Contents

Chapter 1. Introduction
  1.1. Euclid, The Elements (300 BC) 5
  1.2. Symmetries 7
  1.3. The algebraization of geometry 8
  1.4. Groups and symmetries 8
  1.5. Projective geometry 9
  1.6. The parallel postulate 9
  1.7. Exercises 9

Chapter 2. Isometries
  2.1. Isometries 11
  2.2. Geometric classification of isometries of the plane 12
  2.3. Algebraic classification of isometries of the plane 13
  2.4. Exercises 14

Chapter 3. Basics of group theory
  3.1. Groups and subgroups 15
  3.2. Homomorphisms 17
  3.3. Quotient groups 18
  3.4. Exercises 20

Chapter 4. Groups of isometries
  4.1. Dihedral groups 23
  4.2. Finite subgroups of $SO(2)$ 24
  4.3. Finite subgroups of $SO(3)$ 25
  4.4. Group actions 25
  4.5. Proof of the classification theorem 27
  4.6. Exercises 30

Chapter 5. Projective geometry
  5.1. Perspective 33
  5.2. Projective space 36
  5.3. Spherical model of projective geometry 37
  5.4. Hyperplane model of projective geometry 38
  5.5. Homogeneous coordinates 39
  5.6. Projective transformations 40
  5.7. The projective line 41
  5.8. Cross-ratio and applications 43
  5.9. Main theorem of projective geometry 45
  5.10. Exercises 46
Chapter 6. Hyperbolic geometry 49
  6.1. Circle inversions 49
  6.2. Elementary hyperbolic geometry 50
  6.3. Cross-ratio and Möbius transformations of $\mathbb{CP}^1$ 52
  6.4. Hyperbolic distance 53
  6.5. Angles 55
  6.6. The upper half-plane model 56
  6.7. Hyperbolic isometries 57
  6.8. Exercises 59
CHAPTER 1

Introduction

There are written traces of geometric computations going back to the Babylonians, and to the Egyptians. But the way we think about geometry—and, more generally, mathematics—is directly inherited from the Greeks. It is then that mathematics became a systematic process; from an observation, based on a specific example, one derives a claim (or conjecture), to then offer a rigorous proof. Said proof, to be valid, can only rely on a pre-established set of definitions and rules (axioms or postulates). Let’s see a concrete example. Consider odd numbers, 1, 3, 5, 7, ... and their squares 1, 9, 25, 49, ... As we continue the list, we quickly observe that the square of an odd number is also an odd number. Hence our first proposition:

**Proposition 1.** If $n$ is odd, then $n^2$ is odd.

**Proof.** That $n$ is odd means that it can not be divided by 2. In other words, it can be written as $n = 2k + 1$ (for some $k \geq 0$). Then

$$n^2 = (2k + 1)^2 = 4k^2 + 4k + 1$$

is again of the form $2m + 1$, hence odd. □

1.1. Euclid, The Elements (300 BC)

In the 13 volumes of *The Elements*, Euclid takes inventory of all known geometry of his time, based on a list of 130 definitions (what is... a point, a line, parallel lines, circles, angles, etc) and the following five postulates, based on the use of ruler and compass:

(1) Given two points, one can trace a line passing through them.
(2) A segment can be extended in a line.
(3) Given a segment, one can draw a circle.
(4) All right angles are congruent (have the same shape).
(5) Given a line and a point, there exists exactly one line parallel to the first that passes through the point.

This last postulate is called the Parallel Postulate, and is of particular importance, as we will see later on. The rest of the 13 volumes of *The Elements* consists of propositions and their proofs. Here is the first proposition in the first volume:

**Proposition 2** (Euclid, I.1). Given a line $AB$, one can construct an equilateral triangle with base $AB$.

**Proof.**
Most proofs are 'constructive,' using only ruler and compass – unsurprising, since this is essentially as much as is allowed by the five postulates!

**Proposition 3** (Euclid, I.2). *Given AB and a point C, one can find a point D such that the segments AB and CD are isometric (have the same length).*

**Proof.**

(Using the first, second and third postulates.)

Two segments (or angles) are isometric if, when you superimpose them, they coincide. We are more accustomed to think of length as a number but there was no such notion for the Greeks, who only thought in terms of integers and fractions. The construction (and acceptance) of the real numbers came much later.

**Proposition 4** (Euclid, I.15). *If two lines intersect, opposite angles coincide.*

**Proof.**

(Using the fourth postulate.)

**Proposition 5** (Euclid, I.32). *The sum on the internal angles of a triangle is equal to the sum of two right angles.*

**Proof.**
Proposition 6 (Pythagora’s theorem, Euclid I.47). In right-angled triangles, the square on the side opposite the right angle equals the sum of the squares on the sides containing the right angle, i.e. $a^2 + b^2 = c^2$.

Proof. Using four copies of the same right-angled triangle, we can construct a square:

![Diagram of a square constructed from four right-angled triangles]

We can compute the area of this square in two different ways; $(a + b)^2$ or $4ab + c^2$, and

$$(a + b)^2 = 2ab + c^2 \iff a^2 + b^2 = c^2.$$ 

Pythagora’s theorem led to the discovery of irrational numbers; if $a = b = 1$, then $c^2 = 2$.

Proposition 7. $\sqrt{2}$ can not be written as the ratio of two integers; it is “irrational.”

Proof. Exercise!

1.2. Symmetries

Euclidean geometry can be summarized in a nutshell as the study of geometric properties that are preserved by symmetries. Among the many geometric questions that fascinated the Greeks, there is the existence of perfectly symmetric planar figures: the regular polygons.

Definition 8. A polygon\(^1\) is regular if all sides have the same length, and all angles are the same.

It is obvious how to construct a regular $n$-gon. Fix a length, draw a segment of that length, then turn by an angle of $\frac{2\pi}{n}$, then go the length, turn again, etc. But can these symmetric figures be constructed only by ruler and compass? We have already seen how to construct an equilateral triangle (Euclid I.1). The square is also pretty obvious. The task really becomes non-trivial once we get to the regular pentagon. Euclid’s proof of the constructibility of the regular pentagon is often hailed as one of his most beautiful proof. Here is the idea. Euclid relied on the fact that

Proposition 9. One can construct, with ruler and compass, an isosceles triangle with angles $\alpha$, $2\alpha$ and $2\alpha$.

Proof. See the exercises below.

The angle $\alpha$ can be specified, since we know that $5\alpha = \pi$. By glueing two of these triangles, one can construct a new triangle, call it $T$, with an angle $\theta = 2\alpha = \frac{2\pi}{5}$. This proves the constructibility of the regular pentagon; five congruent copies of $T$ can be arranged to form a regular pentagon.

\(^1\) Literally, which has several (poly) sides (gons).
Now, going further: the regular hexagon can be constructed (and this is easier to prove than for the pentagon), but the regular heptagon... can not. Yet to prove that this is impossible, one had to wait the 19th century, and the development of algebraic tools in geometry.

1.3. The algebraization of geometry

The construction of the regular heptagon – or the proof that it is not constructible with ruler and compass – eluded Euclid and his contemporaries, as did the following other three famous problems of antiquity:

- Squaring the circle: given a circle of area $\pi$, can one construct a square of area $\pi$?
- Doubling the cube: given a cube of volume 1, can one construct a cube of volume 2?
- Trisecting the angle: given an angle $\alpha$, can it be divided in three equal parts?

The change of paradigm came with Descartes (1596–1650) and the introduction of the Cartesian coordinates. Along the real line, with its point of origin at 0, points are associated to a (real) number. Descartes generalized this bijection to the plane (resp. the space) and pairs (resp. triples) of real numbers. Under this bijection, points, lines, and circles are expressed by numbers and equations, and geometric constructions become a problem of solving algebraic equations.

**Theorem 1** (Descartes). A coordinate point $(x, y)$ can be constructed with ruler and compass if and only if $x$ and $y$ can be written as finite expressions containing only integers, the four arithmetic operations, and the square root.

And so, to prove that the cube can not be duplicated, one must show that the cubic root $2^{1/3}$ can not be written as an arithmetic sum of integers and square roots. To show that the angle can not be trisected, one needs to find an angle whose trisection induces a cubic trigonometric equation, with no rational solution. We now know that all of this is impossible by Galois theory. (For the quadrature of the circle, one had to wait for Lindemann to prove in the late 1800s that $\pi$ was transcendental, i.e. it is the solution of no polynomial equation with integer coefficients.)

1.4. Groups and symmetries

Let $\triangle$ denote the equilateral triangle. It presents a number of symmetries; it doesn’t change if we rotate the plane by an angle of $\frac{2\pi n}{3}$, or if we reflect across a line bisecting any one of the three angles of $\triangle$.

Symmetry can be formalized to mean an invertible function mapping $\triangle$ to a congruent (identical in form) triangle that preserves distances. Such a map is called an isometry.

**Definition 10** (Isometry). Let $E \subseteq \mathbb{R}^2$ be a subset of the plane. An isometry of $E$ is a map $f : E \to E$ such that for all $x, y \in E$, $|f(x) - f(y)| = |x - y|$.

Rotations and reflections of $\mathbb{R}^2$ and of $\triangle$ are examples of isometries. (This is a nice exercise in linear algebra.) In fact, these are the only isometries of $\triangle$.

**Proposition 11.** Let $\triangle$ denote the equilateral triangle. Let $S$ be the reflection across a line bisecting one of the angles of $\triangle$. Let $R_3$ denote the rotation of the plane by an angle of $\frac{2\pi}{3}$. Let $\text{Isom}(\triangle)$ denote the set of all isometries of $\triangle$, then

$\text{Isom}(\triangle) = \{\text{id}, R_3, R_3^2, S, S \circ R_3, S \circ R_3^2\}$,
where $\circ$ is the usual composition of functions and $R_3^2 = R_3 \circ R_3$, i.e. the rotation by an angle of $\frac{4\pi}{3}$.

**Proof.** Let $f \in \text{Isom}(\triangle)$. Since $f$ is distance-preserving, it must map corners of $\triangle$ to corners of $\triangle$. Combinatorially, there are only $3! = 6$ ways of interchanging the corners. □

The set $\text{Isom}(\triangle)$ itself exhibits some symmetries; observe that each action (reflecting, rotating) can be undone, e.g. $R_3 \circ R_3^2 = \text{id}$, $S \circ S = \text{id}$, etc. In fact, $\text{Isom}(\triangle)$ is an example of an algebraic structure modeled on these very basic symmetries; the group. The structural properties of a group – the object of next chapter – and the study of group actions will make it easier for us to understand the symmetries of more complicated objects, such as the icosahedron.

### 1.5. Projective geometry

### 1.6. The parallel postulate

### 1.7. Exercises

**Exercise 1:** Construct using ruler and compass an equilateral triangle, a square, and a regular hexagon.

**Exercise 2:** Construct using ruler and compass a golden triangle, i.e. an isosceles triangle such that the ratio of the longest side to the shortest side of a golden triangle is

$$
\varphi = \frac{1 + \sqrt{5}}{2}.
$$

A golden triangle then has the form

---

---

-- these are the triangles used by Euclid to construct a regular pentagon.

**Exercise 3:** Let $ABC$ denote the triangle with corners $A$, $B$ and $C$, and corresponding inner angles $\alpha = \angle(AC, AB)$, $\beta = \angle(AB, BC)$, $\gamma = \angle(AC, BC)$.
(1) If the side lengths $AC$ and $BC$ are equal, show, using Euclid’s postulates, that $\alpha = \beta$.
(2) Prove that the converse statement is also true.

**Exercise 4:** Pythagoras’s theorem led to the discovery of irrational numbers. Prove that $\sqrt{2}$ is irrational, i.e. that it can not be written as the ratio of two integers.

**Exercise 5:** Let $\Pi$ be the plane passing through $p_0 = (x_0, y_0, z_0)$ and whose direction is given by the normal vector $\vec{n}$. Show that reflection across $\Pi$ can be written

$$ S(\vec{x}) = \vec{x} - 2\frac{\vec{n} \cdot \vec{x}}{\|\vec{n}\|} \vec{n}. $$

Then show that this map is an isometry of $\mathbb{R}^3$.

**Exercise 6:** Show that $\text{Isom}(\triangle)$ is a group with respect to the composition of functions.

**Exercise 7:** In this exercise, we will prove the triangle inequality: in any triangle, the sum of the lengths of any two of the sides is greater than the length of the remaining side.

(1) Let $ABC$ denote the triangle with corners $A$, $B$ and $C$. We want to prove that $AB \leq AC + BC$. Explain why we may assume that $AB \geq AC, BC$.
(2) Assume the following Proposition is true:

**Proposition 1.** Let $abc$ denote the triangle with corners $a$, $b$ and $c$, and corresponding inner angles $\alpha = \angle(ac, ab)$, $\beta = \angle(ab, bc)$, $\gamma = \angle(ac, bc)$. If $\gamma > \alpha$ then $ab > bc$.

Show using this result that $AB \leq AC + BC$.
(3) Prove the above proposition.
CHAPTER 2

Isometries

2.1. Isometries

Definition 12. An isometry \( f : \mathbb{R}^n \to \mathbb{R}^n \) is a distance-preserving map, i.e. for all \( x, y \in \mathbb{R}^n \),
\[ |f(x) - f(y)| = |x - y|. \]

If \( A \subset \mathbb{R}^n \) is a subset, then \( f : \mathbb{R}^n \to \mathbb{R}^n \) is an isometry of \( A \) if it is distance-preserving and \( f(A) = A \).

We write \( \text{Isom}(\mathbb{R}^n) \) the set of all isometries of \( \mathbb{R}^n \). Observe that compositions of isometries are also isometries; indeed, if \( f, g \in \text{Isom}(\mathbb{R}^n) \), then
\[ |f(g(x)) - f(g(y))| = |g(x) - g(y)| = |x - y| \]
for all \( x, y \in \mathbb{R}^n \).

Isometries fall in one of two categories; they are either orientation-preserving, or orientation-reversing. Rotations are orientation-preserving, but reflections are orientation-reversing because they "switch" the plane. The set of all orientation-preserving isometries is denoted \( \text{Isom}^+(\mathbb{R}^n) \).

The polarization identities
\[ \langle f(x), f(y) \rangle = \frac{1}{2} \left( |f(x) + f(y)|^2 - |f(x)|^2 - |f(y)|^2 \right) \]
\[ = \frac{1}{2} \left( |f(x)|^2 + |f(y)|^2 - |f(x) - f(y)|^2 \right) \]
where \( \langle x, y \rangle \) denotes the dot-product of \( x \) and \( y \), are useful to prove that a function \( f \) is an isometry.

Proposition 13. If \( f \) is an isometry that preserves the origin, \( f(0) = 0 \), then it preserves the dot-product
\[ \langle f(x), f(y) \rangle = \langle x, y \rangle . \]

Proof. Since \( |f(x)| = |f(x) - f(0)| = |x| \),
\[ \langle f(x), f(y) \rangle = \frac{1}{2} \left( |f(x)|^2 + |f(y)|^2 - |f(x) - f(y)|^2 \right) = \frac{1}{2} \left( |x|^2 + |y|^2 - |x - y|^2 \right) = \langle x, y \rangle . \]

Proposition 14. Every isometry \( f \) can be written as the composition \( f = T \circ F \) of a translation and an isometry \( F \) that fixes the origin.

Proof. Let \( f \in \text{Isom}(\mathbb{R}^n) \), let \( o \) denote the origin. Let \( v \) be the vector from \( o \) to \( f(o) \), and let \( T_{-v}(x) = x - v \) be the translation of the plane by \(-v\). Clearly, this translation is an isometry:
\[ |T_{-v}(x) - T_{-v}(y)| = |x - v - y + v| = |x - y| . \]
Then the composition $F = T_{-v} \circ f$ is also an isometry and fixes the origin:

$$F(o) = T_{-v} \circ f(o) = o.$$ 

\[ \square \]

Isometries are exactly the functions that preserve the building blocks of Euclidean geometry: lines and angles.

**Proposition 15.** Isometries preserve lines.

**Proof.** Let $f$ be an isometry. Let $a, b, c$ be three distinct, collinear points. That is, $a, b, c$ lie on a line $L$. Suppose that $c$ lies between $a$ and $b$. Then

$$|a - b| = |a - c| + |c - b|.$$ 

Let the isometry $f$ now act on the plane: since it is an isometry, the equation above becomes

$$|f(a) - f(b)| = |f(a) - f(c)| + |f(c) - f(b)|.$$

That is, $f(a), f(b), f(c)$ are collinear. This shows that the line $L$ is mapped to another line $f(L)$. \[ \square \]

**Proposition 16.** Isometries preserve angles.

**Proof.** Suppose instead that $a, b, c$ are not collinear, and let $\alpha$ be the angle between the line $L$ through the vector $u = \vec{ac}$ and the line $L'$ through the vector $v = \vec{ab}$. Since $f$ preserves lines, it will send the angle $\alpha$ to an angle $f(\alpha)$. Suppose for the moment that $f$ is an isometry that fixes the origin. We claim that $\alpha = f(\alpha)$. Indeed, the angle $\alpha$ is determined by the cosinus identity

$$\cos(\alpha) = \frac{u \cdot v}{|u||v|} = \frac{\langle f(u), f(v) \rangle}{|f(u)||f(v)|} = \cos(f(\alpha)),$$

where the second equality follows from the fact that isometries that preserve the origin, preserve the dot-product (cf. Proposition 14). Since any isometry preserves the origin up to a translation (cf. Proposition 15), and translations clearly preserve angles, this concludes. \[ \square \]

### 2.2. Geometric classification of isometries of the plane

**Theorem 2.** Every isometry of the plane is of one of the following forms:

1. orientation-preserving isometries: identity, translation or rotation
2. orientation-reversing isometries: reflection, or glide reflection

Another way of classifying the isometries of the plane would have been through their fixed point sets. Recall that if $f$ is a function, its set of fixed points is $\{x : f(x) = x\}$. The identity map fixes all points of the plane, while translation fixes none. Rotations fix a single point; that around which the rotation takes place. Reflection fixes each point on the line across which reflection takes place, while glide reflection fixes no points, yet fixes the line (as a 1-dimensional object) across which the reflection takes place (but not the points along that line). The latter observation is why glide-reflections are often included in the classification of isometries of the plane, although they are defined by the composition of a translation and a reflection.
Lemma 19. A function \( f : \mathbb{R}^n \to \mathbb{R}^n \) is orthogonal if it preserves the dot-product; \( \langle f(x), f(y) \rangle = \langle x, y \rangle \). A function \( f : \mathbb{R}^n \to \mathbb{R}^n \) is called linear if it preserves all lines passing through the origin, i.e. if \( f(ax + y) = af(x) + f(y) \) for all \( a \in \mathbb{R} \), \( x, y \in \mathbb{R}^n \).

Proposition 17. A function \( f : \mathbb{R}^n \to \mathbb{R}^n \) is linear orthogonal if and only if it is an isometry that fixes the origin.

Proof. Let \( f \) be linear orthogonal. Since it is linear, \( f(o) = o \). Let \( x, y \in \mathbb{R}^n \). Then
\[
|f(x) - f(y)|^2 = |f(x - y)|^2 = \langle f(x - y), f(x - y) \rangle = \langle (x - y), (x - y) \rangle = |x - y|^2.
\]
Conversely, we have seen that isometries preserve lines. Hence if \( f(o) = o \), the isometry is linear. We have also remarked earlier that an isometry that fixes the origin must preserve the dot-product.

Corollary 18. Every isometry \( f : \mathbb{R}^n \to \mathbb{R}^n \) is the composition of an orthogonal linear function and a translation.

2.3. Algebraic classification of isometries of the plane

Recall that a function \( f : \mathbb{R}^n \to \mathbb{R}^n \) is orthogonal if it preserves the dot-product. Let \( p \neq o \) be another point of the plane.

If \( f(p) = p \), then \( f \) preserves the line \( L \) going through \( o \) and \( p \). Hence \( f \) is either a reflection \( S_L \) across the line \( L \) or the identity. (It can not be a glide-reflection, since points on the line are fixed.)

If \( f(p) \neq p \), let \( \theta \) be the angle between the two lines \( op \) and \( of(p) \). Then
\[
R_{-\theta} \circ f(p) = p \quad \text{and} \quad R_{-\theta} \circ f(o) = o,
\]

hence \( R_{-\theta} \circ f \) fixes the line through \( o \) and \( p \) and is again either a reflection or the identity.

Summarizing, we only have the following options: \( f = \text{id} \) (identity), \( f = T \) (translation), \( f = T \circ S_L \) (reflection or glide-reflection), \( f = T \circ R_{\theta} \) (rotation), \( f = T \circ R_{\theta} \circ S_L \) (reflection or glide-reflection).

A consequence of this classification theorem is, for instance, that the composition of two rotations about two different points is either a rotation about a third point or a translation.
A matrix of this form is a rotation: \[
\begin{pmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{pmatrix}
\begin{pmatrix}
1 \\
0
\end{pmatrix}
= \begin{pmatrix}
\cos \theta \\
\sin \theta
\end{pmatrix}.
\] If on the other hand \( A \in O(2) \) has determinant \( \det(A) = -1 \), then \( A \) is of the form
\[
A = \begin{pmatrix}
-a & b \\
b & a
\end{pmatrix} = \begin{pmatrix}
-1 & 0 \\
0 & 1
\end{pmatrix} \begin{pmatrix}
a & -b \\
b & a
\end{pmatrix}.
\]
The matrix \( \begin{pmatrix}
-1 & 0 \\
0 & 1
\end{pmatrix} \) is the reflection of the plane across the \( x \)-axis: \( \begin{pmatrix}
-1 & 0 \\
0 & 1
\end{pmatrix} \begin{pmatrix}
x \\
y
\end{pmatrix} = \begin{pmatrix}
x \\
-y
\end{pmatrix} \). Hence \( A \) is the composition of a reflection and a rotation, hence is a reflection.

Note that the determinant picks up whether the isometry is orientation-preserving or not. In fact, \( \det(A) = 1 \) indicates that \( A \) preserves the orientation, while if \( \det(A) = -1 \), the orientation is reversed.

**Proposition 21.** Isometries are bijective.

**Proof.** Let \( f \) be an isometry. If \( f \) does not fix the origin, then there exists some translation \( T \) such that the isometry \( F = T \circ f \) fixes the origin. Then \( F \) is linear orthogonal and has a matrix representative \( A \in O(n) \). Since \( \det(A) \neq 0 \), this matrix is invertible, and this implies that \( F \) is invertible. Since translation is a bijective function – the inverse function to \( T_{-v} \) is \( T_v \) –, the isometry \( f \) must be bijective. \( \square \)

**Corollary 22.** \( \text{Isom}(\mathbb{R}^n) \), together with the composition of functions, is a group.

### 2.4. Exercises

**Exercise 1:** Show that isometries preserve center of mass.

**Exercise 2:** Show that there exists no isometry in \( \text{Isom}(\mathbb{R}^2) \) with two fixed points.

**Exercise 3:** Let \( f \in \text{Isom}(\mathbb{R}^2) \).

1. Show that if \( f \) is a composition of three reflections, then \( f \) is not a rotation.
2. Show that if \( f \) is a composition of two reflections, then \( f \) is a rotation if and only if \( f \) has a fixed point.
CHAPTER 3

Basics of group theory

3.1. Groups and subgroups

**Definition 23 (Group).** A group \((G, \cdot)\) is a non-empty set \(G\) together with a map (called a law of composition)

\[
G \times G \to G, \quad (g, h) \mapsto g \cdot h
\]

such that:

1. The law of composition is associative:
   
   \[
   (g \cdot h) \cdot k = g \cdot (h \cdot k).
   \]

2. The set \(G\) contains an identity element \(e\) such that for all \(g \in G\),
   
   \[
   e \cdot g = g \cdot e = g.
   \]

3. Every element \(g \in G\) has an inverse \(g^{-1} \in G\) such that
   
   \[
   g \cdot g^{-1} = g^{-1} \cdot g = e.
   \]

A group is called

- **abelian** if its law of composition is commutative, i.e. \(gh = hg\) for all \(g, h \in G\),
- **finite** if it contains only a finite number of elements. That number is then called the order of the group.

For a law of composition to make sense on a set \(G\), it is important that its image is in \(G\), i.e. \(g \cdot h \in G\) if \(g, h \in G\). For example, addition or multiplication are laws of composition of the integers, but division is not \(-2/3 \notin \mathbb{Z}\). If \(G\) is a set of functions, then the composition of functions is a law of composition, albeit not necessarily a commutative one.

**Examples 24.**

1. \((\mathbb{R}, \cdot)\) with multiplication is not a group, but \((\mathbb{R}^*, \cdot)\) is.

   **Proof.** The set of all real numbers is obviously nonempty, and multiplication of real numbers is an associative (and even commutative) law of composition. The identity element is 1, since multiplying any real number by 1 leaves it unchanged. Every nonzero real number \(a\) has an inverse, namely \(a^{-1}\), since \(a/a = 1\). However, zero has no inverse. Hence \(\mathbb{R}\) is not a group under multiplication, but \(\mathbb{R}^*\) is. \(\square\)

2. The general linear group \(\text{GL}_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) : \det(A) \neq 0\}\) of \(n \times n\) square matrices with real entries and nonzero determinant, together with matrix multiplication, is a group.

   **Proof.** Observe that the identity matrix \(I_n\) is contained in \(\text{GL}_n(\mathbb{R})\). Matrix multiplication is a law of composition for \(\text{GL}_n(\mathbb{R})\) since for any \(A, B \in \text{GL}_n(\mathbb{R})\),
   
   \[
   \det(AB) = \det(A) \det(B) \neq 0,
   \]
3. BASICS OF GROUP THEORY

i.e. \( AB \in \text{GL}_n(\mathbb{R}) \). Since \( \det(A) \neq 0 \) is equivalent to the fact that the matrix \( A \) is invertible, each element of \( \text{GL}_n(\mathbb{R}) \) has an inverse \( AA^{-1} = I_n \).

(3) \((\text{Isom}(\Delta), \circ)\) is a group with respect to the composition of functions.

**Proof.** Exercise.

**Proposition 25** (Uniqueness of identity and inverse elements). Let \( G \) be a group. Then the identity element \( e \in G \) is unique. More generally, if \( g \in G \), then \( g^{-1} \) is unique.

**Proof.** Suppose \( e \) and \( e' \) are both identities. Then \( e = e \cdot e' = e' \). Suppose \( h \) and \( k \) are both inverse to \( g \), then \( h = e \cdot h = (k \cdot g) \cdot h = k \cdot (g \cdot h) = k \cdot e = k \).

**Proposition 26** (Cancellation). If \( g \cdot h = g \cdot k \) then \( h = k \). If \( g \cdot k = h \cdot k \) then \( g = h \).

**Proof.** Since \( g \in G \), \( g^{-1} \in G \). Then

\[
g \cdot h = g \cdot k \implies g^{-1} \cdot (g \cdot h) = g^{-1} \cdot (g \cdot k) \iff (g^{-1} \cdot g) \cdot h = (g^{-1} \cdot g) \cdot k \iff e \cdot h = e \cdot k \iff h = k.
\]

Proving cancellation on the right works in the exact same way.

An important consequence of the cancellation law is that \( g \cdot h = k \iff h = g^{-1} \cdot k \). **From this point on**, we understand the law of composition implicitly and write \( gh := g \cdot h \).

**Definition 27** (Subgroup). Let \( G \) be a group. A nonempty subset \( H \subset G \), with the same law of composition as \( G \), is a subgroup of \( G \) if:

1. the identity element \( e \) of \( G \) is also contained in \( H \),
2. \( H \) is closed under the law of composition, i.e.

\[
g, h \in H \implies gh \in H \quad \text{and} \quad g \in H \implies g^{-1} \in H
\]

You can check from this definition that \( H \) is a group. Since the identity element in a group is unique, the identity element of \( H \) is the same as for \( G \). Since each inverse is unique, the inverse \( h^{-1} \) of an element \( h \in H \) is the same as its inverse in \( G \). Any group has at least two subgroups, \( \{e\} \) and \( G \). These are called the trivial subgroups. A subgroup \( H \subset G \) is called proper if \( H \neq \{e\}, G \).

**Examples 28.**

1. For each \( n \in \mathbb{Z} \), \( n\mathbb{Z} = \{\ldots, -2n, -n, 0, n, 2n, \ldots\} \) is a subgroup of \((\mathbb{Z}, +)\).

**Proof.** For any \( n \), \( 0 \in n\mathbb{Z} \). Fix an integer \( n \), and let \( a, b \in n\mathbb{Z} \). We can write \( a = nk \), \( b = nl \) for some \( k, l \in \mathbb{Z} \). Then

\[
a + b = n(k + l) \in n\mathbb{Z}, \quad -a = -nk \in n\mathbb{Z}, \quad 0 \in n\mathbb{Z}.
\]

(2) The unit circle \( S^1 \) – i.e., the circle of radius 1 centered at the origin of the plane – is a subgroup of \( \mathbb{C}^* = \mathbb{C} \setminus \{0\} \) with respect to the multiplication of complex numbers.

**Proof.** We leave it as an exercise to show that \( \mathbb{C}^* \) with respect to the multiplication of complex numbers is a group. Any \( z \in S^1 \) can be written in polar coordinates as \( z = e^{i\alpha} \), and \( e^{i\alpha}e^{i\beta} = e^{i(\alpha+\beta)} \), \( (e^{i\alpha})^{-1} = e^{-i\alpha} \) and \( 1 \in S^1 \).
Let $G$ be a group, and let $g \in G$. The smallest subgroup of $G$ containing $g$ is

$$\langle g \rangle = \{g^k : k \in \mathbb{Z}\},$$

where $g^k = g \cdots g$ denotes the composition of $g$ with itself $k$ times. We call $\langle g \rangle$ the subgroup generated by $G$. (Check that it is a subgroup!) More generally, $\langle g_1, g_2, \ldots, g_n \rangle$ is the set of elements of $G$ that can be written as composition of powers of $g_1, g_2, \ldots, g_n$. For example, $\langle 2, 3 \rangle = \{2k + 3l : k, l \in \mathbb{Z}\}$. Observe that since $3 - 2 = 1$, $\langle 2, 3 \rangle = \{1\} = \mathbb{Z}$.

**Definition 29** (Cyclic group). A group is called cyclic if it is generated by a single element.

For example, $\mathbb{Z} = \langle 1 \rangle$, $n\mathbb{Z} = \langle n \rangle$. The group of rotations $\{e, R_n, R_n^2, \ldots, R_n^{n-1}\}$, where $R_n$ is a counterclockwise rotation by an angle of $\frac{2\pi}{n}$ that preserves a regular $n$-gon, is an example of a finite cycle group, generated by $R$. On the other hand, $\text{Isom}(\Delta)$ is not cyclic. (Why?)

### 3.2. Homomorphisms

**Definition 30** (Homomorphism). Let $G, G'$ be groups. A (group) homomorphism $f : G \to G'$ is a map that preserves the laws of composition of $G$ and $G'$, i.e. $f(gh) = f(g)f(h)$.

**Proposition 31.** Let $f : G \to G'$ be a homomorphism. Then $f$ maps the identity of $G$ to the identity of $G'$ and $f(g^{-1}) = f(g)^{-1}$.

**Proof.** Let $g \in G$. Let $e$, resp. $e'$, be the identity of $G$, resp. $G'$. Then $f(g) = f(e) = f(e')f(g)$. On the other hand, $f(g) = e'f(g)$. By cancellation, $f(e) = e'$. The statement for the inverse is proven in the same way.

**Examples 32.** Check that the following examples are homomorphisms.

1. $\det : \GL_n \mathbb{R} \to \mathbb{R}^*$
2. $\exp : \mathbb{R} \to \mathbb{R}, t \mapsto e^t$
3. $(\text{norm} | ) : \mathbb{C}^* \to \mathbb{R}^*$
4. (identity map) $G \to G$, $g \mapsto g$
5. (zero map) $G \to H$, $g \mapsto e$
6. (inclusion map) Let $H \subset G$ be a subgroup, $H \to G$, $h \mapsto h$
7. (rotation in the plane, around the origin) $\mathbb{R} \to \SL_2(\mathbb{R})$, $t \mapsto \begin{pmatrix} \cos(t) & -\sin(t) \\
\sin(t) & \cos(t) \end{pmatrix}$. 

A homomorphism $f : G \to G'$ determines two subgroups: the image of the map $\im f = \{g' \in G' : g' = f(g)\}$, and its kernel $\ker f = \{g \in G : f(g) = e\}$. We quickly check that these two sets are really subgroups of $G'$ and $G$ respectively:

$$g'_1, g'_2 \in \im f \implies g'_1 g'_2 = f(g_1)f(g_2) = f(g_1g_2) \in \im f$$

and

$$g' = f(g) \implies g'^{-1} = f(g)^{-1} = f(g^{-1}) \in \im f.$$ 

These subgroups carry some information relative to the map $f$.

**Proposition 33.** A homomorphism $f : G \to G'$ is injective if and only if its kernel is $\{e\}$. A homomorphism is surjective if and only if $\im f = G'$.

**Proof.** Suppose that $f$ is injective, i.e. $f(g) = f(g') \implies g = g'$. Then $f(g) = e = f(e) \implies g = e$. This proves the only if statement. Conversely, $f(g) = f(g') \iff e = f(g)f(g') = f(gg')$. Since $\ker f = \{e\}$, $gg'^{-1} = e$, hence $g = g'$. The second statement is immediate from the definition of a surjective map.
Proposition 34. The inverse of a bijective group homomorphism is itself a group homomorphism.

Proof. Let \( f : G \to G' \) be a bijective homomorphism. Hence it has an inverse (in the function-theoretic sense) \( f^{-1} \). For any \( g, h \in G \),

\[
gh = f(f^{-1}(g))f(f^{-1}(h)) = f(f^{-1}(g)f^{-1}(h)) \implies f^{-1}(gh) = f^{-1}(g)f^{-1}(h).
\]

\( \square \)

Definition 35 (Isomorphism). A bijective homomorphism is called an isomorphism.

Proposition 36. Let \( G \) and \( G' \) be finite groups, and let \( f : G \to G' \) be a group homomorphism. If \( f \) is injective (one-to-one) then \( |G| \leq |G'| \). If \( f \) is surjective (onto), then \( |G| \geq |G'| \).

Proof. Let \( f \) be injective. For each \( g \in G \), \( f(g) \in G' \) and \( f(g) = f(h) \), then \( g = h \). Hence \( |G| \leq |G'| \). If \( f \) is surjective, then for each \( g' \in G' \), there exists an element \( g \in G \) such that \( f(g) = g' \), hence \( |G| \geq |G'| \).

\( \square \)

Being isomorphic is an equivalence relation for groups. In particular, two groups that are isomorphic share the same group-theoretic properties. Here is an example:

Proposition 37. Let \( G \) and \( H \) be isomorphic groups. Then \( G \) is abelian if and only if \( H \) is abelian.

Proof. Suppose that \( G \) is abelian. Let \( f : G \to H \) be an isomorphism. Let \( h_1, h_2 \in H \). Then there exists exactly two elements \( g_1, g_2 \) such that \( f(g_i) = h_i \), \( i = 1, 2 \), and

\[
h_1h_2 = f(g_1)f(g_2) = f(g_1g_2) = f(g_2g_1) = f(g_2)f(g_1) = h_2h_1.
\]

The proof of the converse direction works in the same manner.

\( \square \)

Similarly, you could show that if \( G \cong H \), then \( G \) is cyclic if and only if \( H \) is cyclic, \( G \) is finite if and only if \( H \) is finite, etc.

3.3. Quotient groups

Consider a periodic function \( f(x) = f(x + 1) \), for all \( x \in \mathbb{R} \). Take as an example \( f(x) = \sin(2\pi x) \). Each \( x \in \mathbb{R} \) decomposes uniquely as \( x = \{x\} + [x] \) in a fractional part \( \{x\} \in [0, 1) \) and an integral part \( [x] \in \mathbb{Z} \). The periodicity relation \( f(x) = f(x + 1) \) implies that the value \( f(x) \) is completely determined by the fractional part \( \{x\} \) of \( x \), i.e. \( f(x) = f(\{x\}) \). In a sense, one could fold back the domain \( \mathbb{R} \) of \( f \) onto the subset \( [0, 1) \) and still have a complete understanding of the values of \( f \). In such case, one says that the function \( f \) quotients through \( \mathbb{R}/\mathbb{Z} \), the set of all real numbers modulo the equivalence relation \( x \sim x + 1 \), and which can be written

\[
\mathbb{R}/\mathbb{Z} = \{x + \mathbb{Z} : x \in \mathbb{R}\}.
\]

Observe that \( \mathbb{R}/\mathbb{Z} \) is a set of sets. Indeed, each element of \( \mathbb{R}/\mathbb{Z} \) is itself a set:

\[
x + \mathbb{Z} = \{x + n : n \in \mathbb{Z}\}.
\]

We can equip the set \( \mathbb{R}/\mathbb{Z} \) with a law of composition, the addition of sets:

\[
x + \mathbb{Z} + y + \mathbb{Z} = \{x + m + y + n : m, n \in \mathbb{Z}\} = x + y + \mathbb{Z} \in \mathbb{R}/\mathbb{Z}.
\]

Proposition 38. \( (\mathbb{R}/\mathbb{Z}, +) \) is a group.
Consider now instead of $x \mapsto \sin(2\pi x)$ the periodic map $x \mapsto e^{2\pi ix}$. This map defines a homomorphism from $(\mathbb{R}, +)$ to $(\mathbb{S}^1, \cdot)$, since $e^{2\pi i(x+y)} = e^{2\pi ix}e^{2\pi iy} \in \mathbb{S}^1$, and quotients through $\mathbb{R}/\mathbb{Z}$, since $e^{2\pi i(x+n)} = e^{2\pi ix}$ for all $x \in \mathbb{R}, n \in \mathbb{Z}$. Hence there is a well-defined homomorphism from $(\mathbb{R}/\mathbb{Z}, +)$ to $(\mathbb{S}^1, \cdot)$, given by

$$(\mathbb{R}/\mathbb{Z}, +) \to (\mathbb{S}^1, \cdot), \quad x + \mathbb{Z} \mapsto e^{2\pi ix}.$$ 

**Proposition 39.** The groups $(\mathbb{R}/\mathbb{Z}, +)$ and $(\mathbb{S}^1, \cdot)$ are isomorphic.

**Proof.** We show that the homomorphism above is bijective. Any $z \in \mathbb{S}^1$ can be written in polar coordinates as $z = e^{2\pi ix}$ for some $x \in [0,1)$. Hence the homomorphism is surjective. We see it is injective because $e^{2\pi ix} = 1$ if and only if $x \in \mathbb{Z}$. □

With the latter example in mind, we now define quotient groups abstractly. Let $G$ be a group, and let $H$ be a subgroup. Consider the quotient set

$$G/H = \{gH : g \in G\}$$

for which each element is a set $gH = \{gh : h \in H\}$, where the law of composition is the law of composition associated to the group $G$. Observe that this law of composition induces a law of composition on $G/H$, given by

$$g_1Hg_2H = g_1g_2H,$$

only if $Hg = gH$ for all $g \in G$. (Beware, the latter equality is an equality between sets, and should be understood as: $Hg = gH \iff g^{-1}hg \in H$ for all $g \in G, h \in H$.)

**Definition 40.** (Normal subgroup) A subgroup $N$ of $G$ is called normal if

$$g^{-1}ng \in N$$

for all $g \in G, n \in N$.

**Proposition 41.** Let $G$ be a group, and let $N$ be a normal subgroup of $G$. Then the quotient $G/N$ is a group, called a quotient group.

**Proof.** The quotient $G/N$ inherits the law of composition from $G$, and is well-defined since $N$ is normal. It is automatically associative, since $G$ is a group. The identity element is $N$, and for any $gN$, its inverse is given by $(gN)^{-1} = g^{-1}N$. □

**Proposition 42.** All subgroups of abelian groups are normal.

**Proof.** Let $G$ be abelian, and let $H$ be a subgroup of $G$. Let $g \in G, h \in H$. Then $g^{-1}hg = gg^{-1}h = h \in H$. Hence $H$ is a normal subgroup of $G$. □

**Proposition 43.** Let $G, G'$ be groups and let $f : G \to G'$ be a homomorphism. Then $\ker(f)$ is a normal subgroup of $G$.

**Proof.** Let $g \in G$ and $h \in \ker(f)$. Then $f(g^{-1}hg) = f(g)^{-1}f(h)f(g) = f(g)^{-1}ef(g) = e$, hence $g^{-1}hg \in \ker(f)$. □

We have seen that any isometry of $\mathbb{R}^n$ is of the form $f(x) = (T_b \circ A)(x) = Ax + b$, i.e. the composition of a translation $T_b$ and an orthogonal matrix $A \in \mathcal{O}(n)$. In terms of symmetry, translations are not interesting; they shift the space, but don’t exhibit any symmetry.
Proposition 44. Let $T(R^n)$ be the set of all translations of $R^n$, i.e.

\[ T(R^n) = \{ T_v : R^n \to R^n, T_v(x) = x + v : v \in R^n \}. \]

Then:

1. $(T(R^n), \circ)$ is a normal subgroup of $(\text{Isom}(R^n), \circ)$.
2. The quotient group $\text{Isom}(R^n)/T(R^n)$ is isomorphic to $(O(n), \cdot)$, the group of special orthogonal matrices with respect to matrix multiplication.

PROOF. (1) Exercise. (2) Consider the element 

\[ [f] = \{ f \circ g : g \in T(R^n) \} \subset \text{Isom}(R^n)/T(R^n) \]

for $f \in \text{Isom}(R^n)$, and decompose it into $f = T_v \circ A$ the composition of a translation $T_v$ and a rotation $A \in O(n)$. Observe that for any vector $w \in R^n$,

\[ (T_v \circ A)(w) = Aw + v = A(w + A^{-1}v) = (A \circ T_{A^{-1}v})(w). \]

Hence

\[ [f] = \{ f \circ g : g \in T(R^n) \} = \{ A \circ (T_{A-1}v \circ g) : g \in T(R^n) \} = \{ A \circ h : h \in T(R^n) \}. \]

To finish the proof, show that the map $\varphi([f]) = A$ is an isomorphism. \qed

3.4. Exercises

Exercise 1: Let $G$ be a group.

1. What is the inverse of the identity element $e \in G$?
2. Show that $(gh)^{-1} = h^{-1}g^{-1}$.
3. Show that if $g^{-1} = g$ for all $g \in G$, then $G$ is abelian.

Exercise 2: In which of the following cases is $H$ a subgroup of $G$?

1. $G = \text{GL}_n(C) = \{ A \in M_n(C) : \det(A) \neq 0 \}$ with matrix multiplication, and $H = \text{GL}_n(R)$.
2. $G = R^\ast = R \setminus \{0\}$ with multiplication, and $H = \{ -1, 1 \}$.
3. $G = Z$ with addition, and $H = Z_{\geq 0}$, the set of nonnegative integers.
4. $G = R^\ast$ with multiplication, and $H = R_{> 0}$, the of positive reals.

Exercise 3: Show that the (vector) space of orthogonal $n \times n$ matrices

\[ O(n) = \{ A \in M_n(R) : A^T A = I \} \]

is a group with respect to matrix multiplication.

Exercise 4: Describe all groups that contain no proper subgroup.

Exercise 5: Let $\varphi : Z \to Z$ be a homomorphism.

1. Show that the image of $\varphi$ is exactly $\langle \varphi(1) \rangle$.
2. Using the previous fact, determine which homomorphisms $\varphi : Z \to Z$ are injective, which are surjective, and which are isomorphisms.
3. Prove that all subgroups of $Z$ are of the form $nZ$ for some $n \geq 1$. 
Exercise 6: Consider the set of all real Möbius transformations, that is
\[ G = \left\{ f(x) = \frac{ax + b}{cx + d} : a, b, c, d \in \mathbb{R}, \ ad - bc \neq 0 \right\} \]
together with \( \circ \), the composition of functions.

1. Show that \((G, \circ)\) is a group.
2. Find a subgroup of \( G \) that is isomorphic to \((\mathbb{R}, +)\).
3. Show that \( \varphi : (\text{SL}_2(\mathbb{R}), \cdot) \to (G, \circ), \left( \begin{array}{cc} a & b \\ c & d \end{array} \right) \mapsto \frac{ax + b}{cx + d} \) is a group homomorphism, where
   \[ \text{SL}_2(\mathbb{R}) = \left\{ \left( \begin{array}{cc} a & b \\ c & d \end{array} \right) : a, b, c, d \in \mathbb{R}, \ ad - bc = 1 \right\} \]
is the special linear group of real 2\times2 matrices, with respect to matrix multiplication.
4. Determine the kernel and the image of \( \varphi \).
5. Determine all Möbius transformations of order 2, i.e. all \( f \in G \) such that \( f^2 = e \).

Exercise 7: Find all subgroups of \( \text{Isom}(\triangle) \). Which are normal? Describe the resulting quotient groups.

Exercise 8: Let \((p, q)\) denote the greatest common divisor of two integers \( p, q \in \mathbb{Z} \). Prove that \((\langle p, q \rangle, +) = ((\langle p, q \rangle, +)\). Then prove that all subgroups of \((\mathbb{Z}, +)\) are of the form \((n\mathbb{Z}, +)\) for some \( n \geq 1 \).

Exercise 9:

1. Show that \(O(n)/SO(n)\) is a group and that the determinant map induces an isomorphism \(O(n)/SO(n) \cong (\{\pm 1\}, \cdot)\).
2. Show that \((\{\pm 1\}, \cdot)\) is a cyclic group of order two.
3. Using that ‘being isomorphic’ is an equivalence relation (and in particular transitive), show that \(O(n)/SO(n)\) is isomorphic to any cyclic subgroup \(\langle S \rangle \subset O(n)\) generated by a reflection.
4. Conclude that up to reflection, any orthogonal matrix is a rotation (i.e. is in \(SO(n)\)).

Exercise 10: Show that \(\text{Isom}^+(\mathbb{R}^n)/T(\mathbb{R}^n)\) is isomorphic to \(SO(n)\).
CHAPTER 4

Groups of isometries

Proposition 45. If \( A \subset \mathbb{R}^n \), then the set of isometries of \( A \)
\[
\text{Isom}(A) = \{ f \in \text{Isom}(\mathbb{R}^n) : f(A) = A \}
\]
is a subgroup of \( \text{Isom}(\mathbb{R}^n) \).

Proof. Let \( f, g \in \text{Isom}(A) \). Then \( g(x) \in A \) and \( f(g(x)) \in A \). Since \( f \) is an isometry, it is invertible, and since \( f(A) = A \), \( f^{-1}(A) = A \). \( \square \)

The heuristic for the study of Euclidean symmetries lies in the correspondence
subgroups of \( \text{Isom}(\mathbb{R}^n) \) \( \longleftrightarrow \) symmetries of figures in \( \mathbb{R}^n \).

4.1. Dihedral groups

Let \( n \in \mathbb{N} \), and let \( P_n \) denote the regular \( n \)-gon. We have already seen that
\[
\text{Isom}(P_3) = \text{Isom}(\triangle) = \{ e, R_3, R_3^2, S, SR_3, SR_3^2 \}
\]
where \( R_3 \) is the counterclockwise rotation of \( \triangle \) by \( \frac{2\pi}{3} \), and \( S \) is the reflection across a line bisecting one of the angles.

Definition 46 (Dihedral group). The finite dihedral group of order \( 2n \) is the group
\[
D_n = \langle R_n, S \rangle
\]
generated by the rotation \( R_n = R_{\frac{2\pi}{n}} \) around the origin by an angle of \( \frac{2\pi}{n} \) and the reflection \( S \) across one of the corners of \( P_n \), with respect to the composition of functions.

Lemma 47. \( R_nSR_n = S \).

Proof. Exercise. \( \square \)

Proposition 48.
\[
D_n = \{ e, R_n, R_n^2, \ldots, R_n^{n-1}, S, SR_n, \ldots, SR_n^{n-1} \}
\]

Proof. The generators \( R_n \) and \( S \) have the following properties:
\[
R_n^n = S = e, \quad R_nSR_n = S.
\]

Any element \( g \in D_n \) is a word in the generators \( R_n, S \). That is, it is of the form
\[
g = R_n^{k_1}SR_n^{k_2}SR_n^{k_3} \cdots SR_n^{k_m},
\]
where \( k_1, k_m \in \{0, \ldots, n-1\} \), \( k_2, \ldots, k_{m-1} \in \{1, \ldots, n-1\} \). Then
\[
g = (R_n^{k_1}S)R_n^{k_2}SR_n^{k_3} \cdots SR_n^{k_m} = (SR_n^{-k_1})R_n^{k_2}S \cdots R_n^{k_m} = S^2R_n^{-k_2+k_3}SR_n^{k_4} \cdots R_n^{k_m} = R_n^{k_1-k_2+k_3}SR_n^{k_4} \cdots R_n^{k_m} = \ldots = S^aR_n^b
\]

In particular, each element of $D_n$ is either a rotation or a reflection. From there, it is easy to see that $D_n$ is a group, and moreover – from the presentation above –, that it is a finite group of order $2n$. Finally, it follows from the classification theorem of isometries of the plane that $D_n = \text{Isom}(P_n)$.

4.2. Finite subgroups of $SO(2)$

Recall that matrices of $SO(2)$ are rotations. Let $G$ be a nontrivial finite subgroup of $SO(2)$, and let $n = |G|$ be its order. Let $R_\alpha \in G$ be some non-trivial rotation, by an angle $\alpha \in \mathbb{R}/2\pi\mathbb{Z}$. Then $G$ contains at least

$$R_\alpha, R_\alpha^2, \ldots, R_\alpha^n = \text{id} \subset G$$

and the existence of some $m$ such that $R_\alpha^m = \text{id}$ follows from the fact that $G$ is finite. Choose $R_\alpha$ to be the minimal rotation in $G$, that is the rotation of smallest angle.

**Claim 49.** $G = \langle R_\alpha \rangle$

**Proof.** Let $R_\theta \in G$ be an arbitrary element of $G$. By the division algorithm, we may write $\theta$ as

$$\theta = n\alpha + \beta$$

for some $n \in \mathbb{N}$, $\beta \in [0, \alpha)$. Then

$$R_\theta = R_\alpha^n \circ R_\beta \implies R_\beta = R_\alpha^{-n} \circ R_\theta \in G \implies \beta = 0$$

by minimality of the angle $\alpha$. Hence $R_\theta = R_\alpha^n$. \hfill $\square$

Observe that this implies that the angle $\alpha$ for the generating rotation $R_\alpha$ must be

$$\alpha = \frac{2\pi}{n}.$$

Note also that if the angle of rotation is an irrational multiple of $2\pi$, then $\langle R_\alpha \rangle$ is infinite. (Why?) We have shown:

**Theorem 3.** Any finite subgroup $G \subset SO(n)$, $|G| = n$, is cyclic, generated by a rotation of angle $\frac{2\pi}{n}$.

Consider two rational (i.e. by a rational multiple of $2\pi$) rotations $R_\alpha$, $R_\beta$, and the group $\langle R_\alpha, R_\beta \rangle$ they generate. Does the theorem imply that this group is infinite? No;

$$\langle R_\alpha, R_\beta \rangle = \{R_\alpha^m \circ R_\beta^n : m, n \in \mathbb{Z}\} = \{R_{m\alpha + n\beta} : m, n \in \mathbb{Z}\}$$

We are assuming that $\alpha$ and $\beta$ are rational multiples of $2\pi$; write $\alpha = \frac{p}{q}$, $\frac{r}{s}$. Then

$$\langle R_\alpha, R_\beta \rangle = \{R_{\frac{mp + nr}{qs}} : m, n \in \mathbb{Z}\} = \{R_{\frac{\gamma k}{qs}} : k \in \mathbb{Z}\} = \langle R_\gamma \rangle$$

for $\gamma = 2\pi \frac{1}{qs}$.
4.3. Finite subgroups of $SO(3)$

We first describe matrices in $SO(3)$.

**Lemma 50.** If $A \in SO(3)$ then 1 is an eigenvalue of $A$.

**Proof.** We want to show that $\det(A - I) = 0$. This follows from

$$\det(A - I) = \det(A - I) \det(A^T) = \det(AA^T - A^T) = (-1)^3 \det(A - I) = -\det(A - I).$$

Hence $A$ can be conjugated to a matrix of the form

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{pmatrix}.$$ 

Let $C = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. One checks immediately that $B^T B = I$ implies that $C^T C = I$ and $\det(B) = \det(C) = 1$. Hence $C \in SO(2)$ and

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}.$$ 

Observe that $B$ is the rotation of the plane $y-z$ around the $x$-axis.

**Proposition 51.** Any matrix of $SO(3)$ is a rotation of $\mathbb{R}^3$ around an axis passing through the origin.

**Theorem 4.** Any finite subgroup of $SO(3)$ is isomorphic to one of the following groups:

1. $C_n$, $n \geq 1$
2. $D_n$, $n \geq 2$
3. $\text{Isom}^+(T)$, where $T$ is the regular tetrahedron
4. $\text{Isom}^+(C)$, where $C$ is the cube
5. $\text{Isom}^+(I)$, where $I$ is the regular icosahedron

Let $g \in SO(3)$, $g \neq 0$. Consider the action of the rotation $g$ on the sphere $S^2 = \{x \in \mathbb{R}^3 : \|x\| = 1\}$; $g$ rotates all points of $S^2$ around an axis, and leaves two points fixed, namely the two points on the axis where the axis intersects the sphere. If $g$ is the rotation around the $z$-axis, then these two points are the north and south poles of $S^2$. More generally, let $G \subset SO(3)$ be a finite subgroup and define

$$P = \{x \in S^2 : \exists g \in SO(3), g \neq e \text{ s.t. } g(x) = x\}.$$ 

Since the group $G$ is finite, $P$ must be a finite set. We call $P$ the set of poles of $G$. To prove the classification theorem above, we need to understand how $G$ acts on $P$.

### 4.4. Group actions

**Definition 52.** If $G$ is a group and $S$ is a non-empty set, we say that $G$ acts on $S$ (we write $G \curvearrowright S$) if there is a well-defined map $G \times S \to S$ such that $(gh)(s) = g(h(s))$ and $e(s) = s$.

The general example to motivate this definition for us will be the following. Let $A \subset \mathbb{R}^n$. The action map is $\text{Isom}(A) \times A \to A, (f, x) \mapsto f(x)$, and is well-defined since $f(A) = A$. The two formal requirements to have a group action are here obvious from the definitions:

$$(f \circ g)(x) = f(g(x)) \quad \text{id}(x) = x.$$
Definition 53. An action \( G \curvearrowright S \) is transitive if for any two elements \( x, y \in S \), there exists a group element \( g \in G \) such that \( g(x) = y \).

For example, the group of translations \( T(\mathbb{R}^n) \) acts transitively on \( \mathbb{R}^n \). On the other hand, \( SO(2) \) does not act transitively on \( \mathbb{R}^2 \). Recall that any element in \( SO(2) \) is a rotation around the origin. In particular the origin is fixed by \( SO(2) \) while any other point can only be moved along the circle centered at the origin on which it lies.

Definition 54. Let \( G \curvearrowright S \) and let \( x \in S \). The orbit of \( x \) for the \( G \)-action is the set

\[
\mathcal{O}(x) = \{ g(x) : g \in G \}.
\]

The stabilizer of \( x \) for the \( G \)-action is the subgroup

\[
\text{Stab}(x) = \{ g \in G : g(x) = x \}.
\]

Let \( x \in \mathbb{R}^2 \). Then the orbit of \( x \) under the action of the group of translations of \( \mathbb{R}^2 \) is \( \mathcal{O}(x) = \mathbb{R}^2 \). The orbit of \( x \) under the action of \( SO(2) \) is the circle centered at the origin on which \( x \) lies (or \( \{0\} \) if \( x = 0 \)). The stabilizer of \( x \) under the action of \( SO(2) \) is \( \text{Stab}(x) = \{1\} \) if \( x \neq 0 \) or \( SO(2) \) if \( x = 0 \).

Exercise 55. Show that a group action is transitive if and only if it has only one orbit.

From these two examples, we observe that \( \mathbb{R}^2 \) can be completely ‘covered’ by the orbits of the group action. More generally:

Proposition 56. If \( G \curvearrowright S \), then \( S \) is the disjoint union of all orbits \( \mathcal{O}(x) \), \( x \in S \).

Proof. We will only show that if \( \mathcal{O}(x) \cap \mathcal{O}(y) \neq \emptyset \), then \( \mathcal{O}(x) = \mathcal{O}(y) \). By assumption, let \( f(x) = g(y) \) be an intersection point of the two orbits. Then \( (g^{-1}f)(x) = y \). This shows that \( y \in \mathcal{O}(x) \). We claim that \( \mathcal{O}(y) \subseteq \mathcal{O}(x) \). Let \( h(y) \in \mathcal{O}(y) \). Then \( h(y) = (hg^{-1}f)(x) \in \mathcal{O}(x) \). One shows that \( \mathcal{O}(x) \subseteq \mathcal{O}(y) \) in the same way. \( \square \)

Exercise 57. Show that the group \( \text{Isom}(\mathbb{R}^2) \) acts on the set \( S \) of all triangles in \( \mathbb{R}^2 \). Let \( \triangle \in S \). Then \( \text{Stab}(\triangle) = \text{Isom}(\triangle) \). Show that the orbit \( \mathcal{O}(\triangle) \) is the set of all triangles in \( \mathbb{R}^2 \) that are congruent to \( \triangle \).

Theorem 5 (Orbit-stabilizer theorem). Let \( G \) be a finite group, and let \( S \) be a set, such that \( G \) acts on \( S \). Let \( x \in S \). Then \( |G| = |\mathcal{O}(x)||\text{Stab}(x)| \).

Proof. Consider the evaluation map

\[
\varphi : G \to \mathcal{O}(x), \ g \mapsto g(x).
\]

If \( y \in \mathcal{O}(x) \), then its preimage under \( \varphi \) is

\[
\varphi^{-1}(y) = \{ g \in G : g(x) = y \}.
\]

The group \( G \) is covered by the union

\[
G = \bigcup_{y \in \mathcal{O}(x)} \varphi^{-1}(y).
\]

We claim that this union is disjoint. Suppose that there exist \( y, y' \in \mathcal{O}(x) \) such that \( \varphi^{-1}(y) \cap \varphi^{-1}(y') \neq \emptyset \). That is, there exists a group element \( g \in G \) such that \( g(x) = y \) and \( g(x) = y' \). Then \( y = y' \). This proves that \( G \) can be written as disjoint union of preimages of \( \varphi \), and

\[
|G| = \sum_{y \in \mathcal{O}(x)} |\varphi^{-1}(y)|
\]
Observe that trivially \( x \in \mathcal{O}(x) \) since \( e(x) = x \). Then \( \varphi^{-1}(x) = \text{Stab}(x) \). Let \( g \in G \). Then
\[
\varphi^{-1}(g(x)) = \{ f \in G : f(x) = g(x) \} = \{ f \in G : (g^{-1}f)(x) = x \}.
\]
Observe that there is a one-to-one correspondence between the latter set and \( \text{Stab}(x) \). (Exercise: write this correspondence down.) This implies \( |\varphi^{-1}(g(x))| = |\text{Stab}(x)| \). Since any \( y \in \mathcal{O}(x) \) is of the form \( y = g(x) \),
\[
|G| = \sum_{y \in \mathcal{O}(x)} |\varphi^{-1}(y)| = |G| = \sum_{y \in \mathcal{O}(x)} |\text{Stab}(x)| = |\text{Stab}(x)| \sum_{y \in \mathcal{O}(x)} 1 = |\text{Stab}(x)||\mathcal{O}(x)|.
\]

**Exercise 58.** Use the orbit-stabilizer theorem to prove Lagrange’s theorem: if \( H \) is a subgroup of a finite group \( G \), then \( |H| \) divides \( |G| \).

**Proposition 59.** Let \( T \) denote the regular tetrahedron. Then \( |\text{Isom}^+(T)| = 12 \).

**Proof.** Since an isometry preserves lines and angles, it also preserve the corners of \( T \). Since there are only four corners, the group \( \text{Isom}^+(T) \) is finite, and it acts on the set \( S \) of the corners of \( T \). We may thus apply the orbit-stabilizer theorem. Choose \( x \) to be the top corner of \( T \):

![Regular Tetrahedron](image)

The group \( \text{Isom}^+(T) \) can not contain any translation. (More generally, if \( A \subset \mathbb{R}^n \) is a bounded domain, then any isometry of \( A \) must preserve it and a translation won’t.) Hence any element in the group can be represented by a matrix in \( O(3) \), and more precisely in \( SO(3) \), since we restrict to orientation-preserving isometries. The orbit \( \mathcal{O}(x) = S \) since any corner of \( T \) can be reached from \( x \) by a rotation. Similarly, the only rotations that fix \( x \) are the rotations of the base of \( T \); the rotations by 120, 240, and 360 degrees. Hence \( |G| = 4 \cdot 3 = 12 \).

**Exercise 60.** Let \( C \) be the cube, \( I \) the regular icosahedron. Show that \( |\text{Isom}^+(C)| = 24 \), \( |\text{Isom}^+(I)| = 60 \).

### 4.5. Proof of the classification theorem

**Theorem 6.** Any finite subgroup of \( SO(3) \) is isomorphic to one of the following groups:

1. \( C_n \), \( n \geq 1 \)
2. \( D_n \), \( n \geq 2 \)
3. \( \text{Isom}^+(T) \), where \( T \) is the regular tetrahedron
4. \( \text{Isom}^+(C) \), where \( C \) is the cube
5. \( \text{Isom}^+(I) \), where \( I \) is the regular icosahedron

Let \( g \in SO(3) \), \( g \neq 0 \). Consider the action of the rotation \( g \) on the sphere \( S^2 = \{ x \in \mathbb{R}^3 : \|x\| = 1 \} \); \( g \) rotates all points of \( S^2 \) around an axis, and leaves two points fixed, namely the two points on the axis where the axis intersects the sphere. If \( g \) is the rotation around
the z-axis, then these two points are the north and south poles of $S^2$. More generally, let $G \subset SO(3)$ be a finite subgroup and define 
\[ P = \{ x \in S^2 : \exists g \in SO(3), \ g \neq e \ \text{s.t.} \ g(x) = x \}. \]
Since the group $G$ is finite, $P$ must be a finite set. We call $P$ the set of poles of $G$. To prove the classification theorem above, we need to understand how $G$ acts on $P$.

**Claim 61.** $G \rtimes P$.

**Proof.** We first note that for any $x \in S^2$, $(gh)(x) = g(h(x))$ and $e(x) = x$ hold. We show that $(g, x) \mapsto g(x)$ is well-defined, i.e. that $g(x) \in P$. Since $x \in P$, there exists $h \in G$ non-trivial such that $h(x) = x$. Then $(ghg^{-1})(g(x)) = g(h(x)) = g(x)$, i.e. $g(x) \in P$. □

We will start by counting the pairs
\[ \#\{(g, x) \in G \times P : g \neq e, \ g(x) = x\}. \]
Observe that if we fix $g \neq e$ in $G$, then
\[ \#\{x \in P : g(x) = x\} = 2 \]
counts the two poles fixed by the axis around which $g$ rotates. Adding this over all elements of $G$ minus the identity element,
\[ \#\{(g, x) \in G \times P : g \neq e, \ g(x) = x\} = 2(|G| - 1). \]
On the other hand, if we fix $x \in P$, then
\[ \#\{g \in G : g \neq e, \ g(x) = x\} = |\text{Stab}(x)| - 1, \]
and thus
\[ 2(|G| - 1) = \sum_{x \in P} (|\text{Stab}(x)| - 1). \]
We can expand the sum on the right-hand side (RHS) further. Recall that $P$ decomposes into a disjoint union of orbits of the $G$-action (Proposition 57). Since $P$ is finite, there can only be a finite number of orbits, and we may write
\[ P = O(x_1) \cup O(x_2) \cup \cdots \cup O(x_k). \]

**Lemma 62.** If $y \in O(x)$, then $O(y) = O(x)$.

**Proof.** By definition, there exists $g \in G$ such that $g(x) = y$. Let $h \in G$. Then $h(y) = (hg)(x) \in O(x)$ and $h(x) = (hg^{-1})(y) \in O(y)$. □

By the orbit-stabilizer theorem, the lemma implies that $|\text{Stab}(x)| = |\text{Stab}(y)|$. Hence
\[ 2(|G| - 1) = \sum_{x \in P} (|\text{Stab}(x)| - 1) = \sum_{i=1}^{k} \sum_{x \in O(x_i)} (|\text{Stab}(x)| - 1) = \sum_{i=1}^{k} |O(x_i)| (|\text{Stab}(x_i)| - 1). \]
By dividing both sides of the equation by $|G|$,
\[ 2 \left( 1 - \frac{1}{|G|} \right) = \sum_{i=1}^{k} \left( 1 - \frac{1}{|\text{Stab}(x_i)|} \right). \]
We will assume that $G$ is non-trivial, hence $|G| \geq 2$ and the LHS is $< 2$. Since $x_i \in P$, $|\text{Stab}(x_i)| \geq 2$ (because there exists some $g \neq e$ such that $g(x_i) = x_i$). Hence the RHS is $\geq k/2$. In conclusion, $1 \leq k \leq 3$ – there are at most three orbits.
Suppose there is only $k = 1$ orbit. The equation becomes

$$2 \left( 1 - \frac{1}{|G|} \right) = 1 - \frac{1}{|\text{Stab}(x_1)|}.$$  

But now the RHS is $< 1$ while the LHS is $\geq 1$. This is impossible. Hence $2 \leq k \leq 3$. Suppose there are $k = 2$ orbits,

$$2 \left( 1 - \frac{1}{|G|} \right) = 1 - \frac{1}{|\text{Stab}(x_1)|} + 1 - \frac{1}{|\text{Stab}(x_2)|} = 2 - \frac{1}{|\text{Stab}(x_1)|} - \frac{1}{|\text{Stab}(x_2)|}.$$  

By Lagrange’s theorem, since $\text{Stab}(x_i)$ is a subgroup of $G$, $\frac{|G|}{|\text{Stab}(x_i)|} \in \mathbb{N}$, and

$$2 = \frac{|G|}{|\text{Stab}(x_1)|} + \frac{|G|}{|\text{Stab}(x_2)|}$$  

implies that $|\text{Stab}(x_1)| = |\text{Stab}(x_2)| = |G|$. This says that $G$ has two fixed points $x_1$ and $x_2$. Then $G$ must be generated by rotations around the axis that passes through $x_1$ and $x_2$. Since $G$ is finite, it follows (exactly as in the proof of the classification theorem for $SO(2)$) that $G$ is generated by a single rotation, of angle $\frac{2\pi}{|G|}$. Hence $G \cong C_{|G|}$.

Finally, suppose that there are $k = 3$ orbits. Then

$$1 + \frac{2}{|G|} = \frac{1}{|\text{Stab}(x_1)|} + \frac{1}{|\text{Stab}(x_2)|} + \frac{1}{|\text{Stab}(x_3)|}.$$  

Since the LHS is $> 1$, at least one of the $\text{Stab}(x_i)$ must contain only two elements. We will assume without limitation of generality (wlog) that $|\text{Stab}(x_1)| = 2$. Then we are left with

$$\frac{1}{2} + \frac{2}{|G|} = \frac{1}{|\text{Stab}(x_2)|} + \frac{1}{|\text{Stab}(x_3)|}. \quad (4.1)$$  

By the same reasoning, since the LHS is $> 1/2$, at least one of $\text{Stab}(x_i)$ must contain $\leq 3$ elements.

Suppose first that wlog $|\text{Stab}(x_2)| = 2$. Then $|\text{Stab}(x_3)| = |G|/2$ – this implies that $|G|$ must be even. So in this case, we have

$$|G| \text{ even}, \quad |\text{Stab}(x_1)| = 2, \quad |\text{Stab}(x_2)| = 2, \quad |\text{Stab}(x_3)| = \frac{|G|}{2}.$$  

By the above, $x_3$ has an orbit of 2 poles, call them $A$ and $B$. Then the stabilizer of $A$ has order $|G|/2$ and hence the other half of the elements of $G$ interchange $A$ and $B$. So the elements of $G$ are either rotations around the line $L$ passing through $A$ and $B$, or rotations by angle $\pi$ around a line perpendicular $L$. We conclude that $G$ preserves the regular polygon $P_n$ with $n = |G|/2$. Hence $G \cong D_n$.

Finally suppose that wlog $|\text{Stab}(x_2)| = 3$, and $|\text{Stab}(x_3)| \geq 3$. Then Eq. (4.2) becomes

$$\frac{1}{6} + \frac{2}{|G|} = \frac{1}{|\text{Stab}(x_3)|}. \quad (4.2)$$  

Since the LHS is $> 1/6$, $|\text{Stab}(x_3)| \leq 5$. We have three possibilities to explore: $|\text{Stab}(x_3)| = 3, 4, 5$. Using Eq. (4.2) we can fill out the following table:

| $|\text{Stab}(x_1)|$  | $|\text{Stab}(x_2)|$  | $|\text{Stab}(x_3)|$  | $|G|$ |
|------------------|------------------|------------------|-----|
| 2                | 3                | 3                | 12  |
| 2                | 3                | 4                | 24  |
| 2                | 3                | 5                | 60  |
30 4. GROUPS OF ISOMETRIES

We’ll only look at the first case; \( x_3 \) has an orbit of \( 12/3 = 4 \) poles, call them \( A, B, C, D \). The stabilizer of \( A \) has order 3 and acts by rotation on \( B, C, D \). This means that \( B, C, D \), when connected, form an equilateral triangle. Arguing in the same way, any triple in \( \{A, B, C, D\} \) gives rise to an equilateral triangle. Hence \( A, B, C, D \) are the corners of a regular tetrahedron \( T \). Since \( \text{SO}(3) \) preserves \( T \), and \( G \subset \text{SO}(3) \), it follows that \( G \subset \text{Isom}^+(T) \). Then \( |G| = 12 = |\text{Isom}^+(T)| \) implies that \( G = \text{Isom}^+(T) \).

4.6. Exercises

Exercise 1: The dihedral group \( D_n \) is generated by a reflection \( S \) and a rotation \( R_n \). Show that \( R_nSR_n = S \), for each \( n \in \mathbb{N} \).

Exercise 2: Show that the action of \( \text{Isom}(\mathbb{R}^n) \) on \( \mathbb{R}^n \) is transitive.

Exercise 3: Show that the group \( \text{Isom}(\mathbb{R}^n) \) acts on the set of lines in \( \mathbb{R}^n \). Determine the stabilizer of a line.

Exercise 4: The group \( \text{GL}_n(\mathbb{R}) \) acts on \( \mathbb{R}^n \) by matrix multiplication on the left.

(1) Describe the decomposition of \( \mathbb{R}^n \) into orbits.
(2) What is the stabilizer of \( e_1 = (1 0 \ldots 0) \)?
(3) Show that \( \text{Stab}(e_1) \) is a normal subgroup.
(4) Describe the quotient group \( \text{GL}_n(\mathbb{R})/\text{Stab}(e_1) \).

Exercise 5: The translation of \( \mathbb{R}^n \) by a vector \( v \in \mathbb{R}^n \) is described by the map
\[
T_v : \mathbb{R}^n \to \mathbb{R}^n, \quad T_v(x) = x + v.
\]
Let \( T(\mathbb{R}^n) = \{T_v : v \in \mathbb{R}^n\} \) be the set of all translations of \( \mathbb{R}^n \).

(1) Show that \( (T(\mathbb{R}^n), \circ) \) is a normal subgroup of \( (\text{Isom}^+(\mathbb{R}^n), \circ) \).
(2) Show that \( (T(\mathbb{R}^n), \circ) \cong (\mathbb{R}^n, +) \).

Exercise 6: Let \( I \) denote the regular icosahedron. Compute \( |\text{Isom}^+(I)| \).

Exercise 7: Let \( L_1 \) and \( L_2 \) be two lines in \( \mathbb{R}^2 \) that intersect with an angle of \( \alpha = \frac{2\pi}{n} \) \( (n \in \mathbb{N}) \).
Let \( S_1 \) and \( S_2 \) be the reflections across \( L_1 \) resp. \( L_2 \). Describe the subgroup \( \langle S_1, S_2 \rangle \subset O(2) \).

Exercise 8: Let \( G \) act transitively on a finite set \( S \). Show that if \( S \) contains more than two elements, then there exists \( g \in G \) with no fixed point.

Exercise 9: Let \( G \) be a finite subgroup, and let \( H \subset G \) be a subgroup. In this exercise, we prove Lagrange’s theorem; if \( G \) is a finite group then the order of any subgroup \( H \subset G \) divides the order of \( G \), i.e. there exists \( n \in \mathbb{N} \) such that \( |G| = n|H| \).

(1) Show that \( H \) acts on \( G \) by left composition, i.e. \( H \times G \to G, (h, g) \mapsto hg \) defines a group action.
(2) What is the orbit \( O_H(e) \) of the identity element \( e \in G \)?
(3) Let \( g \in G \) be distinct from \( e \). Show that the map \( H \to O_H(g), h \mapsto hg \) is bijective.
(4) Prove Lagrange’s theorem.
Exercise 10: In this exercise, we will prove the first Sylow theorem. Sylow’s theorems (there are three) offer a partial converse to Lagrange’s theorem, namely they partially answer the following question; if \( n \) divides the order of a finite group \( G \), does \( G \) contain a subgroup \( H \) of order \( n \)?

**Theorem 63 (First Sylow Theorem).** Let \( G \) be a finite group, whose order is \( |G| = p^n h \) for a prime number \( p \) and \( h \in \mathbb{N} \) not divisible by \( p \). Then \( G \) has a subgroup of order \( p^n \).

1. Consider the set \( M = \{ S \subset G : |S| = p^n \} \) of subsets of \( G \) of cardinality \( p^n \), and show that \( g.S = \{ gs : s \in S \} \) defines an action of \( G \) on \( M \).
2. Show that \( p \) does not divide \( |M| \).
3. Deduce that there exists an orbit \( O(S) \) of the action such that \( p \) does not divide \( |O(S)| \).
4. Show that \( G = \bigcap_{S' \in O(S)} S' \).
5. Deduce that \( |O(m)| = h \) and conclude that \( G \) has a subgroup of order \( p^n \).
CHAPTER 5

Projective geometry

Projective geometry is the study of the geometric properties that are preserved by projections, in the same way as Euclidean geometry is the study of geometric properties that are preserved by isometries.

5.1. Perspective

The study of projective geometry initiates with the question of how to represent 3-dimensional objects on a canvas, or a sheet of paper, or: a plane. An example we are familiar with is the cube under ‘parallel projection:’

Observe that

- Parallel edges are represented by parallel lines
- Straight lines are represented by straight lines
- Midpoints (the midpoint of each face, the midpoint of each edge) are preserved

On the other hand, this is not very realistic – depending on where I stand with respect to the cube, it looks different to me. One can see a cube from below, from above, frontally, depending on the position of the cube in space relative to the observer. A way to represent the cube ‘realistically’ is through ‘perspective projection:’

Although this representation seems more realistic, it clearly does not preserve parallel lines anymore. In the following example, we – the observer – are standing on railroad tracks staring to the horizon. In our perspective, the tracks, which are obviously parallel, seem to converge to a vanishing point on the horizon.
However, if we – the observer – go to look down at our feet, the tracks again will again look like they are parallel to one another.

The main properties of perspective were summarized by the Florentine architect Brunelleschi (1377–1446) as follow:

- The horizon appears as a line
- Straight lines are preserved
- Sets of parallel lines meet at a vanishing point
- Lines that are parallel to the picture appear parallel and hence do not converge to a vanishing point.

Let us consider as a last example the planar figure

Observing how it is projected once in perspective

we can observe that

- Straight lines, intersections and tangency points are preserved,
- Angles, circles, midpoints, parallel lines are not preserved

under perspective projections. Formally, perspective projection is defined as follows:

**Definition 64.** Let $\Pi \subset \mathbb{R}^3$ be a plane, $O \notin \Pi$ be an observer, $X \subset \mathbb{R}^3$ be an object. The perspective of $X$ on $\Pi$ as seen from $O$ is obtained by the projection

$$\pi : X \to \Pi, \ P \mapsto P' = OP \cap \Pi,$$
where \( OP \) is the straight line connecting \( O \) to \( P \).

We illustrate this definition in light of the railroad track picture.

In this figure, the observer’s eye is at \( O \), the plane of projection \( \Pi \) is the \( y\)-\( z \)-plane in \( \mathbb{R}^3 \). Again, the observer is standing on a (very schematic) railroad track, which runs on the \( x\)-\( y \)-plane along the \( x \)-axis.

Continuing with this example, we can give an explicit parametrization of the map \( \pi : X \to \Pi \),

\[
P = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \mapsto P' = OP \cap \Pi = \begin{pmatrix} 0 \\ y' \\ z' \end{pmatrix}.
\]

Suppose that \( O = \begin{pmatrix} x_0 \\ 0 \\ z_0 \end{pmatrix} \). In terms of vector geometry,

\[
OP = \begin{pmatrix} x - x_0 \\ y \\ -z_0 \end{pmatrix}, \quad \text{and} \quad OP' = \begin{pmatrix} -x_0 \\ y' \\ z' - z_0 \end{pmatrix}.
\]

Our goal is to express \( y', z' \) in terms of \( x, y, x_0, z_0 \) such as to give an explicit formula for the map \( \pi \) above. To do so, we need to relate these variables by a system of equations. Observe that trivially, since \( OP \) and \( OP' \) lie on the same line, the cross-product \( OP \times OP' \) must be 0, i.e.

\[
OP \times OP' = \begin{pmatrix} x - x_0 \\ y \\ -z_0 \end{pmatrix} \times \begin{pmatrix} -x_0 \\ y' \\ z' - z_0 \end{pmatrix} = \begin{pmatrix} y(z' - z_0) + z_0y' \\ -(x - x_0)(z' - z_0) + z_0x_0 \\ (x - x_0)y' + x_0y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.
\]

Using the last two coordinates, if \( x \neq x_0 \),

\[
y' = \frac{-x_0y}{x - x_0}, \quad \text{and} \quad z' = \frac{z_0x}{x - x_0}.
\]

Remarks:

1. This solution exists only if \( x \neq x_0 \). Intuitively, this marks the limit of the observer’s peripheral vision; the observer can only observe what is in front of him.
2. We can read off what happens as the observer stares at the railroad track at the horizon; then as \( x \to -\infty \), \( y' \to 0 \), \( z' \to z_0 \).
Hence the projection map is given by

\[
\pi \left( \begin{array}{c} x \\ y \\ 0 \\ 0 \\ \end{array} \right) = \left( \begin{array}{c} 0 \\ -x_0 y/(x-x_0) \\ z_0 x/(x-x_0) \\ \end{array} \right)
\]

whenever \( x \neq x_0 \).

A more general computation along the same lines shows that in general, coordinates of projected points are rational fractions. As a result, the projection map is not linear, which prevents us from applying the usual tools from vector geometry/linear algebra. Fortunately, there is an easy way to correct this situation; replace \( P \) by

\[
\left( \begin{array}{c} x \\ y \\ 0 \\ 1 \\ \end{array} \right) \in \mathbb{R}^4
\]

and consider instead of \( P' \) the line

\[
L_{P'} = \left\{ \lambda \left( \begin{array}{c} 0 \\ y' \\ z' \\ 1 \\ \end{array} \right) : \lambda \in \mathbb{R} \right\}
\]

in \( \mathbb{R}^4 \) that passes through the origin, where \( y', z' \) are as in (5.1). Pick the point

\[
(x - x_0) \left( \begin{array}{c} 0 \\ y' \\ z' \\ 1 \\ \end{array} \right) = \left( \begin{array}{c} 0 \\ -x_0 y \\ z_0 x \\ x - x_0 \\ \end{array} \right) \in L_{P'},
\]

which is obtained by matrix multiplication

\[
\left( \begin{array}{cccc} 0 & 0 & 0 & 0 \\ 0 & -x_0 & 0 & 0 \\ z_0 & 0 & 0 & 0 \\ 0 & 0 & x - x_0 & 0 \\ \end{array} \right) \left( \begin{array}{c} x \\ y \\ z_0 x \\ x - x_0 \\ \end{array} \right) = \left( \begin{array}{c} 0 \\ -x_0 y \\ z_0 x \\ x - x_0 \\ \end{array} \right).
\]

In other words, via the sequence

\[
P = \left( \begin{array}{c} x \\ y \\ 0 \\ \end{array} \right) \mapsto \left( \begin{array}{c} x \\ y \\ 0 \\ 1 \\ \end{array} \right) \mapsto \left( \begin{array}{c} 0 \\ -x_0 y \\ z_0 x \\ x - x_0 \\ \end{array} \right) \mapsto \left( \begin{array}{c} 0 \\ -x_0 y/(x-x_0) \\ z_0 x/(x-x_0) \\ \end{array} \right) = P',
\]

we get to perform the projection as a linear transformation (i.e. using matrix multiplication).

### 5.2. Projective space

**Definition 65.** The real \( n \)-dimensional projective space is the set

\[
\mathbb{R}P^n = \{ \text{all lines in } \mathbb{R}^{n+1} \text{ passing through the origin} \}.
\]

It will be useful to describe lines formally; a line \( L \) in \( \mathbb{R}^{n+1} \) is completely determined by a vector \( v \in \mathbb{R}^{n+1} \) for which

\[
L = \{ \lambda v : \lambda \in \mathbb{R} \}.
\]
In the terminology of linear algebra, $L$ is the 1-dimensional subspace spanned by the vector $v$.

We first examine $\mathbb{RP}^1$. By definition, the points in $\mathbb{RP}^1$ are lines in $\mathbb{R}^2$ that pass through the origin. Be ready for some mental gymnastics:

<table>
<thead>
<tr>
<th>Projective geometry</th>
<th>Euclidean geometry</th>
</tr>
</thead>
<tbody>
<tr>
<td>point</td>
<td>line</td>
</tr>
</tbody>
</table>

By analogy,

- $\mathbb{RP}^1$ is called the projective line, and the points along the projective line are all the lines in $\mathbb{R}^2$ that pass through the origin;
- $\mathbb{RP}^2$ is called the projective plane. A point in $\mathbb{RP}^2$ is a line in $\mathbb{R}^3$ that passes through the origin, and a line in $\mathbb{RP}^2$ is a plane in $\mathbb{R}^3$ that passes through the origin.

Let $L \in \mathbb{RP}^2$ be a projective point. Then $L$ is a line in $\mathbb{R}^3$ that passes through the origin. For instance, take $L_1$ to be the $x$-axis and $L_2$ to be the $y$-axis. Then $L_1$ and $L_2$ lie on a common plane $P$, namely the $x$-$y$-plane. In projective geometry, $P$ is the projective line on which the projective points $L_1$ and $L_2$ lie. In summary, we have a dimension mismatch when passing between projective and Euclidean geometry:

<table>
<thead>
<tr>
<th>Projective geometry</th>
<th>Euclidean geometry</th>
</tr>
</thead>
<tbody>
<tr>
<td>point</td>
<td>line</td>
</tr>
<tr>
<td>line</td>
<td>plane</td>
</tr>
<tr>
<td>plane</td>
<td>3-dim. linear subspace</td>
</tr>
</tbody>
</table>

The latter example is a projective manifestation of Euclid’s first postulate (which states that one can always trace a line that passes through two given distinct points). This is true more generally:

**Proposition 66.** Given any two distinct projective points in $\mathbb{RP}^n$, there exists a unique projective line that passes through them.

**Proof.** A projective point is a line passing through the origin in $\mathbb{R}^{n+1}$. Consider then two different lines in $\mathbb{R}^{n+1}$, which both pass through the origin. Then these are contained in a unique plane that passes through the origin. \(\square\)

We next show that there is no projective analogue of Euclid’s fifth postulate. Recall that Euclid’s fifth postulate allows to construct parallel lines. As we have seen, projections do not necessarily preserve parallel lines – think of the railroad tracks. In fact, in that example, the tracks converge to a single point, called vanishing point in the lexicon of perspective drawing.

**Proposition 67.** In a projective plane, any two distinct projective lines meet in a unique point.

**Proof.** Consider two planes in $\mathbb{R}^3$ that meet at the origin. Then these two planes must intersect, and their intersection is a unique line. Since both planes pass through the origin, the line of intersection must also pass through the origin. \(\square\)

### 5.3. Spherical model of projective geometry

Every set of parallel lines is equipped with an additional point (at infinity), incident to these lines. The line at infinity is incident to each point at infinity, and to only those points.
5.4. Hyperplane model of projective geometry

**Definition 68.** In a \(n\)-dimensional space, a hyperplane is a \((n-1)\)-dimensional subspace.

For instance: hyperplanes in an affine plane are lines, hyperplanes in an affine space are planes, hyperplanes in the projective plane \(\mathbb{RP}^2\) are projective lines, i.e. Euclidean planes passing through the origin, etc.

The hyperplane model for \(\mathbb{RP}^2\) consists in identifying each projective point with either a point on an affine hyperplane or a ‘point at infinity.’ (We will only consider the case of the projective plane, although the model described in this section can be applied more generally to any \(n\)-dimensional projective space.)

In the picture below, we fix the hyperplane \(\{z = 1\}\). Each line through the origin either intersects \(\{z = 1\}\) or lies in the \(x\)-\(y\)-plane. We rephrase as follows: each projective point (which we draw in purple) of \(\mathbb{RP}^2\) corresponds either to a point in \(\{z = 1\}\) or a point at infinity. Similarly, each plane through the origin either intersects \(\{z = 1\}\) along a line or is the \(x\)-\(y\)-plane. In our hyperplane model, each projective line is identified with either a line on \(\{z = 1\}\) or the ‘line at infinity’ (a.k.a. the \(x\)-\(y\)-plane).

Here is how to make sense of the notion of ‘point at infinity.’ Let \(P_1\) and \(P_2\) be two distinct planes in \(\mathbb{R}^3\) that intersect along the \(x\)-axis, none of which are the \(x\)-\(y\)-plane. In \(\mathbb{RP}^2\), we thus have two projective lines intersecting at the projective point \([1, 0, 0]\). In the hyperplane model above, these two projective lines correspond to two parallel lines on \(\{z = 1\}\) that intersect at the point at infinity given by \([1, 0, 0]\).
Indeed, recall that by Proposition 68, lines in projective geometry always intersect. For parallel lines as depicted in the figure above, this is achieved by having them meet at an abstract (or ideal) point at infinity!

**Exercise 69.** Illustrate Proposition 67 and Proposition 68 using the hyperplane model. Consider all possible cases!

It should be clear that the properties described above do not depend on the particular choice of the affine hyperplane, as long as this hyperplane does not pass through the origin. But one can as how do two different hyperplane model relate? That is, suppose that $H_1$ and $H_2$ are two different hyperplanes in $R^3$, away from the origin. Let the parallel planes $\hat{H}_1$, $\hat{H}_2$ passing through the origin denote the corresponding lines at infinity. The any projective point $x \in \mathbb{P}^2$ is identified to a point $x_1 \in H_1 \cup \hat{H}_1$ and to a point $x_2 \in H_2 \cup \hat{H}_2$. We want to understand the relation between $x_1$ and $x_2$.

**Exercise 70.** Draw all possible cases and show that the map $x_1 \mapsto x_2$ is a perspicivitity from the origin.

### 5.5. Homogeneous coordinates

Let $L \in \mathbb{P}^n$. That is $L$ is a line passing through a vector $v \in R^{n+1}$, given in Cartesian coordinates by $v = (v_0, v_1, \ldots, v_n)$.

Let’s write $L = [v]$, and define the homogeneous coordinates of $L$ to be $[v] = [v_0, v_1, \ldots, v_n]$.

**Lemma 71.** Let $v, w \in R^{n+1}$. Then $[v] = [w]$ if and only if $\exists \lambda \in R$ such that $\lambda v = w$.

**Proof.** Exercise. □

As a result, for any $\lambda \in R \setminus \{0\}$, $[v] = [\lambda v]$ or equivalently, $[v_0, \ldots, v_n] = [\lambda v_0, \ldots, \lambda v_n]$. That is, these coordinates are invariant under homogeneous scaling, and hence called homogeneous coordinates. For example,

$$[2, 4, 6] = [1, 2, 3] = [\frac{1}{3}, \frac{2}{3}, 2]$$

and this point in $\mathbb{P}^2$ corresponds to the line in $R^3$ determined by the vector $(1, 2, 3)$.

Using homogeneous coordinates, we can give an algebraic characterization of projective space:

$$\mathbb{P}^n = \{[x_0, x_1, \ldots, x_n] : x_i \in R, \text{not all } x_i's = 0\}.$$

**Question 72.** Why does this definition require that not all $x_i's = 0$?

In the hyperplane model of $\mathbb{P}^2$, we distinguish the points $[x, y, z]$ with $z \neq 0$, which we identify with points on $\{z = 1\}$ from the points at infinity $[x, y, 0]$. We can algebraically decompose the projective plane as follows:

$$\mathbb{P}^2 = \{[x, y, z] : x, y, z \in R, xyz \neq 0\}$$

$$= \{[x, y, z] : x, y, z \in R, z \neq 0\} \cup \{[x, y, 0] : x, y \in R, xy \neq 0\}$$

Since $[x, y, z] = [\frac{x}{z}, \frac{y}{z}, 1]$ if $z \neq 0$, we can further identify the first set with $\{[x, y, 1] : x, y \in R\}$. Note that the function given by $[x, y, 1] \mapsto (x, y, 1)$ defines a bijection and hence identifies points in the latter set with point in the hyperplane $\{z = 1\}$. On the other hand, the second set can be identified with the projective line $\{[x, y] : x, y \in R, xy \neq 0\}$. 

5.6. Projective transformations

Recall that we have the following dictionary:

<table>
<thead>
<tr>
<th>Projective geometry</th>
<th>Euclidean geometry</th>
</tr>
</thead>
<tbody>
<tr>
<td>point</td>
<td>line through origin</td>
</tr>
<tr>
<td>line</td>
<td>plane through the origin</td>
</tr>
<tr>
<td>$k$-dim. subspace</td>
<td>$(k + 1)$-dim. linear subspace</td>
</tr>
</tbody>
</table>

Accordingly, we expect the following correspondence between transformations of projective/Euclidean geometry

$$f : \mathbb{P}^n \rightarrow \mathbb{P}^n \quad \leftrightarrow \quad f : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^{n+1}$$

that preserves subspaces

Recall that a linear map $f : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^{n+1}$ is exactly a map that preserves linear subspaces. Note however that if $f$ is not injective, say there exists a non-zero vector $v \in \mathbb{R}^{n+1}$ such that $f(v) = 0$, then $f$ will ‘collapse’ the line $\{\lambda v : \lambda \in \mathbb{R}\}$ to the point at origin; in fact, for all $\lambda \in \mathbb{R}$, $f(\lambda v) = \lambda f(v) = 0$. We don’t want this to happen, and hence will require that $f$ be a linear injective transformation of $\mathbb{R}^{n+1}$. It is a basic theorem of linear algebra that a linear injective transformation is invertible (hint: apply the dimension theorem). This means that the standard matrix representative $A$ of $f$ is a matrix in

$$\text{GL}_{n+1}(\mathbb{R}) = \{ A \in M_{(n+1)\times(n+1)}(\mathbb{R}) : \det(A) \neq 0 \},$$

the general linear group of $(n + 1) \times (n + 1)$ matrices.

Let $A \in \text{GL}_{n+1}(\mathbb{R})$. Correspondingly, we define $f_A : \mathbb{P}^n \rightarrow \mathbb{P}^n$, $f_A([v]) = [Av]$. Observe that $A = \lambda B$ for some $\lambda \in \mathbb{R}^*$ if and only if $f_A = f_B$.

**Definition 73.** A projective transformation $f : \mathbb{P}^n \rightarrow \mathbb{P}^n$ is a map of $\mathbb{P}^n$ such that $f([v]) = [Av]$ for some $A \in \text{GL}_{n+1}(\mathbb{R})$.

Observe that it now follows from the definition that

**Proposition 74.** A projective transformation preserves projective lines and intersection points.

**Examples 75.**

1. Let $A = \left(\begin{smallmatrix} 3 & 2 \\ 1 & 0 \end{smallmatrix}\right)$. Then $f_A([x, y]) = [3x + 2y, x]$. More generally, if $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ with $ad - bc \neq 0$, then $f_A([x, y]) = [ax + by, cx + dy]$.

2. Recall the projection map $f([x, y, 0, 1]) = [0, x_0 y, -z_0 x, x - x_0]$ from earlier. The corresponding matrix in $\text{GL}_4(\mathbb{R})$ is (up to scaling)

$$A = \left(\begin{array}{cccc} 0 & 0 & 1 & 0 \\ 0 & x_0 & 0 & 0 \\ -z_0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -x_0 \end{array}\right).$$

Observe that $\det(A) = -x_0^2 z_0 \neq 0$ if and only if $x_0, z_0 \neq 0$. Recalling the setup of this railroad track examples, this means that the observer should be at a distance of the tracks and of the projection screen for the projection map to make sense as a projective transformation.
Proposition 76. The set of all projective transformations of $\mathbb{RP}^n$ forms a group, called the projective general linear group, and denoted $\text{PGL}_n(\mathbb{R})$.

Proof. We show that this is a group with respect to the composition of functions. The identity element is induced by the identity matrix. By definition, the composition of $f_A, f_B \in \text{PGL}(2, \mathbb{R})$ is given by $f_A(f_B([v])) = f_A([Bv]) = [ABv]$. From this equation, we moreover see that each projective transformation $f_A$ is invertible, with inverse $f_A^{-1} = f_A^{-1}$. □

5.7. The projective line

Let $[x, y] \in \mathbb{RP}^1$. Note that if $y \neq 0$, then $[x, y] = [\frac{x}{y}, 1]$, while otherwise $[x, 0] = [1, 0]$. Hence we have the decomposition

$$\mathbb{RP}^1 = \{(x, y) : x, y \in \mathbb{R}, x, y \text{ not both } 0\} = \{[x, 1] : x \in \mathbb{R}\} \cup [1, 0],$$

which describes the hyperplane model

Here, all lines in $\mathbb{R}^2$ (segments of which are pictured in blue) intersect the hyperplane $\{y = 1\}$ (the purple points) except for the line lying on the $x$-axis, which is the point at infinity $[1, 0]$. A projective transformation $f : \mathbb{RP}^1 \to \mathbb{RP}^1$ is of the form

$$f([x, y]) = [ax + by, cx + dy].$$

There is a one-to-one correspondence $\mathbb{RP}^1 \to \mathbb{R} \cup \{\infty\}$ given by

$$[x, y] \mapsto \begin{cases} \frac{x}{y} & \text{if } y \neq 0 \\ \infty & \text{if } y = 0. \end{cases}$$

Under this correspondence, we can express the projective transformation $f$ given above by the map $f : \mathbb{R} \cup \{\infty\} \to \mathbb{R} \cup \{\infty\}$ given by

$$f(x) = \begin{cases} \frac{ax + b}{cx + d} & \text{if } x \neq -\frac{d}{c} \\ \infty & \text{if } x = -\frac{d}{c} \end{cases}$$

for $x \in \mathbb{R}$ and

$$f(\infty) = \begin{cases} \frac{a}{c} & \text{if } c \neq 0 \\ \infty & \text{if } c = 0. \end{cases}$$

Such a function, which we will sometimes write as $f(x) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} (x)$ is called a Möbius transformation. Note that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} (x) = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} (x)$$
and
\[
\begin{bmatrix}
a & b \\
c & d
\end{bmatrix}(x) = \begin{bmatrix}
\lambda a & \lambda b \\
\lambda c & \lambda d
\end{bmatrix}(x)
\]
for any \( \lambda \in \mathbb{R}^* \). In other words, projective transformations of \( \mathbb{RP}^1 \) can be identified with Möbius transformations, and as a result \( \text{PGL}_2(\mathbb{R}) \) is identified with the group of Möbius transformations.

Using this very explicit parametrization, we prove our first result about projective transformations.

**Theorem 7.** Let \( x, y, z \in \mathbb{R} \cup \{ \infty \} \) be three distinct points. Then there exists a unique Möbius transformation \( f \) such that \( f(x) = \infty, f(y) = 0, f(z) = 1 \).

**Proof.** Since \( x, y \) and \( z \) are distinct, we may assume that \( x, y \neq \infty \). We want to determine \( a, b, c, d \) such that \( f = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \). We have the following equations
\[
\begin{bmatrix}
a & b \\
c & d
\end{bmatrix}(x) = ax + b \\
\begin{bmatrix}
a & b \\
c & d
\end{bmatrix}(y) = ay + b
\]
\[
\text{cx + d = 0} \iff \text{ay + b = 0}
\]
We need to distinguish the cases \( z = \infty \) and \( z \neq \infty \). Suppose first that \( z = \infty \), then
\[
\begin{bmatrix}
a & b \\
c & d
\end{bmatrix}(\infty) = a/d = 1 \iff a = c \neq 0.
\]
We can conclude from this system of equations that
\[
\begin{bmatrix}
1 & -y \\
1 & -x
\end{bmatrix}
\]
is the required Möbius transformation, since
\[
\begin{bmatrix}
1 & -y \\
1 & -x
\end{bmatrix}(x) = x - y \\
\begin{bmatrix}
1 & -y \\
1 & -x
\end{bmatrix}(y) = y - x
\]
\[
\text{cx + d} = 0 \iff \text{ay + b = 0}
\]
and it is indeed uniquely defined. If \( z \neq 0 \), then
\[
\begin{bmatrix}
a & b \\
c & d
\end{bmatrix}(z) = \frac{az + b}{cz + d} = 1 \iff az + b = cz + d.
\]
Hence we have the system of equations
\[
\begin{cases}
\text{cx + d} = 0 \\
\text{ay + b} = 0 \\
\text{az + b} = cz + d
\end{cases}
\]
and its solution yields the uniquely determined Möbius transformation
\[
\begin{bmatrix}
\frac{y - x}{y - z} & -x \\
\frac{1}{y - z} & -z
\end{bmatrix}.
\]
\( \square \)

**Corollary 77.** Let \( (x, y, z) \) and \( (x', y', z') \) be two triplets of pairwise distinct points in \( \mathbb{R} \cup \{ \infty \} \), then there exists a unique Möbius transformation \( f \) such that \( f(x) = x', f(y) = y', f(z) = z' \).
5.8. Cross-ratio and applications

Proof. Let \( f_1 \) be the unique Möbius transformation such that \( f_1(x) = \infty, f_1(y) = 0, f_1(z) = 1 \), and let \( f_2 \) be the unique Möbius transformation such that \( f_2(x') = \infty, f_2(y') = 0, f_2(z') = 1 \). Then \( f_2^{-1} \circ f_1 \) is uniquely determined and maps \( x \) to \( x' \), \( y \) to \( y' \), and \( z \) to \( z' \). □

Finally we note that the uniqueness part of the theorem is not true anymore if we only consider two distinct points. In fact, there exist infinitely many Möbius transformations that fix \( 0 \) and \( \infty \):

\[
\begin{bmatrix}
  n & 0 \\
  0 & 1/n
\end{bmatrix}
\]

is the cross-ratio determined by \( x, y, z \). (As a rule of thumb, if one of \( x, y, z \) or \( t \) is \( \infty \), you may use the shortcut \( \infty / \infty = 1 \); e.g.

\[
\frac{(\infty - x)(t - y)}{(\infty - y)(y - x)} = \frac{t - y}{y - x}.
\]

As we will see in the next section, the cross-ratio is ‘the’ invariant of projective transformations of lines, in the same way that distance was the invariant of orthogonal linear maps in Euclidean geometry.

5.8. Cross-ratio and applications

The cross-ratio is a geometric object, and can be defined as such. In this section, we will only consider the cross-ratio of four collinear points.

Definition 78. Let \( A, B, C, D \) be four collinear points. Let \( |AB| \) denote the length of the segment between \( A \) and \( B \). Then

\[
[A, B, C, D] = \frac{(C - A)(D - B)}{(D - A)(C - B)}
\]

is the cross-ratio of \( A, B, C, D \).

Be careful, as you can see from the definition, the order of \( A, B, C, D \) in the cross-ratio matters. The following list of properties can be checked from the definition of the cross-ratio by direct computation:

1. \([\infty, 0, 1, t] = t\)
2. \([A, B, C, D] = 0 \iff A = C \text{ or } B = D\)
3. \([A, B, C, D] = 1 \iff A = B\)
4. \([A, B, C, D] = [A, B, C, D'] \iff D = D'\)
5. ...
6. \([A, B, C, D] = [f(A), f(B), f(C), f(D)]\) for every real Möbius transformation \( f \).

Exercise 79. Check that these statements are true.

As an example, let \( A, B, C, D \) be points on a line \( L \) and \( A', B', C', D' \) be points on another line \( L' \) such that the two quadruples are in perspective.
This means that there exists a perspective projection \( f \) centered in a point \( P \) such that 
\[ f(A) = A', \ f(B) = B', \ f(C) = C', \ f(D) = D'. \]
Since \( f \) is a projection from \( L \) to \( L' \), it can be written as a Möbius transformation (cf. last section), and since the cross-ratio is invariant under Möbius transformations,
\[ [A, B, C, D] = [A', B', C', D']. \]

**First application of the cross-ratio.** While perspective projections yield for us two-dimensional representations of three-dimensional objects, the cross-ratio sometimes allows to go the other way around; obtain information on a three-dimensional object from its representation on a plane. To give a concrete example, think of a photograph representing you standing right next to a building.

You can measure the cross ratio \( c = [A', B', C', \infty] \) from the picture, and by the invariance property, \( c = [A, B, C, \infty] = \frac{C - A}{B - A} \). Hence you can recover from the photograph the height of the building \( C - A = c(B - A) \).

**Second application of the cross-ratio.**

**Theorem 8 (Desargues’ theorem).** Let \( ABC \) and \( A'B'C' \) be two triangle in perspective as pictured below. Then \( A'', B'', \) and \( C'' \) are collinear.

**Proof.** We may suppose that \( A'' \) and \( B'' \) are distinct points (otherwise we can directly draw a line connecting \( A'' = B'' \) and \( C'' \) by Proposition 67. Observe that if we can show that \( A''B'' \cap BC = A''B'' \cap B'C' \), then this would imply that \( C'' \) lies on the line \( A''B'' \). We will prove the equivalent (cf. (4) above) statement:
\[ [A'', B'', D'', A''B'' \cap BC] = [A'', B'', D'', A''B'' \cap B'C'']. \]

Observe that \( A'', B'', D'', A''B'' \cap BC \) are collinear, and that they are in perspective from \( B \) with the four collinear points \( A'', A, D, C \). Hence
\[ [A'', B'', D'', A''B'' \cap BC] = [A'', A, D, C]. \]
5.9. Main theorem of projective geometry

On the other hand, $A'', A, D, C$ are in perspective from $P$ with the four collinear points $A', A', D', C'$. Hence
\[
[A'', A, D, C] = [A'', A', D', C'].
\]
Finally, $A'', A', D', C'$ are in perspective from $B'$ with the four collinear points $A'', B'', D'', A'' B'' \cap B'C'$. We therefore have
\[
[A'', B'', D'', A'' B'' \cap BC] = [A'', B'', D'', A'' B'' \cap B'C']
= [A'', A, D, C] = [A'', A', D', C']
= [A'', B'', D'', A'' B'' \cap B'C'].
\]

5.9. Main theorem of projective geometry

We gave a very explicit proof that, given three distinct points $x, y, z \in \mathbb{R} \cup \{\infty\}$, there exists a unique Möbius transformation $f$ such that $f(x) = \infty$, $f(y) = 0$, $f(z) = 1$. This turns out to be special case of a much more general feature of projective geometry.

Observe that the three points $x, y, z$ correspond to three distinct vectors $v_1, v_2, v_3$ in $\mathbb{R}^2$ (unique only up to scaling) that are further characterized by the fact that any two of them are linearly independent. For example, if $x = \infty$, $y = 0$, $z = 1$, then we can choose as the corresponding vectors $e_1$, $e_2$ and $e_1 + e_2$, respectively.

**Exercise 80.** Explain why any two of $v_1, v_2, v_3$ in the figure above need to be linearly independent, but not all $v_1, v_2, v_3$ are.

Then (?) states that

**Theorem 9** (*‘Projective’* reformulation of ??). Let $[v_1], [v_2], [v_3]$ be three distinct points in $\mathbb{RP}^1$ such that any two of $v_1, v_2, v_3$ are linearly independent. Then there exists a unique projective transformation $f \in \text{PGL}_2(\mathbb{R})$ such that
\[
f([v_1]) = [e_1] = [1, 0], \quad f([v_2]) = [e_2] = [0, 1], \quad f([v_3]) = [e_1 + e_2] = [1, 1].
\]

You might recall that this is very similar to an important theorem in linear algebra: given two ordered bases $\beta_1 = \{v_1, v_2, \ldots, v_n\}$ and $\beta_2 = \{w_1, w_2, \ldots, w_n\}$ of an $n$-dimensional vector space, there exists a unique linear map $A \in \text{GL}_n(\mathbb{R})$ such that $A(v_i) = w_i$ for $i = 1, 2, \ldots, n$. In other words, there exists a unique change of coordinates. We digress quickly to recall why this theorem is important with the following example. Let $S$ be the reflection across the $x$-axis in $\mathbb{R}^2$. That is, $S(x, y) = (x, -y)$, and $S$ has standard matrix representative
\[
A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\]
Let $L$ be an arbitrary line in $\mathbb{R}^2$ passing through the origin and let $S_L$ be the reflection across that line. What is the matrix representative of $S_L$? The easy way to answer this question is to take the canonical basis $\{e_1, e_2\}$, and rotate it to $\beta' = \{\tilde{e}_1, \tilde{e}_2\}$ such that $\tilde{e}_1$ now lies on $L$. With respect to the basis $\beta'$, $S_L$ has matrix representative $A$, and to find the matrix representative with respect to $\beta$, we need to conjugate $A$ by the change of coordinates matrix given by the theorem.

**Theorem 10** (*Main theorem of projective geometry*). Let $[v_1], [v_2], \ldots, [v_{n+2}]$ be distinct points in $\mathbb{RP}^n$ such that any $n + 1$ vectors in $v_1, v_2, \ldots, v_{n+2} \in \mathbb{R}^{n+1}$ are linearly independent.
Exercise 3: Projective transformation, the lines. Recall that in the hyperplane model, \( [e_1], [e_2], \ldots, [e_{n+2}] \) a projective frame. Then there exists a unique projective transformation \( f \in \text{PGL}(n+1, \mathbb{R}) \) such that

\[
f([v_1]) = [e_1], \quad f([v_2]) = [e_2], \quad f([v_{n+1}]) = [e_{n+1}], \quad f([v_{n+2}]) = [e_1 + e_2 + \cdots + e_{n+1}].
\]

**Proof.** By assumption, \( v_1, v_2, \ldots, v_{n+1} \) are linearly independent. Therefore, these vectors form a basis for \( \mathbb{R}^{n+1} \). Since \( v_n \) is distinct from all, it is written as a non-trivial linear combination

\[
v_n = \sum_{i=1}^{n+1} a_i v_i \quad (a_i \in \mathbb{R}).
\]

Define \( \overline{v_i} = a_i v_i \) for \( i = 1, \ldots, n+1 \). Note that \( \overline{v_1}, \overline{v_2}, \ldots, \overline{v_{n+1}} \) is also a basis for \( \mathbb{R}^{n+1} \), and with respect to this basis, \( v_n = \overline{e_1} + \cdots + \overline{e_{n+1}} \). By the theorem of linear algebra cited above, there exists a unique \( A \in \text{GL}(n+1, \mathbb{R}) \) such that \( A\overline{e_i} = e_i \) for \( i = 1, \ldots, n+1 \).

Consider the induced projective transformation \( f([v]) = [Av] \). Then

\[
f([v_1]) = f([\overline{v_1}]) = [A\overline{e_1}] = [e_1]
\]

for \( i = 1, \ldots, n+1 \) and \( f([v_{n+2}]) = [Av_n] = [A\overline{e_1} + \cdots + A\overline{e_{n+1}}] = [e_1 + \cdots + e_{n+1}] \).  

**Desargues’ theorem revisited.** Observe that \( A'', B'', C \) all lie on a same plane. Think of this plane as the hyperplane in the hyperplane model. By Theorem 10, there exists a projective transformation \( f \) such that \( f(A'') = [e_1], \quad f(B'') = [e_2], \quad f(B) = [e_3], \quad f(C) = [e_4] \).

Recall that in the hyperplane model, \( [e_1] \) and \( [e_2] \) are points at infinity. Hence, under the projective transformation, the lines \( f(AC) \) and \( f(A'C') \) meet at infinity. This means that they run parallel to one another. The same must be true of the lines \( f(AB) \) and \( f(A'B') \).

As a consequence, \( f(BC) \) and \( f(B'C') \) must also be parallel and meet at a point at infinity. Hence \( f(A''), f(B''), f(C'') \) all lie on the line at infinity, and since projective transformations are invertible and preserve lines, we conclude that \( A'', B'', C'' \) are collinear.

This type of deduction of a general theorem from a special case is one of the beautiful and characteristic features of projective geometry. The idea is simple; a hard proof can be made easy, because projective transformations can make complicated figures simple by sending points at infinity.

5.10. **Exercises**

**Exercise 1:** What does a circle look like in perspective? Consider all possible cases.

**Exercise 2:** Let \( v, w \in \mathbb{R}^{n+1} \). Show that \( [v] = [w] \) if and only if \( \exists \lambda \in \mathbb{R} \) such that \( \lambda v = w \).

**Exercise 3:** Find the equation of the projective line that passes through the projective points \([1, 2, 3]\) and \([2, -1, 4]\).

**Exercise 4:** Determine whether the projective points \([2, 1, 3]\), \([1, 2, 1]\) and \([-1, 4, -3]\) are collinear.

**Exercise 5:** We have shown in class that in a projective plane, any two distinct projective lines meet in a projective point. Show that this point is unique.
Exercise 6: Illustrate the following fact using the hyperplane model of $\mathbb{RP}^2$: any two distinct projective lines in a projective plane meet in a unique point.

Exercise 7:

(1) Is $f([x,y,z]) = [2x, y + 3z, z]$ a projective transformation?

(2) What is the image of the line $x + 2y + 3z = 0$ under the projective transformation $f([x,y,z]) = [2x + y, -x + z, y + z]$?

Exercise 8: Consider a quadrangle $ABCD$ in $\mathbb{RP}^2$, and the associated intersection points $A' = CB \cap AD$, $B' = BD \cap AC$, $C' = AB \cap CD$ as in the drawing below.

The goal of this exercise is to show that $A', B', C'$ are not collinear, but $AB \cap A'B'$, $AC \cap A'C'$, $BC \cap B'C'$ are. Using that projective transformations preserve collinearity and intersection points, this can be done in two different ways:

(1) There exists a projective transformation $p$ sending $ABCD$ to the quadrangle with vertices $p(A) = [1, 0, 0]$, $p(B) = [0, 1, 0]$, $p(C) = [0, 0, 1]$, $p(D) = [1, 1, 1]$. Since projective transformations preserve intersection points,

$$p(A') = p(B)p(C) \cap p(A)p(D)$$

$$p(B') = p(B)p(D) \cap p(A)p(C)$$

$$p(C') = p(A)p(B) \cap p(C)p(D).$$

Projective transformations also preserve collinearity. This means that $A'$, $B'$, $C'$ are collinear if and only if $p(A')$, $p(B')$, $p(C')$ are collinear. Start by checking that $p(A') = [0, 1, 1]$, $p(B') = [1, 0, 1]$, $p(C') = [1, 1, 0]$.

(2) There exists a projective transformation $\pi$ that sends $B'$ and $C'$ to points at infinity.

Draw what the projected quadrangle $\pi(ABCD)$ looks like in the hyperplane model.

Exercise 9: Take $A \in \text{GL}_2(\mathbb{R})$ with $\det(A) = -1$. Show that the corresponding Möbius transformation $[A] \in \text{PGL}_2(\mathbb{R})$ fixes two points.

We denote these two points $x_1$, $x_2$. Show that the following statements are equivalent:

(a) $[A]^2 = \text{id}$

(b) $[x_1, x_2, x, [A](x)] = -1$

(c) $\text{tr}(A) = 0$

Exercise 10: Let $A, B, C, D$ be collinear.
(1) Show that if \( A \neq B, C \neq D, \) and \([A, B, C, D] = -1\), then \([A, B, D, C] = -1\).

Hence, when \( A \neq B, C \neq D, \) and \([A, B, C, D] = -1,\) \( C \) and \( D \) are called conjugates. In fact, they are called ‘harmonic conjugates’ for the following reason. The harmonic average \( z \) of \( x, y \) is given by the equation

\[
\frac{2}{z} = \frac{1}{x} + \frac{1}{y}.
\]

(2) Show that \((B - A)\) is the harmonic average of \((C - A)\) and \((D - A)\) if and only if \([A, B, C, D] = -1\).

(3) Let \( A \neq B, \) and let \( \infty \) be the point at infinity on the line \( AB. \) Show that the harmonic conjugate of \( \infty \) is the midpoint \( M \) of the segment \([A, B]\).

**Exercise 11:**

In the drawing above, let \( \ell_A = PA, \ell_B = PB, \ell_C = PC, \ell_D = PD. \) In the spirit of the duality between lines and points that is ever-present in projective geometry, one can define the cross-ratio of the lines \( \ell_A, \ell_B, \ell_C, \ell_D \) – concurrent in the point \( P \) – by

\[
[A, B, C, D] = \frac{\sin(\angle(\ell_A, \ell_C)) \sin(\angle(\ell_B, \ell_D))}{\sin(\angle(\ell_A, \ell_D)) \sin(\angle(\ell_B, \ell_C))}.
\]

Show that \([A, B, C, D] = [\ell_A, \ell_B, \ell_C, \ell_D].\)

**Exercise 12:** Let \( L, L' \) be two lines in \( \mathbb{R}P^2, \) and let \( f : L \to L' \) be a projective transformation. Show that \( f \) is a perspective projection if and only if \( f(L \cap L') = L \cap L'.\)

**Exercise 13:** Let \( A, B, C \) be on a line, and let \( D, E, F \) on another line. Consider the cross-intersection points \( X = AE \cap BD, Y = BF \cap CE, Z = CD \cap AF. \) Prove that \( X, Y, Z \) are collinear.

**Exercise 14:** Let

\[
\text{GL}_2^+(\mathbb{R}) = \{A \in \text{GL}_2(\mathbb{R}) : \det(A) > 0\}, \quad \text{SL}_2(\mathbb{R}) = \{A \in \text{GL}_2(\mathbb{R}) : \det(A) = 1\}.
\]

Show that \( \text{PGL}_2^+(\mathbb{R}) = \text{PSL}_2(\mathbb{R}).\)
CHAPTER 6

Hyperbolic geometry

Recall Euclid’s parallel postulate: given a line \( L \) and a point \( p \) that is not on \( L \), there exists a unique line \( L' \) passing though \( p \) that is parallel to \( L \). Unlike the other four postulates, this looks like a proposition one could actually prove. Euclid, followed by many prominent mathematicians across the centuries, did indeed try, and all failed. So let’s see what happens when the parallel postulate fails. First of all, this can happen for two reasons:

**Failure of existence.** Suppose that given \( L \) and \( p \) as above, any line through \( p \) will intersect \( L \). Seems weird? Consider the following model. Let \( S^2 \) denote the unit sphere in \( \mathbb{R}^3 \). On this sphere, we decide that lines are exactly great circles. (A great circle is an equator, such that if you slice through it, you cut the sphere in two equal parts.) All great circles intersect one another. We will not pursue spherical geometry in this course.

**Failure of uniqueness.** Suppose that given \( L \) and \( p \) as above, there are at least two different lines \( L', L'' \) passing through \( p \) that are both parallel to \( L \). A simple model in which we can observe this phenomenon is the disk model. Let \( D = \{ z \in \mathbb{C} : |z| < 1 \} \) be the unit disk in the complex plane, whose boundary is the unit circle \( S^1 = \{ z \in \mathbb{C} : |z| = 1 \} \). We define lines in \( D \) to be either linear or circular arcs in \( D \) that meet \( S^1 \) orthogonally (i.e. at a right angle).

6.1. Circle inversions

Let \( C \) be a circle of center \( z_0 \in \mathbb{C} \) and radius \( r \). We will use the notation \( C = C_{z_0}(r) \). A circle inversion is the analogue of the reflection across a line for circles. Hence we want the map

\[
i_C : C \setminus \{z_0\} \to C \setminus \{z_0\}
\]

to

- invert the inside and the outside of \( C \),
- be self-dual, i.e. \( i_C^2(z) = (i_C \circ i_C)(z) = z \),
- fix \( C \).

If \( C = S^1 \), the map that does this is

\[
i_{S^1}(z) = \frac{1}{z} = \frac{z}{|z|^2}.
\]

(In polar coordinates, \( z = re^{i\theta} \) is reflected across \( S^1 \) to \( \frac{1}{r}e^{i\theta} \) along the line passing through 0 and \( z \).) By change of coordinates, any circle \( C = C_{z_0}(r) \) can be translated and scaled to \( S^1 \) via the map \( z \mapsto \frac{z-z_0}{r} \). We derive the general formula

\[
i_C(z) = \frac{r^2}{z-z_0} + z_0
\]

for the inversion with respect to the circle \( C = C_{z_0}(r) \).
Exercise 81. Check that \(|i_C(z)||z| = r^2\) for all \(z \in \mathbb{C}\).

Let’s look at how circle inversions operate with the following examples (without proof).

1. Any line through \(z_0\) is preserved, i.e. \(i_C(L) = L\).
2. A line tangent to \(C\) is reflected to a circle inside \(C\) passing through \(z_0\) and tangent to \(C\).
3. Circles that pass through \(z_0\) are reflected to lines.
4. Circles that don’t pass through \(z_0\) are reflected to circles.
5. Orthogonal circles are preserved, i.e. if \(C'\) is orthogonal to \(C\), then \(i_C(C') = C'\).

The last property is the most important for us. It implies that hyperbolic lines are preserved under inversion with respect to \(S^1\). Can we talk about angles more generally? The answer is yes.

Consider two circular arcs \(a_1, a_2\) that intersect at a point \(x\). Since the \(a_1\) is circular, there is exactly one line that is tangent to \(a_1\) at \(x\). There is another such line for \(a_2\) at \(x\). Denote these lines \(\ell_1\) and \(\ell_2\) respectively. Then \(\ell_1\) and \(\ell_2\) meet at \(x\), and we have two supplementary angles. [FIG] The convention is the following: let \(\alpha\) be the counterclockwise angle from \(\ell_1\) to \(\ell_2\). (As a result, \(\pi - \alpha\) is the clockwise angle from \(\ell_1\) to \(\ell_2\).) We’ll use as a black box the following result:

**Theorem 11.** Circle inversions preserve the magnitude of angles between curves but reverse the orientation.

To conclude this section, notice that circle inversion maps

\[
i_C : \mathbb{C} \setminus \{z_0\} \to \mathbb{C} \setminus \{z_0\}, \quad z \mapsto \frac{r^2}{z - z_0} + z_0
\]

naturally extend to the ‘extended complex plane’ \(\mathbb{C} \cup \{\infty\}\) if one considers

\[
i_C(z_0) = \infty, \quad i_C(\infty) = z_0
\]

and \(i_C\) defined everywhere else by (6.1). We will soon see how circle inversions relate to (complex) Möbius transformations.

**Definition 82.** A generalized circle in the extended complex plane \(\mathbb{C} \cup \{\infty\}\) is either a circle in \(\mathbb{C}\) or a line together with the point at infinity.

It is perhaps helpful to visualize this definition with the help of the picture corresponding to (2) above. The moral is: circle inversions map generalized circles to generalized circles.

### 6.2. Elementary hyperbolic geometry

**Definition 83.** A geodesic (hyperbolic line) in \(\mathbb{D}\) is either a straight line passing through the origin, or a circular arc meeting \(S^1\) orthogonally.

Geodesics are fixed by \(i_{S^1}\). Let \(\ell\) be a geodesic, and denote by \(i_\ell\) the inversion with respect to the circle determined by \(\ell\). Then the restriction of \(i_\ell\) to \(\mathbb{D}\) fixes \(\ell\) and \(S^1\) and reflects all points of \(\mathbb{D}\) to points of \(\mathbb{D}\) across \(\ell\).

**Lemma 84.** Let \(z \in \mathbb{D}\) away from the origin. Then there exists a unique geodesic \(\ell\) such that \(i_\ell(z) = 0\).

**Proof.** We need to find a circle \(C = C_{z_0}(r)\) such that (1) \(C\) is orthogonal to \(S^1\), (2) \(i_C(z) = 0\). In general, how do we know whether a circle \(C\) is orthogonal to \(S^1\)?
Let $x$ be a point where $C$ and $S^1$ intersect. Observe that the intersection is orthogonal if and only if the triangle through the points $x$, 0, $z_0$ has a right angle. By Pythagoras theorem, this is equivalent to $1 + r^2 = |z_0|^2$. Hence $C_{z_0}(r)$ is orthogonal to $S^1$ if and only if $1 + r^2 = |z_0|^2$. For (2) to hold,

$$r^2 + z_0(z - z_0) = r^2 + z_0z - |z_0|^2 = 0.$$ 

Using (1), this is equivalent to

$$z_0 = \frac{1}{\bar{z}} = i_{S^1}(z)$$

and $r$ is then also uniquely determined by $r = (|z_0|^2 - 1)^{1/2}$. Set $\ell = C \cap D$.

**Proposition 85.** Let $A, B \in D$. There exists a unique geodesic passing through $A$ and $B$.

**Proof.** If $A, B, 0$ are aligned, the geodesic is the straight line passing through them. Otherwise, by Lemma 85, there exists a unique geodesic $L$ such that $i_L(A) = 0$. Let $\ell'$ be the straight line passing through $i_L(B)$ and 0. Then $i_L(\ell') = \ell$ is a generalized circle passing through $A$ and $B$ and intersecting orthogonally $S^1$, and is uniquely determined. □

Now that we can construct geodesic segments, we can look at triangles.

**Proposition 86.** The sum of the inner angles of a hyperbolic triangle is less than $\pi$.

**Proof.** Let $ABC$ be a triangle in $D$. Up to a circle inversion (Lemma 85), we may assume that $A = 0$. Look at the resulting picture: two of the triangles three edges are straight segment, while the last edge bends towards the origin of $D$. This bend implies that the sum of the angles is less than $\pi$. Since angle magnitudes and edges are preserved by reflection, this is also true of $ABC$. □

**Exercise 87.** Show that the parallel postulate holds if and only if the sum of the inner angles of a triangle is $\pi$.

Observe that the sum of the angles of triangles in spherical geometry has an excess $> \pi$. We have the complete classification

<table>
<thead>
<tr>
<th>Geometry</th>
<th>Sum of Inner Angles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Euclidean geometry</td>
<td>$\pi$</td>
</tr>
<tr>
<td>Hyperbolic geometry</td>
<td>$&lt; \pi$</td>
</tr>
<tr>
<td>Spherical geometry</td>
<td>$&gt; \pi$</td>
</tr>
</tbody>
</table>

**Definition 88.** Two geodesic segments $AB$, $A'B'$ in $D$ are called congruent if there exists a geodesic $\ell$ such that $i_\ell(AB) = A'B'$.

**Proposition 89.** Let $ABC$ be a geodesic triangle, and $\alpha = \angle(ABC)$, $\beta = \angle(ACB)$. If $\alpha = \beta$, then $AB$ and $AC$ are congruent.

**Proof.** By a simple extension of Lemma 85, there exists a geodesic $\ell$ such that $i_\ell(B) = C$. If $\alpha = \beta$, then $A$ must lie on $\ell$, and hence $i_\ell(AB) = AC$. □

To compare the above results with what we know in affine/Euclidean geometry, note that

- The role of reflections across lines is replaced by reflections across circles.
- New behavior appears wherever the parallel postulate is used.
- There is a notion of congruence and angle.

Our next goal is to introduce an explicit distance function, so that we can actually measure lengths and angles.
6.3. Cross-ratio and Möbius transformations of \( \mathbb{CP}^1 \)

**Proposition 90.** Let \( a, b, c \in \mathbb{C} \cup \{\infty\} \) be three distinct points in the extended complex plane. There exists a unique generalized circle passing through \( a, b, c \).

This is obvious. Say \( a, b, c \) are not aligned, then construct the triangle \( abc \), its midpoint \( z_0 \), point your compass at \( z_0 \) and draw the circle through \( a, b, c \). This produces a unique circle

\[
C = \{ z \in \mathbb{C} : |z - z_0| = r \}.
\]

The goal of this section is to prove

**Proposition 91.** Let \( C \) be the circle passing through three distinct points \( a, b, c \). Then \( z \in C \) if and only if \( [a, b, c, z] \in \mathbb{R} \). In other words,

\[
C = \{ z \in \mathbb{C} : [a, b, c, z] \in \mathbb{R} \}.
\]

A complex Möbius transformation is a map of the extended complex plane

\[
\mathbb{C} \cup \{\infty\} \to \mathbb{C} \cup \{\infty\}, \quad z \mapsto \frac{az + b}{cz + d}
\]

such that \( \left( \begin{array}{cc} a & b \\ c & d \end{array} \right) \in \text{GL}_2(\mathbb{C}) \) (and \( \infty \mapsto a/c, -d/c \mapsto \infty \)).

**Proposition 92.** Complex Möbius transformations preserve generalized circles.

**Proof.** A Möbius transformation can be decomposed into the composition of a finite number of inversions \( z \mapsto 1/z \), scalings \( z \mapsto \alpha \cdot z \), and translation \( z \mapsto z + \beta \). Indeed, if \( a, b, c, d \neq 0 \), we can decompose \( z \mapsto \frac{az + b}{cz + d} \) into

\[
z \mapsto cz + d \mapsto \frac{1}{cz + d} + t = \frac{ctz + dt + 1}{cz + d} \mapsto az + \frac{a}{D}(dt + 1).
\]

If \( D = ad - bc \), set \( t = a/D \). Then the decomposition above reads

\[
z \mapsto cz + d \mapsto \frac{1}{cz + d} \mapsto \frac{1}{cz + d} + \frac{a}{D} = \frac{az + b}{cz + d}.
\]

(Note that this is assuming that \( c \neq 0 \). We leave it to the reader to show that such a decomposition is also possible in the case when \( c = 0 \).) Scalings and translations of generalized circles are still generalized circles. As for the inversion \( z \mapsto z^{-1} \), observe that it can be written as the composition

\[
i_{S^1}(z) = \frac{1}{z}
\]

of a circle inversion and complex conjugation. Both also preserve generalized circles. \( \Box \)

**Lemma 93.** Let \( a, b, c \in \mathbb{C} \cup \{\infty\} \) be distinct. Then there exists a unique Möbius transformation \( f : \mathbb{C} \cup \{\infty\} \to \mathbb{C} \cup \{\infty\} \) such that \( f(a) = \infty, f(b) = 0, f(c) = 1 \). In fact, \( f \) is given by

\[
f(z) = [a, b, c, z] = \frac{(c - a)(z - b)}{(b - a)(z - c)}.
\]

**Proof.** Exercise. \( \Box \)

The formula above defines the cross-ratio of four points in the (extended) complex plane. In particular, \( [a, b, c, z] \in \mathbb{C} \). As application of Proposition 93, we give a ‘projective proof’ of Proposition 91.
**Projective proof of Proposition 91.** Let $f$ be the Möbius transformation that maps $a, b, c$ to the real line. Then $C = f^{-1}(R)$ is a generalized circle passing through $a, b, c$. Since $f$ is uniquely determined, so is $C$. □

**Proof of Proposition 92.** Let $f$ be the Möbius transformation as in the proof of Proposition 91. Since $f$ is Möbius, it will map all points on $C$ to points on the real axis $R$ (cf. Proposition 93). Since $f(a), f(b), f(c), f(z) \in R$, their cross-ratio is real, and by invariance of the cross-ratio under Möbius transformations,

$$[a, b, c, z] = [f(a), f(b), f(c), f(z)] \in R.$$  

Conversely, since $f(a) = \infty$, $f(b) = 0$, $f(c) = 1$, we compute that $f(z) = [a, b, c, z]$. Now, if $f(z)$ lies on the real line, it will get mapped back $f^{-1}(f(z)) = z$ to a point on $C$. □

### 6.4. Hyperbolic distance

Given two points $A, B \in D$, drawing the geodesic connecting them yields two more points on the boundary of the disk, i.e. $S^1$. We name $P$ the point on the side of $B$ and $Q$ the point on the side of $A$. Then $P, Q \in S^1$. Moreover, note that since $P, Q$ are the extremities of a geodesic, $P \neq Q$. In this manner, $P$ and $Q$ are completely determined by $A$ and $B$. Thus the map

$$D \times D \to R, \quad (A, B) \mapsto [A, B, P, Q]$$

is well-defined (and real-valued by Proposition 92).

**Exercise 94.** The image of the above assignment is $\geq 1$ for all $A, B \in D$. Moreover, it is precisely $1$ if and only if $A = B$.

**Definition 95.** The hyperbolic distance between two points $A, B \in D$ is given by

$$d(A, B) = \log [A, B, P, Q].$$

The distance function $d$ defines a ‘metric’ on $D$: we can check that $d(A, B) = 0$ if and only if $A = B$, $d(A, B) = d(B, A)$ and that it satisfies the triangle inequality $d(A, B) \leq d(A, C) + d(B, C)$ for all points $A, B, C \in D$. In fact, it is this last inequality that motivates the presence of the logarithm in our definition. In fact, observe that if $A, B, C$ are collinear (in the sense that they line on the same geodesic), then

$$d(A, B) + d(B, C) = \log [A, B, P, Q][B, C, P, Q]$$

$$= \log \left( \frac{(P - A)(Q - B)(P - B)(Q - C)}{(P - B)(Q - A)(Q - B)(P - C)} \right) = \log [A, C, P, Q] = d(A, C).$$

**Exercise 96.** Check that $d$ satisfies the triangle inequality.

**Definition 97.** Two hyperbolic segments are congruent if there exists a hyperbolic reflection $i_\ell$ such that $i_\ell(AB) = CD$.

**Proposition 98.** Let $AB, CD$ be two hyperbolic segments. These are congruent if $d(A, B) = d(C, D)$.

**Proof.** Suppose that these are congruent. Then $i_\ell(A) = i_\ell(C)$, $i_\ell(B) = D$, $i_\ell(P) = P'$, $i_\ell(Q) = Q'$. In terms of cross-ratio we have

$$[C, D, P', Q'] = [i_\ell(A), i_\ell(B), i_\ell(P), i_\ell(Q)].$$
The question becomes, how does the cross-ratio behave under a circle inversion? We observe that any circle inversion can be written as the composition of a Möbius transformation and complex conjugation \( c(z) = \overline{z} \);

\[
i_C(z) = \frac{r^2}{\overline{z} - z} + z_0 = \frac{z_0 \cdot \overline{z} + (r^2 - |z_0|^2)}{\overline{z} - z} = \left[ \begin{array}{cc} z_0 & r^2 - |z_0|^2 \\ 1 & -\overline{z_0} \end{array} \right] (\overline{z}) =: f(\overline{z}).
\]

(Remark that if \( C \) is orthogonal to \( S^1 \) – as is our case – \( r^2 - |z_0|^2 = -1 \). See the proof of Lemma 85 for an explanation.) Since the determinant of the matrix above is \( = -|z_0|^2 + |r|^2 \neq 0 \), \( f \) is indeed a Möbius transformation. Then

\[
[C, D, P', Q'] = [i_\ell(A), i_\ell(B), i_\ell(P), i_\ell(Q)] = [f(A), f(B), f(P), f(Q)]
\]

\[
= [A, B, P, Q] = [A, B, P', Q] = [A, B, P, Q],
\]

where the last equality holds because \([A, B, P, Q] \in \mathbb{R} \). \( \square \)

In the rest of this section, we will attempt to build up our intuition of how ‘distorting’ this metric \( d \) is by studying circles in \( D \).

**Proposition 99.** Let \( z \in D \), where \( z = re^{i\theta} \) in polar coordinates. Then \( d(0, z) = \log \left( \frac{1 + r}{1 - r} \right) \).

**Proof.** Note that \( P = e^{i\theta} \) and \( Q = e^{i\theta}e^{i\pi} = -e^{i\theta} \). Then

\[
d(0, z) = \log[0, re^{i\theta}, e^{i\theta}, -e^{i\theta}] = \log \left( \frac{1 + r}{1 - r} \right).
\]

\( \square \)

Let \( C \) be the circle centered at 0 passing through the point \( z \) as in the previous Proposition. Then

\[
C = \{z : |z| = r\} = \left\{ z : d(0, z) = \log \left( \frac{1 + r}{1 - r} \right) \right\}.
\]

Observe that as \( r \to 1 \), \( d(0, z) \to \infty \). This motivates the following definition.

**Definition 100.** The unit circle \( S^1 = \{z : |z| = 1\} \) is called the boundary at infinity of the hyperbolic disc model, which we also denote \( \partial_\infty D = S^1 \).

What happens if we consider a sequence of successive circles \( C_n = \{z \in D : d(0, z) = R_n\} \) such that \( R_{n+1} - R_n = R \)? If \( d(0, z) = |z| \) was our usual Euclidean distance function, then these successive circles would be equidistant. In the hyperbolic metric however, there must be a contraction in the distance between successive circles the closer we get to the boundary at infinity. This contraction can be explicitly computed. First observe that \( R_n = nR \). Let \( r_n \) be the Euclidean radius of the circle \( C_n \), i.e. \( r_n = |z| \) for any \( z \in C_n \). By the computation above,

\[
R_n = \log \left( \frac{1 + r_n}{1 - r_n} \right) \iff r_n = \frac{e^{nR} - 1}{e^{nR} + 1} = \tanh \left( \frac{nR}{2} \right).
\]

Then the contraction is measured by \( r_{n+1} - r_n \), and from the behavior of \( \tanh \), we know it to decay exponentially as \( n \to \infty \).

Now that we have observed the distortion effect of the hyperbolic metric, we might wonder whether a circle \( C = \{z \in D : d(z, z_0) = R\} \), defined with respect to the hyperbolic metric, does actually look like a circle? We have seen this to be the case if \( z_0 = 0 \), but perhaps away from the origin, it looks like an ellipse? Let \( z \in D \). Let \( i_\ell \) be the hyperbolic reflection such that \( i_\ell(z_0) = 0 \).
Exercise 101. Reflection across a hyperbolic line is a hyperbolic isometry, i.e. \( d(i\ell(A), i\ell(B)) = d(A, B) \) for all \( A, B \in \mathbb{D} \).

Then

\[ d(i\ell(z), 0) = d(z, z_0) = R \implies i\ell(C) = \{ z \in \mathbb{D} : d(z, 0) = R \}. \]

Since hyperbolic reflections preserve circles, \( C \) is actually a circle. However the distortion induced by the metric has the effect that its center \( z_0 \) might not coincide with the Euclidean center of \( C \).

6.5. Angles

Recall that the angle at the intersection of two circular arcs is taken to be the angles between the two tangent lines at the intersection point. Hence we measure angles in the same way that we do in Euclidean geometry. How does this relate to hyperbolic geometry?

Proposition 102. Let \( \ell_1, \ell_2 \) be two geodesics with respective extremities \( a, b \) and \( c, d \) on the boundary at infinity. Then \( \ell_1, \ell_2 \) are orthogonal if and only if \( [\ell, a, b, c, d] = -1 \).

Proof. There exists a (unique) Möbius transformation sending \( S^1 \) to the real axis in the extended complex plane \( \mathbb{C} \cup \{ \infty \} \), \( a \) to \( a' = 0 \), \( b \) to \( b' = 1 \), \( d \) to \( d' = \infty \). Then \( c' \in (0, 1) \). Since \( \ell_1, \ell_2 \) are orthogonal to \( S^1 \), their images \( \ell_1', \ell_2' \) must be orthogonal to the real axis. Hence \( \ell' \) is the half-circle passing through 0 and 1 with center 1/2, and \( \ell_2' \) is a vertical half-line going from \( c' \) to infinity.

If \( \ell_1 \) and \( \ell_2 \) are orthogonal, then the vertical line \( \ell_2' \) is orthogonal to the half-circle \( \ell_1' \). This implies that \( c' = 1/2 \). Then by direct computation,

\[ [a, b, c, d] = [a', b', c', d'] = [0, 1/2, 1/2, \infty] = -1. \]

Conversely, \( [a, b, c, d] = [0, 1, c', \infty] = -1 \) if and only if \( c' \) is the harmonic conjugate of \( \infty \). From the homework, we know that \( c' \) is then the midpoint of the segment \( [0, 1] \), i.e. \( c' = 1/2 \). It then follows that \( \ell_1' \) and \( \ell_2' \) are orthogonal, and hence \( \ell_1 \) and \( \ell_2 \) as well. \( \square \)

In the first part of the above proof, we have switched from the disc model of hyperbolic geometry to the upper half-plane model. We will discuss this model some more on Friday, but we can already highlight one of its advantages over the disc model; it is easier to carry explicit computations in the upper half-plane model. Here is an example.

If, more generally, \( \beta \) is the angles between \( \ell_1 \) and \( \ell_2 \) in the sector corresponding to the triangle \( \ell_1 \cap \ell_2, b, d \), then \( \beta \) is also the angle between the half-circle \( \ell_1' \) and the vertical half-line \( \ell_2' \) over \( c' \) in the upper half-plane model. With elementary trigonometry, we can now show that

\[ \sin(\beta - \pi 2) = \frac{c' - 1/2}{1/2} = 2c' - 1. \]

Remark that \( \sin(\beta - \pi/2) = \cos(\beta) \). Note that \( c' \) is completely determined by the cross-ratio;

\[ c' = [\infty, 0, 1, c'] = [d', a', b', c'] = [d, a, b, c]. \]

In conclusion, we have the general formula for angles

\[ [d, a, b, c] = \frac{1 + \cos(\beta)}{2}. \]
6.6. The upper half-plane model

We can write any \( z \in \mathbb{C} \) in rectangular coordinates as \( z = x + iy \), \( x, y \in \mathbb{R} \). We denote the upper half-plane \( \{ z \in \mathbb{C} : y > 0 \} \) by \( \mathbb{H} \).

**Exercise 103.** The Cayley transform

\[
\mathbb{H} \to \mathbb{D}, \quad z \mapsto \frac{z - i}{z + i}
\]

is a bijection. Hyperbolic lines in \( \mathbb{D} \) are mapped by the Cayley transform to vertical half-lines orthogonal to the \( x \)-axis and semi-circles with center on the \( x \)-axis. Check that under this one-to-one correspondence, the boundary at infinity \( \partial_\infty \mathbb{D} = S^1 \) is mapped to \( \partial_\infty \mathbb{H} = \mathbb{R} \cup \{ \infty \} \).

Since it is a Möbius transform – because \( \det \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} = 2i \neq 0 \) – it preserves the hyperbolic distance, angles and generalized circles. In particular, the two models, the upper half-plane and the disc, are equivalent; all results seen so far can be transposed directly to the upper half-plane via the Cayley transform.

**Lemma 104.** The upper half-plane is preserved by the group \( \text{PSL}_2(\mathbb{R}) \), i.e. \( \text{PSL}_2(\mathbb{R})(\mathbb{H}) = \mathbb{H} \).

**Proof.** Let \( f \in \text{PSL}_2(\mathbb{R}) \), \( z = x + iy \in \mathbb{H} \). By definition, \( f \) is of the form \( f(z) = \frac{az + b}{cz + d} \), for \( ad - bc = 1 \). Then \( f(z) \) has imaginary part

\[
\Im(f(z)) = \Im \left( \frac{az + b}{cz + d} \right) = \Im \left( \frac{(ac|z|^2 + bd + adx + bcx) + i(ay - bd)}{|cz + d|^2} \right) = \frac{y}{|cz + d|^2} > 0.
\]

This shows \( \text{PSL}_2(\mathbb{R})(\mathbb{H}) \subseteq \mathbb{H} \). Conversely, let \( z \in \mathbb{H} \), then

\[
z = x + iy = \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y^{1/2} \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ y^{-1/2} & 0 \end{bmatrix} (i).
\]

This shows that \( \mathbb{H} \subseteq \text{PSL}_2(\mathbb{R})(\mathbb{H}) \). \( \square \)

In the next section, we will prove

**Theorem 12.** \( \text{Isom}^+(\mathbb{H}) = \text{PSL}_2(\mathbb{R}) \) and any orientation-reversing isometry of \( \mathbb{H} \) can be written as the composition of an element of \( \text{PSL}_2(\mathbb{R}) \) and a reflection \( i\ell \) across a geodesic.

**Exercise 105.** Using the Cayley transform, show that

\[
\text{Isom}^+(\mathbb{D}) = \left\{ \begin{bmatrix} a & b \\ \bar{b} & \bar{a} \end{bmatrix} : a, b \in \mathbb{C}, \ |a|^2 - |b|^2 = 1 \right\}.
\]

The upper half-plane is well-suited for explicit computations. A good illustration of this is computing the length of a curve with respect to the hyperbolic metric.

**Definition 106.** Let \( c : [0, 1] \to \mathbb{H} \), \( c(t) = x(t) + iy(t) \) be the parametrization of a piecewise \( C^1 \) curve in \( \mathbb{H} \). Its hyperbolic length is defined by

\[
\ell(c) = \int_0^1 \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2} \frac{1}{y(t)} \, dt.
\]

Observe that this definition differs from the usual definition of length encountered in calculus because of the factor \( y(t) > 0 \) in the denominator. This normalization ensures that this definition of length is consistent with the hyperbolic metric. In fact, consider for example the straight path from \( ia \) to \( ib \), which we can easily parametrize by

\[
c(t) = i(bt + a(1 - t)), \quad c(0) = ia, \ c(1) = ib.
\]
Then we can check by direct computation that
\[ \ell(c) = \int_0^1 \frac{(b - a)}{bt + a(1 - it)} dt = \log(b) - \log(a) = \log \frac{b}{a} = d(ia, ib). \]

**Proposition 107.** Let \( f \in \text{Isom}^+(\mathbb{H}) \). Then \( \ell(f \circ c) = \ell(c) \).

**Proof.** Let \( c(t) = x(t) + iy(t) \) and \( f(c(t)) = u(t) + iv(t) \). Then
\[
\ell(f \circ c) = \int_0^1 \frac{|f'(c(t))\dot{c}(t)|}{v(t)} dt = \int_0^1 \frac{|f'(c(t))|\sqrt{x'(t)^2 + y'(t)^2}}{v(t)} dt.
\]
Since \( \text{Isom}^+(\mathbb{H}) = \text{PSL}_2(\mathbb{R}) \), \( f \) is of the form \( f(z) = \frac{az + b}{cz + d} \) for \( ad - bc = 1 \). Then
\[
f'(z) = \frac{a(cz + d) - c(az + d)}{(cz + d)^2} = \frac{ad - bc}{(cz + d)^2} = \frac{1}{(cz + d)^2}.
\]
Hence \( \frac{|f'(z)|}{v(z)} = \frac{1}{y} \) and \( \ell(f \circ c) = \ell(c) \).

**Proposition 108.** For any parametrized curve \( c \),
\[ \ell(c) \geq d(c(0), c(1)). \]
That is, the shortest path between two points is along the unique hyperbolic line that passes through them.

**Proof.** Let \( \ell \) be the hyperbolic line passing through \( c(0) \) and \( c(1) \). Let \( \alpha \) and \( \beta \) be the extremities of this line at the boundary at infinity \( \partial_{\infty} \mathbb{H} \) such that \( \alpha \) is on the side of \( c(0) \) and \( \beta \) is on the side of \( c(1) \). By the main theorem of projective geometry, there exists a projective transformation \( f \in \text{PGL}_2(\mathbb{C}) \) such that \( f(\alpha) = 0 \), \( f(c(0)) = ia \) and \( f(\beta) = \infty \). Since \( f \) preserves hyperbolic lines, it maps \( \ell \) to the vertical geodesic \( i\mathbb{R}_{>0} \). Moreover, as it preserves the boundary at infinity and \( \mathbb{H} \), we have that \( f \in \text{PSL}_2(\mathbb{R}) \) (cf. Lemma 105).) Let \( b \in \mathbb{R}_{>0} \) be such that \( f(c(1)) = ib \), and let \( f(c(t)) = x(t) + iy(t) \). Then
\[
\ell(c) = \ell(f \circ c) = \int_0^1 \frac{\sqrt{x'(t)^2 + y'(t)^2}}{y(t)} dt \\
\geq \int_0^1 \frac{|y'(t)|}{y(t)} dt \\
\geq \int_0^1 \frac{\dot{y}(t)}{y(t)} dt \\
= \log \left( \frac{y(1)}{y(0)} \right) = d(ia, ib) = d(f(c(0)), f(c(1))) = d(c(0), c(1)).
\]

\[ \square \]

### 6.7. Hyperbolic isometries

**Theorem 13.** \( \text{Isom}^+(\mathbb{H}) = \text{PSL}_2(\mathbb{R}) \) and any orientation-reversing isometry of \( \mathbb{H} \) can be written as the composition of an element of \( \text{PSL}_2(\mathbb{R}) \) and a reflection \( i_\ell \) across a hyperbolic line.

**Proof.** We have already seen that Möbius transforms an hyperbolic reflections preserve the hyperbolic metric. Conversely, let \( f \in \text{Isom}(\mathbb{H}) \). We start by showing that \( f \) preserves
hyperbolic lines. Indeed, if \(a, b, c \in \mathbb{H}\) are three distinct ordered points along a hyperbolic line \(\ell\), then
\[
d(f(a), f(c)) = d(a, c) = d(a, b) + d(b, c) = d(f(a), f(b)) + d(f(b), f(c)),
\]
which proves that \(f(a), f(b), f(c)\) are three distinct points along a hyperbolic line \(\ell'\), and \(f(\ell) = \ell'\).

Let \(\ell = i\mathbb{R}_{>0}\) be the vertical half-line connecting 0 and \(\infty\) on the boundary at infinity, and consider the three points \(ia, i, ib \in \ell\) such that \(0 < a < 1 < b < \infty\). Let \(\varphi \in \text{PSL}_2(\mathbb{R})\) be the unique map such that \(\varphi(f(ia)) = ia, \varphi(f(i)) = i, \varphi(f(ib)) = ib\). The map \(F := \varphi \circ f\) is a composition of hyperbolic isometries, hence an isometry itself. Moreover \(F(\ell) = \ell\) and \(F\) fixes the three points \(ia, i, ib\). Since \(F\) is an isometry, this implies that \(F\) not only preserve the line \(\ell\) but fixes each point along that line. There are only two possibilities; either \(F\) is the identity map or it is the hyperbolic reflection across \(\ell\). As a result, \(f\) is either of the form \(\varphi^{-1}\) or \(\varphi^{-1} \circ i\ell\) for some \(\varphi \in \text{PSL}_2(\mathbb{R})\).

We conclude this chapter with the classification of orientation-preserving hyperbolic motions. On the next page, we illustrate these in the upper half-plane.

**Theorem 14.** Hyperbolic isometries can be classified as follows. Let \(f \in \text{PSL}_2(\mathbb{R})\). Then
\[
f(z) = \frac{az + b}{cz + d}
\]
for \(a, b, c, d \in \mathbb{R}\) with \(ad - bc = 1\) and is called...

<table>
<thead>
<tr>
<th>type</th>
<th>alg. criterion</th>
<th>geom. criterion</th>
<th>canonical form</th>
</tr>
</thead>
<tbody>
<tr>
<td>...elliptic</td>
<td>(</td>
<td>a + d</td>
<td>&lt; 2)</td>
</tr>
<tr>
<td>...parabolic</td>
<td>(</td>
<td>a + d</td>
<td>= 2)</td>
</tr>
<tr>
<td>...hyperbolic</td>
<td>(</td>
<td>a + d</td>
<td>&gt; 2)</td>
</tr>
</tbody>
</table>

**Proof.** We will only consider hyperbolic case. The two other cases are dealt with similarly. Let \(f \in \text{Isom}^+(\mathbb{H}) = \text{PSL}_2(\mathbb{R})\). Consider the quadratic equation
\[
(*) \quad f(z) = \frac{az + b}{cz + d} = z \iff cz^2 + (d - a)z - b = 0.
\]
Its discriminant is \(\Delta = (d - a)^2 + 4bc = (a + d)^2 - 4\). The following are equivalent:
\[|a + d| > 2 \iff \Delta > 0 \iff (*)\]
has two real solutions.

Let \(x_\pm\) be these solutions. There exists a unique geodesic in \(\mathbb{H}\) with extremities \(x_-, x_+\) on the boundary at infinity. Let \(z\) be a point on that geodesic. Then there exists a unique \(\varphi \in \text{PSL}_2(\mathbb{R})\) such that \(\varphi(x_-) = 0, \varphi(z) = i, \varphi(x_+) = \infty\). Consider the composition \(F = \varphi \circ f \circ \varphi^{-1} \in \text{PSL}_2(\mathbb{R})\). Then \(F(0) = 0, F(\infty) = \infty\) and \(F(i) =: it \in i\mathbb{R}_{>0}\). On the other hand, the Möbius transform \(a_t\) defined in the box above satisfies \(a_t(0) = 0, a_t(\infty) = \infty\) and \(a_t(i) = it\). The uniqueness part of the main theorem of projective geometry implies \(F = a_t\).

In other terms, every hyperbolic \(f\) is conjugates to a dilation of the form \(a_t\). Similarly, every parabolic \(f\) is conjugated to a translation of the form \(r_x\), and every elliptic \(f\) is conjugated to a rotation of the form \(k_\theta\).
Exercises

Exercise 1: Show that the parallel postulate holds if and only if the sum of inner angles of a triangle is \( \pi \).

Exercise 2: Let \( A, B \) be distinct points in \( D \). Show that there exists a unique geodesic (hyperbolic line) \( \ell \) orthogonal to the (hyperbolic) segment \( AB \) with respect to which \( A \) and \( B \) are reflected, i.e. such that \( i_\ell(A) = B \).

Exercise 3: Let \( \ell, \ell' \) be two parallel geodesics. Show how to construct a geodesic \( \ell'' \) that is perpendicular to \( \ell \) and \( \ell' \). Is this construction always possible? Is it unique?

Exercise 4: Let \( d \) be the hyperbolic distance function. The set
\[
C = \{ z \in \mathbb{C} : d(z, 1/3) = 1 \}.
\]
describes an Euclidean circle. Find its center and radius.
Exercise 5: Let $A \in \mathbb{D}$. Let $\alpha$ be an angle. Prove that there exists a hyperbolic isometry that fixes $A$ and rotates all points in $\mathbb{D}$ by $\alpha$. 