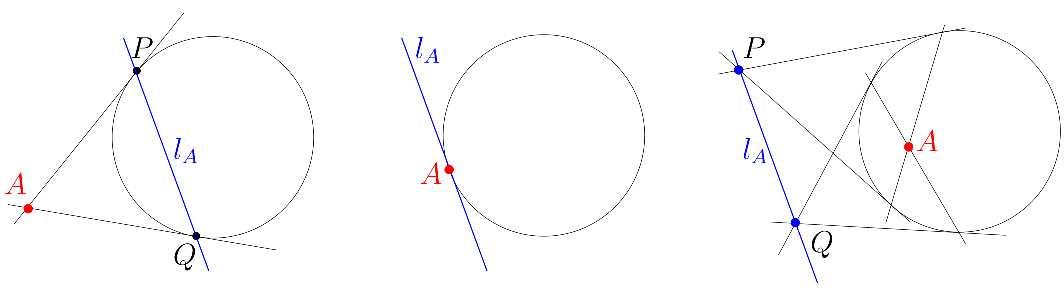


Figure 51: Polar line l_A for a point A inside, on and outside of the conic.

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.

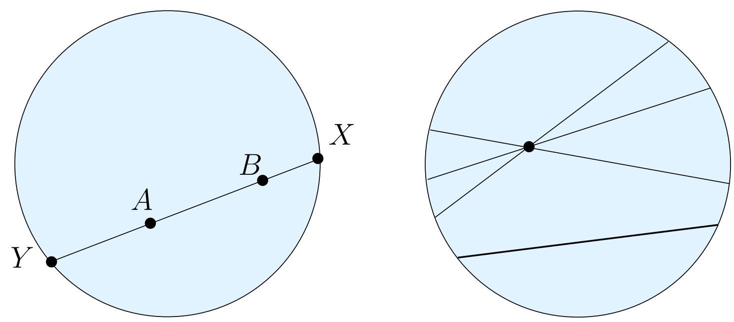


Figure 52: Klein model.

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,

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

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

$$d(a,b) = \frac{1}{2}\left|\ln[a,b,x,y]\right| = \frac{1}{2}\left|ln(\frac{x-a}{x-b}/\frac{y-a}{y-b})\right|.$$

  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).

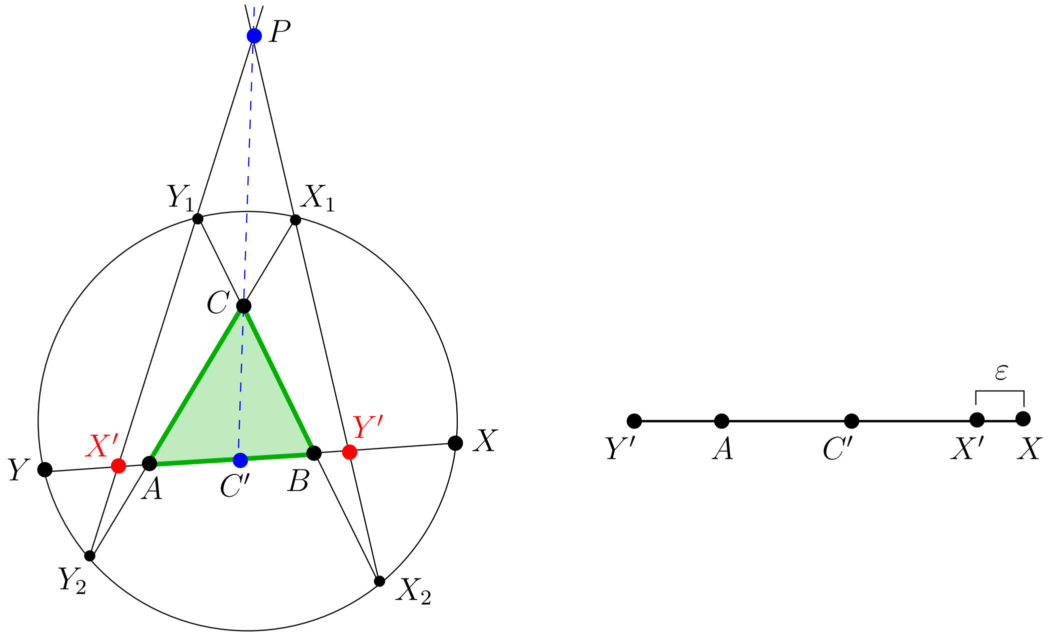


Figure 53: Proof of triangle inequality.

- 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) Claim: $[A, C', X', Y'] > [A, C', X, Y]$ 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

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

which proves the claim.

(4) Finally, we compute:

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

□

Isometries of Klein model

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

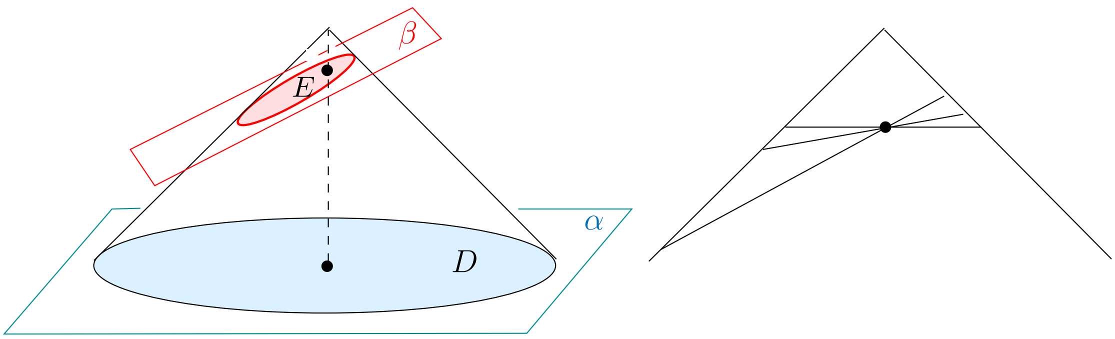


Figure 54: To the proof of Theorem 4.31.

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.

□

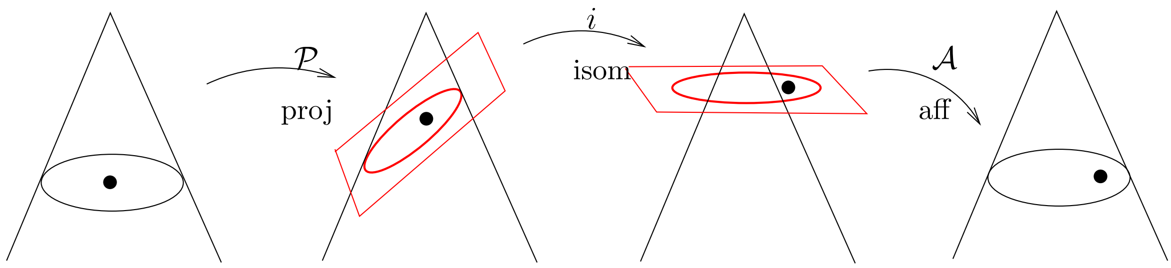


Figure 55: To the proof of Theorem 4.31.

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.

□

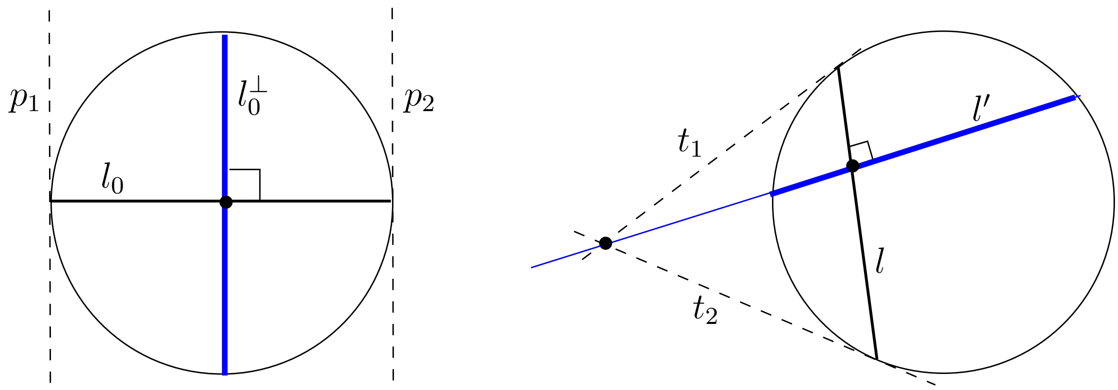


Figure 56: Right angles in the Klein model.

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.

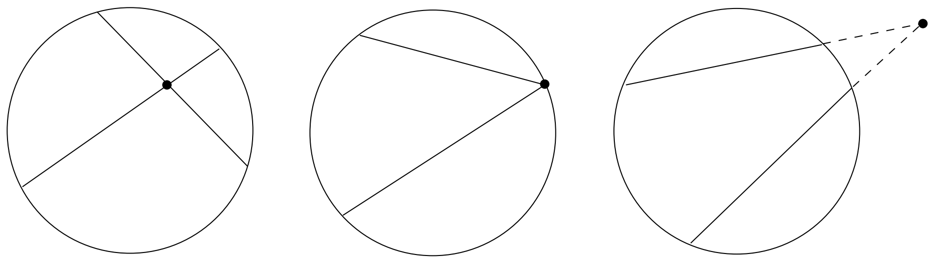


Figure 57: Pairs of lines in the Klein model: intersecting, parallel and ultra-parallel.

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

Proof. See Fig. 58.

□

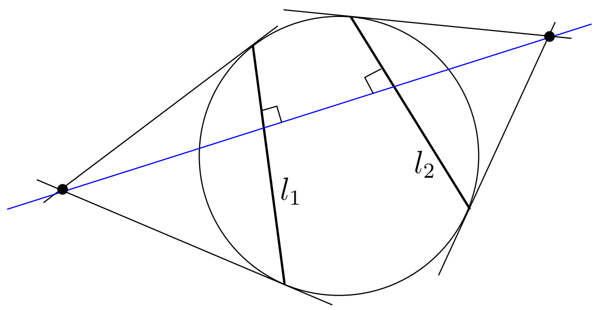


Figure 58: Common perpendicular for any ultra-parallel lines l_1 and l_2.

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.

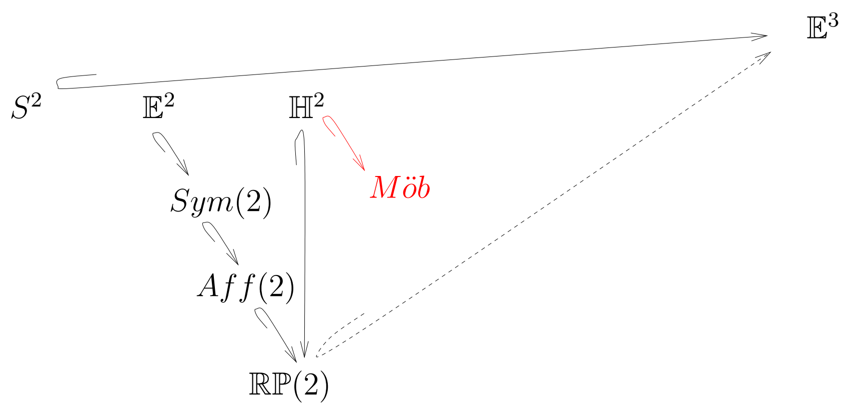


Figure 59: Hierarchy of geometries

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 $g_0 : (0, 1, \infty) \to (z_1', z_2', z_3')$.

Construction: 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).

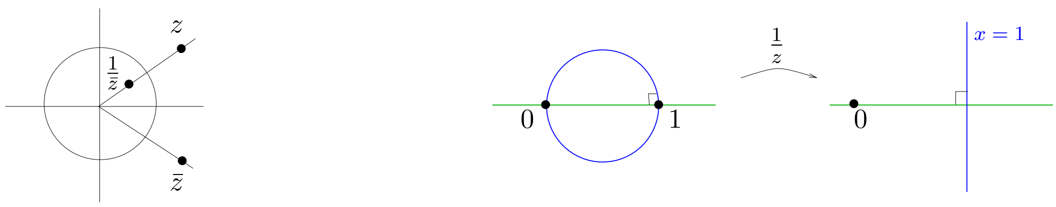


Figure 60: Left: Transformation $z \to 1/\bar{z}$. Right: an angle preserved by $z \to 1/z$.

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 Möb 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 Möb 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).

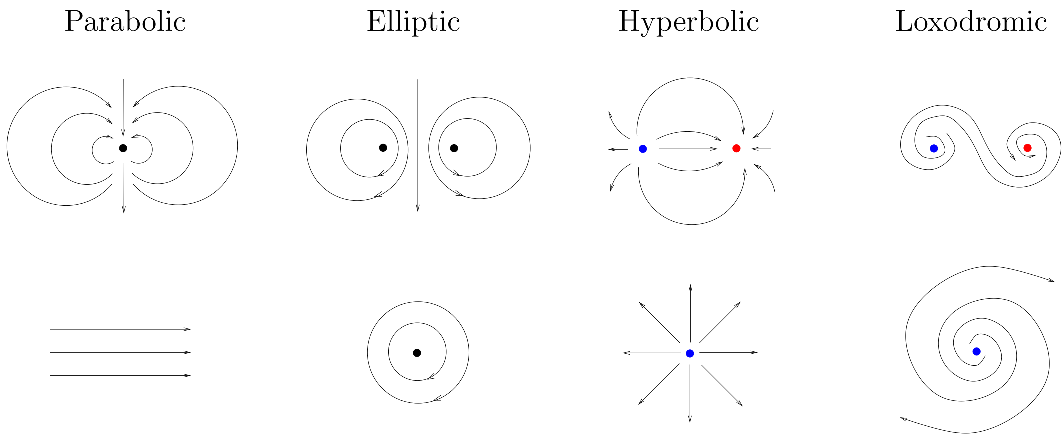


Figure 61: Dynamics of parabolic, elliptic, hyperbolic and loxodromic elements.

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* $\triangle OQ'P'$.

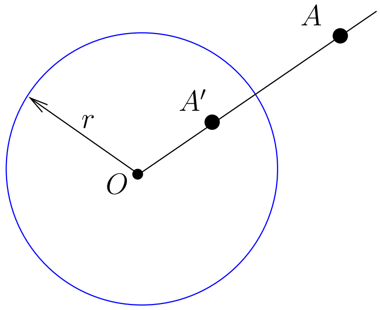


Figure 62: Inversion: $|OA| \cdot |OA'| = r^2$.

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). $\square$

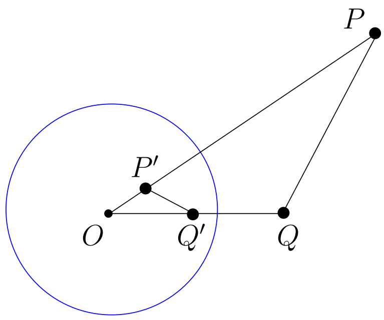


Figure 63: Inversion: $\triangle OPQ \sim \triangle OQ'P'$.

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'.

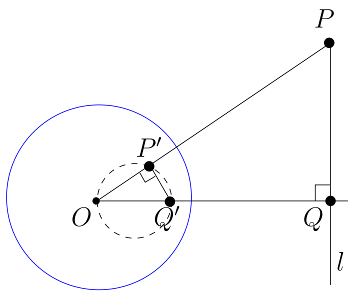


Figure 64: Inversion takes lines not through origin to the circles through origin.

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 $Q'R'$.

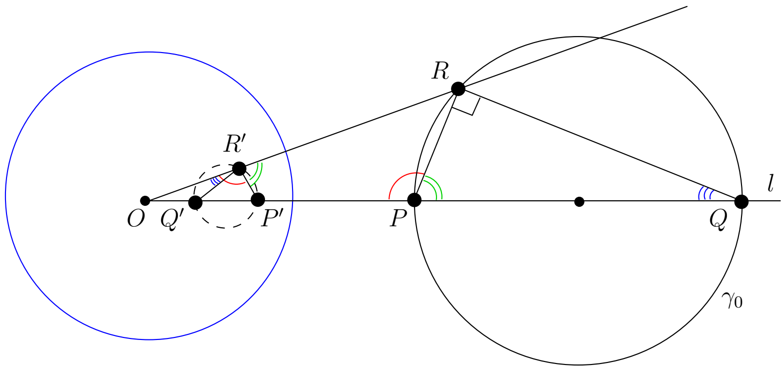


Figure 65: Inversion takes circles not through origin to the circles not through origin.

□

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.

□

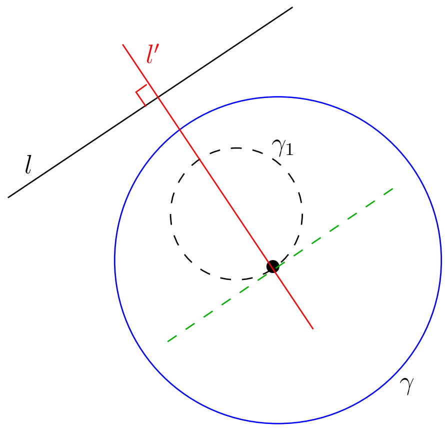


Figure 66: Inversion preserves angles.

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.

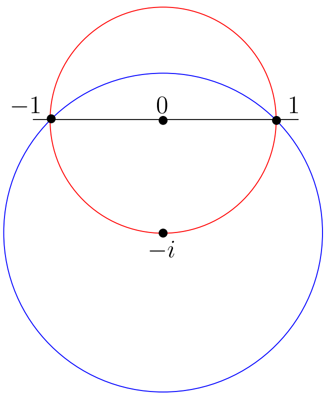


Figure 67: Reflection is a conjugated inversion.

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.

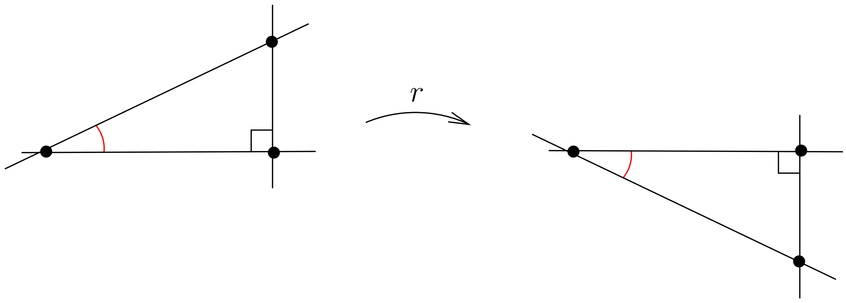


Figure 68: r preserves all real points and preserves angles, hence r is a reflection.

- $az = |a|e^{i \operatorname{Arg} a} z$ 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 reflection 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 $f \in Möb$, $f : a, b, c \to a', b', c'$ 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.

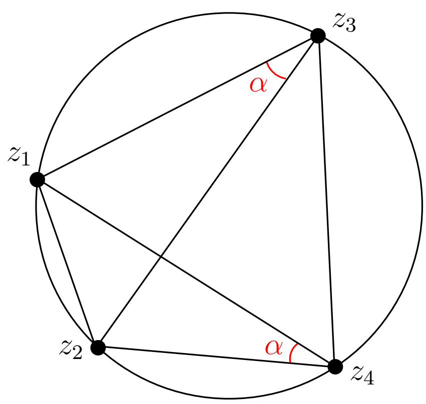


Figure 69: Geometric meaning of real cross-ratio.

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