Math 130 - Classical Geometry

Taught by Man-Wai Cheung
Notes by Dongryul Kim

Spring 2018

This course was taught by Man-Wai (Mandy) Cheung during the spring semester of 2018. The text book was Stillwell's *The Four Pillars of Geometry*. Grading was based on weekly assignments, a midterm paper, and a final paper, and I was the course assistant.

Contents

1 January 22, 2018

This is a class on classical geometry. We are going to start with Euclid's axiom, talk about coordinates and projective geometry, and move to non-Euclidean geometry. We might also talk about finite geometry, geometry over finite fields. There will be weekly homework assignments due on Wednesdays.

1.1 Euclid's construction axioms

Today we will start discussing Euclid's geometry. Euclid was the first person to study geometry from a rigorous standpoint.

> "The laws of nature are but the mathematical thoughts of God."
>
> —Euclid

The course will emphasize proofs. We will start out with axioms, that we accept as true, and from those axiom deduce theorems. Here are Euclid's construction axioms:

(E1) Given two points A and B, we can draw a straight line AB between the two points.

(E2) We can extend a line infinitely long.

(E3) Given two points A and a length r, we can draw a circle with center A and radius r.

Already from these constructions, we can add two lengths together. Consider two lengths $|AB|$ and $|CD|$. Using (E2), we can extend AB infinitely long. Then using (E3), draw the circle with center B and radius $|CD|$. Similarly, we can construct differences between lengths by taking the other intersection.

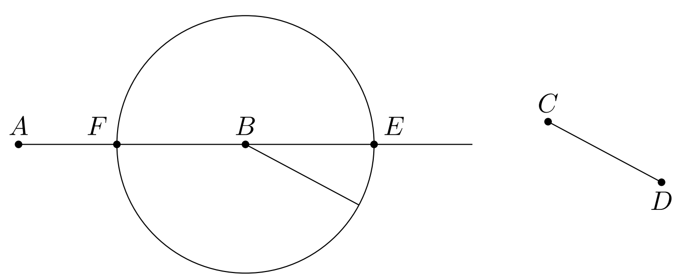

Figure 1: Constructing the sum and difference of lengths $|AB|$ and $|CD|$: we have $|AE| = |AB| + |CD|$ and $|AF| = |AB| - |CD|$.

Why did the Greeks care about these? They need to do tasks such as measure the area of land, so that they could sell or buy at a reasonable price. Or they wanted to build architecture that does not collapse, or they cared about measuring distance between stars in the sky.

Next, let us try to draw an **equilateral triangle**, which is a triangle with all th sides of the same length. Given two points A and B, we can draw a circle centered at A with radius $|AB|$ and also a circle centered at B with radius $|AB|$. If we let C be one intersection of the two circles, then ABC will be an equilateral triangle. Why is this? Because C is on the circle centered at B with radius $|AB|$, we get $|AB| = |BC|$. Likewise, $|AB| = |AC|$, and so $\triangle ABC$ is equilateral.

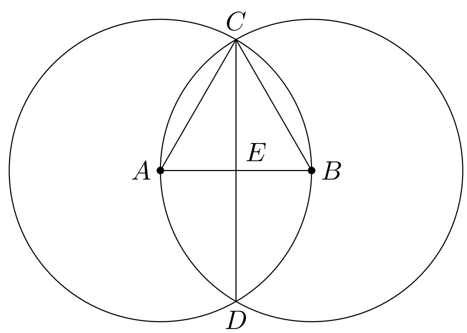

Figure 2: Constructing an equilateral triangle ABC

Given a segment AB, we can also draw a line that is perpendicular to AB and bisects AB; just draw the line connecting the two intersections of the circles. But how do we know that this line actually bisects AB and is perpendicular to AB? To prove this, we need other axioms. I'm not going to give all the axioms today, but that fact follows from SSS and SAS. They roughly say that if two triangles share three side lengths, or two side lengths and the same angle between them, they are congruent. In the above situation, we first show that $\triangle ACD \cong \triangle BCD$ using SSS. From this we see that $\angle ACE = \angle BCE$, and so by SAS we get $\triangle ACE \cong \triangle BCE$. This shows that $AE = BE$ and $\angle AEC = \angle BEC = 90°$.

2 January 24, 2018

The homework is due next Wednesday. Last time, we talked about Euclid's construction axioms. The first two were about straight lines; we can connect two points, and we can extend a line infinitely long. The third was to draw a circle with given center and radius. We used this to add and subtract lengths. We also learned to draw an equilateral triangle of given length using two circles, and used it to bisect segments by a perpendicular line.

2.1 More constructions

Let us now bisect angles. Given an angle, how do we bisect it? First draw a circle of arbitrary radius, and then this intersects the two lines at two points. Then we get a segment, and the perpendicular bisector will bisect the angle.

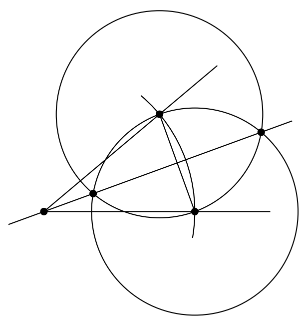

Figure 3: Bisecting an angle

Can we draw line parallel to a given line? We can draw a perpendicular line, and then draw another perpendicular line. For now, we are not going to prove things rigorously, because we don't know how. But here, we're using that $\frac{\pi}{2} + \frac{\pi}{2} = \pi$.

Theorem 2.1 (Thales)**.** *Parallel lines cut any line they cross into segments of proportional length. More precisely,*

$$\frac{|AP|}{|PB|} = \frac{|AQ|}{|QC|}$$

in Figure 4 if PQ and BC are parallel.

We're not going to prove it, but we can use it to do more constructions. Given a segment AB, how can we divide it into n equal parts? We know how to bisect it, and bisect each of them. This gives us a way to divide AB into 2^k pieces, but what if n is odd? We can draw an arbitrary line ℓ passing A, take an arbitrary length, and draw these on the line ℓ so that $|AP_1| = |P_1P_2| = \cdots = |P_{n-1}P_n| = \cdots$. Then we can draw lines through P_i parallel to P_nB, which will divide AB into n parts.

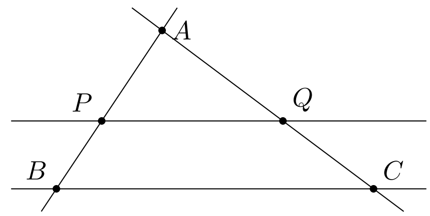

Figure 4: Thales's theorem

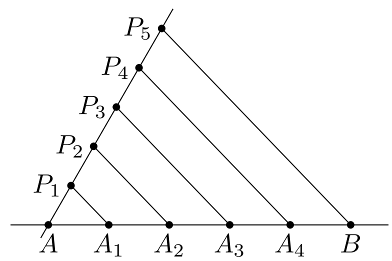

Figure 5: Dividing a segment into n equal parts

2.2 Multiplying and dividing lengths

We would now like to multiply and divide lengths. But the problem is that multiplication of two lengths is not well-defined as a length. For instance, what lengths does 2m times 3m correspond to? So make sense out of it, we need to take an arbitrary length and fix it as 1. Then every length corresponds to a number. Then we can make sense of 2 times 3 as 6, where these are all lengths.

So how should we multiply or divide lengths, where we are given this length 1? Let's first thing about dividing $|CD|$ by $|AB|$. The idea is if one of the lengths in Thales's theorem is 1, then we get an equation that says that some length is the ratio of two lengths. Let us first put AB and CD together so that $B = C$. Draw an arbitrary line through A, and find a point P such that $|AP| = 1$. Draw line parallel to BP passing through D, and let Q be the point where this line meets AP. Then by Thales's theorem,

$$|PQ| = \frac{|PQ|}{|AP|} = \frac{|BD|}{|AB|} = \frac{|CD|}{|AB|}.$$

A similar idea allows us to multiply lengths. How can we construct $|AB| \times |CD|$? Take an arbitrary line through A, pick a point C on this line with $|AC| = 1$, add $|CD|$ to this length, and then draw a parallel line to BC through D. If we denote by E the intersection point, then

$$\frac{|AC|}{|AB|} = \frac{|CD|}{|BE|}$$

shows that $|BE| = |AB||CD|$.

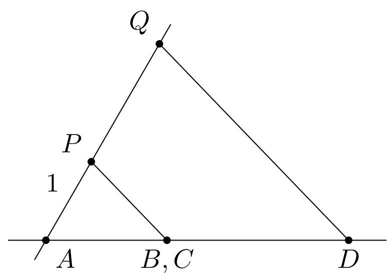

Figure 6: Dividing $|CD|$ by $|AB|$

So far we have learned how to add, subtract, multiply, and divide lengths. Next time we will construct some regular polygons.

3 January 26, 2018

We are starting form Euclid's construction axioms. To repeat again, (1) we can draw a line between two points, (2) we can extend any line segment infinitely in both directions, (3) we can draw a circle with given center and radius. We used these to do addition, subtraction, multiplication, and division of lengths. We also talked about Thales's theorem, which says that parallels cut any lines the cross in proportional line segments.

3.1 Similar triangles

Definition 3.1. We are going to say that two triangles $\triangle ABC$ and $\triangle A'B'C'$ are **similar** if the corresponding angles are the same (i.e., $\angle BAC = \angle B'A'C'$ and $\angle ABC = \angle A'B'C'$ and $\angle ACB = \angle A'C'B'$.) In this case, we write $\triangle ABC \sim \triangle A'B'C'$.

This is also called the AAA condition.

Theorem 3.2 (assuming Thales's theorem)**.** *Similar triangles have proportional side lengths.*

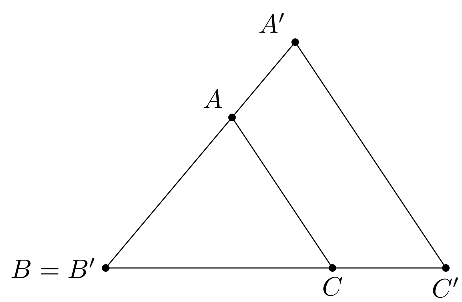

Figure 7: Similar triangles have proportional side lengths

Proof. Let us put $\triangle ABC$ and $\triangle A'B'C'$ together so that $B = B'$ and A, C lies on $B'A', B'C'$. We haven't discussed this yet, but $\angle ACB = \angle A'C'B'$ implies that AC and $A'C'$ are parallel. Then Thales's theorem implies

$$\frac{|AB|}{|A'B'| - |AB|} = \frac{|BC|}{|B'C'| - |BC|}.$$

After algebraic manipulations, we see that $|AB||B'C'| = |BC||A'B'|$, and so

$$\frac{|B'C'|}{|BC|} = \frac{|A'B'|}{|AB|}. \qquad \square$$

There are several things we're assuming for now, but we are going to prove all of them later. For now, just assume that these are true and focus on the constructions.

3.2 Construction of an irrational number

Using addition, subtraction, multiplication, and division, we can construction all the numbers of the form

$$\frac{m}{n}$$

where m and n are integers. Such numbers are called **rational numbers**. There are numbers that are not expressible in such a way, called **irrational numbers**. The number $\sqrt{2}$ is such a number. This number can be constructed in the following way. First draw a square with side length 1. Then let d be the length of one diagonal: $d = |AC|$.

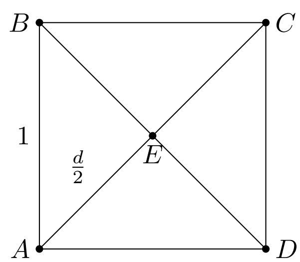

Figure 8: Constructing $\sqrt{2}$

Now my claim is that $d = \sqrt{2}$. To show this, we first note that $\triangle AMD \sim \triangle ADC$. This is because they share an angle at A, and also $\angle AMD = \angle ACD = 90°$. Because similar triangles have proportional sides, we have

$$\frac{d}{1} = \frac{|AC|}{|AD|} = \frac{|AD|}{|AM|} = \frac{1}{\frac{d}{2}},$$

and so $d^2 = 2$. That is, $d = \sqrt{2}$.

Theorem 3.3. *The number $\sqrt{2}$ is not rational, i.e., irrational.*

Proof. We first going to assume that $\sqrt{2}$ is rational and then get a contradiction. This is going to show that $\sqrt{2}$ is not rational. Assume that

$$\sqrt{2} = \frac{m}{n}.$$

Here, we can assume that m and n do not have a common divisor (other than ± 1.) If we square both sides, we get

$$m^2 = 2n^2.$$

This shows that m^2 is even, and so m should be even as well. (Otherwise m would be odd and m^2 would be odd too.) Write $m = 2\ell$. Then $(2\ell)^2 = 2n^2$ and so

$$2\ell^2 = n^2.$$

This again, shows that n is even. So both m and n are even, but we have assumed that m and n do not have a common divisor. Therefore we get a contradiction, so it cannot be the case that $\sqrt{2}$ is rational. That is, $\sqrt{2}$ is irrational. $\square$

3.3 Euclid's axioms continued

Here are the other Euclid's axioms, which will allow us to actually prove stuff.

(E4) All right angles are the same.

(E5) (Parallel postulate) If two lines intersect a third line, in such a way that the sum of the angles on one side is less then π, then the two lines meet on that side when extended.

The parallel postulate looks complicated, but it's clear when drawn out. If basically says that, in Figure 9, if $\alpha + \beta < \pi$ then the two lines will meet on the right side.

Figure 9: The parallel postulate

Now we can prove that the sum of three angles of a triangle is π.

Theorem 3.4. *Let $\triangle ABC$ be a triangle. Then $\angle BAC + \angle ABC + \angle BCA = \pi$.*

Proof. Let us write $\alpha = \angle BAC$ and $\beta = \angle ABC$ and $\gamma = \angle BCA$. Let ℓ be a line that passes C and is parallel to AB. Because ℓ is parallel to AB, we have $\delta + (\pi - \beta) = \pi$ in Figure 10. So $\delta = \beta$, and likewise we can move α to B so that $\alpha + \beta + \gamma$ is the angle on the line ℓ. $\square$

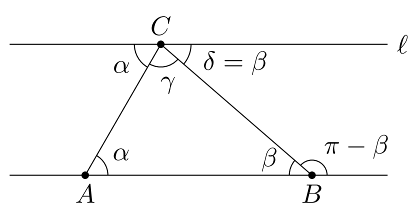

Figure 10: Sum of angles in a triangle is π

Why isn't this an obvious fact? If you look at a triangle on a sphere, where lines are great circles, you will find out that the sum of the angles of a triangle is greater than π. The angle between two lines is a local notion, but

Actually, the fact that the sum of angles in a triangle is equal to π, along with the first 4 axioms and Thales's theorem, implies the parallel axiom. To see this, let m, n be two lines that intersect k at A, B and assume that the angle at A and at B add up to less than π. Draw a line ℓ that is parallel to n and passes through A. For each $P \in \ell$, we can draw a triangle APQ such that $\angle APQ = \frac{\pi}{2}$ and $Q \in m$. All these triangles are similar, and so as AP gets big, the length PQ will get proportionally big. So at some point, it will exceed the distance between ℓ and n. At that point, it will be clear that the line m has already intersected n.

4 January 29, 2018

Today we are going to talk about Euclid's axioms. The first three were construction axioms: a straight line can be draw between two points, a line segment can be extended, there is a circle through centered at any point with given radius. The other two are: all right angles are equal, if you have two lines drawn which intersects a third so that the sum of the angles on one side is less than π, then the two lines meet on that side. There are geometric models in which the first four axioms hold but the fifth one doesn't. On a sphere, the natural notion of a straight line is a great circle. (If you take a flight from one place to another, you fly in this route.) But the fifth parallel postulate doesn't hold. There is a triangle with all angles $\frac{\pi}{2}$, and this can be interpreted as $\alpha + \beta = \pi$ but meeting at that side.

There are other axioms that Euclid didn't state:

- SAS: If two triangles have the same side-angle-side, they are congruent.
- Axiom of equality: If $A = B$ and $B = C$ then $A = C$.

These axioms Euclid just assumes because they seemed too obvious.

4.1 Equivalent form of the parallel postulate

There is another form of the parallel postulate.

(E5') Given a line ℓ and a point P not on ℓ, there is a unique parallel line through ℓ (i.e., doesn't meet ℓ).

You can prove from this that the sum of the angles of a triangle is π. Actually, this fact is equivalent to the parallel postulate.

Proposition 4.1. *An equivalent way to sate that parallel postulate is that the sum of the angles of a triangle is π.*

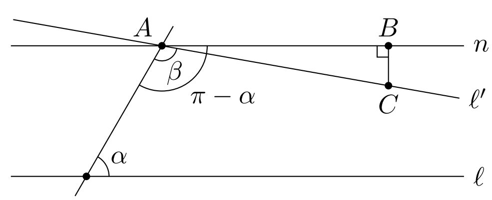

Figure 11: Equivalent formulation of the parallel axiom

Proof. We want to show that if $\alpha + \beta < \pi$, then ℓ and ℓ' meet on the right side. Draw another line n that passes A and has angle $\pi - \alpha$, and let B be a point on n, on the right side of A. Let C be the point on ℓ' with $\angle ABC = \frac{\pi}{2}$. By Thales's theorem, as the length of AB increases, the length of BC increases as well. So BC eventually has length greater than the distance between n and ℓ. □

Here, we've assumed that the distance of a segment perpendicular to two parallel lines always has the same length. I think we will prove this fact later. Also, note that the axiom that "sum of the angles of a triangle is π" is false if we're on a sphere. There is a triangle with angles $\frac{\pi}{2}$, and then the sum is $\frac{\pi}{2} + \frac{\pi}{2} + \frac{\pi}{2} = \frac{3\pi}{2}$. In fact, on the sphere, the sum of the angles is always greater than π.

4.2 Congruence axioms

Definition 4.2 (Euclid)**.** Two geometric figures are **congruent** if on can be moved to fit exactly on the other.

A more formal way of saying this would be the following.

Definition 4.3. Two triangles are **congruent** if their corresponding angles and side lengths are equal.

This means that there is a correspondence between the vertices of the two triangle, coming from the "fitting on the other", such that the corresponding angles and side lengths are equal. The SAS axiom states that we don't need to check all.

(SAS) When checking congruence, we only need to check two adjacent sides and the angel between the two.

The other axioms (ASA) and (SSS) follow from (SAS).

Definition 4.4. An **isosceles triangle** is a triangle with two equal sides.

Proposition 4.5. *The opposite angles to the two equal sides of an isosceles triangles are the same.*

Proof. If $AC = BC$, then we have a congruence $\triangle ACB \equiv \triangle BCA$ by (SAS). So $\angle CAB = \angle CBA$. □

We have constructed the perpendicular bisector before, but didn't really prove that it is a perpendicular bisector. We can now prove this.

Proposition 4.6. *The line CD is a perpendicular bisector for AB.*

Proof. We have $\angle ACD = \angle ADC$ because $AC = AD$. Similarly, we have $\angle ACB = \angle ADB$. We'll continue next time. □

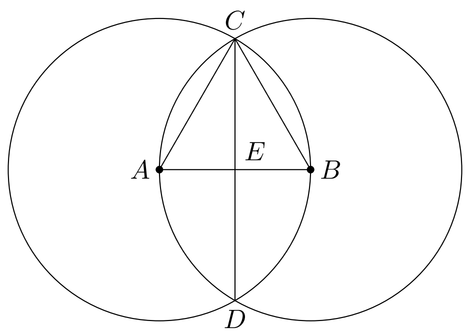

Figure 12: Constructing the perpendicular bisector of AB

5 January 31, 2018

Today I want to talk about area.

5.1 Area

First consider the algebraic fact

$$(a+b)^2 = a^2 + 2ab + b^2.$$

This has a geometric interpretation. The square of a number is the are of the square with that side-length. So the area of a square with side $a+b$ can be divided into a square with side a, a square with side b, and two rectangles of side a and b.

Using this, we can figure out the area of a parallelogram or a triangle. By definition, the area of a rectangle with side a and b has area ab. But what about a parallelogram?

Definition 5.1. A **parallelogram** is the region between two sets of two parallel lines.

Proposition 5.2. *The area of a parallelogram is its base times height.*

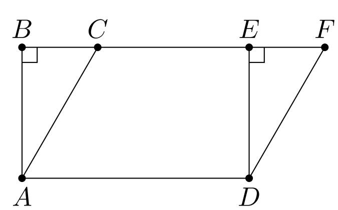

Figure 13: Area of a parallelogram

Proof. The idea is that we can cut up and rearrange it into a rectangle. We have

$$\text{Area}(ACFD) = \text{Area}(ABFD) + \text{Area}(ABC).$$

But by ASA, we have a congruence $\triangle ABC \cong \triangle DEF$, so $\text{Area}(ABC) = \text{Area}(DEF)$. So

$$\text{Area}(ACFD) = \text{Area}(ABFD) + \text{Area}(DEF) = \text{Area}(ABDE).$$

Now this is a rectangle, so its area is the product of the two sides. □

Proposition 5.3. *The area of a triangle is $\frac{1}{2}$ times base times height.*

Proof. Given any triangle ABC, we can find a point D that makes $ACDB$ a parallelogram. Then $\angle CBD = \angle BCA$ and $\angle CBA = \angle BCD$ and so $\triangle ABC \cong \triangle DCB$ by ASA. This shows that $\text{Area}(ABC) = \text{Area}(BCD$ and so

$$\text{Area}(ABC) = \frac{1}{2}\,\text{Area}(ABDC).$$

Then it follows from the previous proposition. □

5.2 The Pythagorean theorem

From the discussion on area, we can deduce the Pythagorean theorem.

Theorem 5.4 (Pythagorean theorem)**.** *The sum of the square of the two shorter sides of a right triangle is the square of the longer side.*

There are many different proofs, but here is one proof not in the book.

Proof. Let the lengths of the sides be $a, b < c$. We want to show that $a^2+b^2 = c^2$, and we draw triangles as in Figure 14.

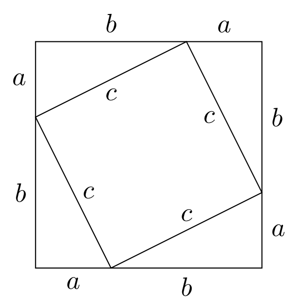

Figure 14: The Pythagorean theorem

We can calculate the area of the big square in two different ways. First, it is just $(a + b)^2$, because it has side length $a + b$. On the other hand, it is the sum of the areas of all the small regions, which is $c^2 + 4 \cdot \frac{1}{2}ab$. So

$$(a + b)^2 = 4\Big(\frac{1}{2}ab\Big) + c^2 = 2ab + c^2,$$

and it follows that $a^2 + b^2 = c^2$. □

5.3 Angles in the circle

Last class, we showed that the angles opposite equal edges in an isosceles triangle are equal. This fact seems obvious, but it has a surprising consequence. This is the power of the formalism we are working with.

Proposition 5.5. *Let A and B be points on a circle. For all points C on the circle on the same side of AB, the angle $\angle ACB$ is the same.*

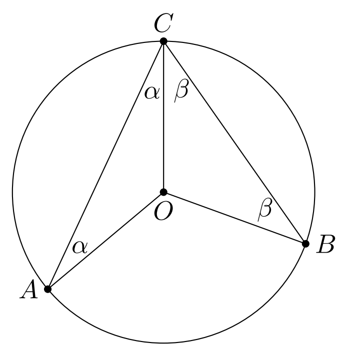

Figure 15: Angles in a circle

Proof. Let O be the center of the circle. By the isosceles triangle proposition, we have $\angle OCA = \angle OAC = \alpha$. Since the sum of the angles of $\triangle OAC$ is π, it follows that $\angle AOC = \pi - 2\alpha$. Similarly, if we write $\angle OCB = \angle OBC = \beta$, we conclude $\angle BOC = \pi - 2\beta$. These two facts imply that

$$\angle AOB = 2\pi - (\pi - 2\alpha) - (\pi - 2\beta) = 2\alpha + 2\beta = 2(\alpha + \beta) = 2\angle ACB.$$

That implies that $\angle ACB = \frac{1}{2}\angle AOB$ is independent of C. □

There is one special case I would like to mention. What happens if AB is the diameter? In this case, $\angle AOB = \pi$ and so $\angle ACB = \frac{\pi}{2}$ for every C.

This fact can be used to construct square roots. Given a length h, how can we construct a segment of length $\sqrt{h}$? This is a bit tricky. First extend the segment of length h and put a length 1 next to it. Then draw a circle with diameter $1 + h$.

If you use the Pythagorean theorem, you get three equations

$$h^2 + b^2 = a^2, \quad 1 + b^2 = c^2, \quad a^2 + c^2 = (1 + h)^2.$$

You can then do the algebra and get $b = \sqrt{h}$. But there is a more geometric explanation. If we look at the angles, you see from $\angle ABC + \angle ACB = \frac{\pi}{2}$ and

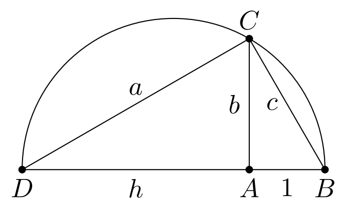

Figure 16: Constructing the square root

$\angle ACB + \angle ACD = \frac{\pi}{2}$ that $\angle ABC = \angle ACD$. Also $\angle BAC = \angle CAD = \frac{\pi}{2}$. From this see that the two triangles are similar:

$$\triangle ABC \sim \triangle ACD.$$

Then

$$\frac{1}{AC} = \frac{AB}{AC} = \frac{AC}{AD} = \frac{AC}{h},$$

and it follows that $AC = \sqrt{h}$.

Now we know now to construct $+, -, \times, \div, \sqrt{}$ using straightedge and compass constructions. But there are some things Euclid missed to make an axiom. He assumed that certain circles intersected when it is not provable that they would intersect. Hilbert around 1900 wrote down 15 axioms to complete Euclid's axioms to be totally rigorous. Here is one example of Hilbert's axioms:

(L3) There exist three points that don't lie on a line.

You might have some construction that is based on a point that lies on a certain line, and it is not provable that this construction is possible only with Euclid's axioms.

6 February 2, 2018

We have talked about the parallel axiom and the congruence axioms.

Definition 6.1. Two triangles are **congruent** if their corresponding angles and side-length are equal.

The congruence axioms then say that two triangles are congruent if two sides and the angle between them are equal. From this, we obtained a formula for the area of a parallelogram, and then a formula for the area of a triangle. We also showed that two angles of an isosceles triangle are equal, and from this showed that an angle $\angle ACB$ in a circle is determined only by A and B (and the side on which C lies).

6.1 Alternative proof of the Pythagorean theorem

Let me first give another proof of the Pythagorean theorem. For a right triangle $\triangle ABC$, we want to show that $a^2 + b^2 = c^2$.

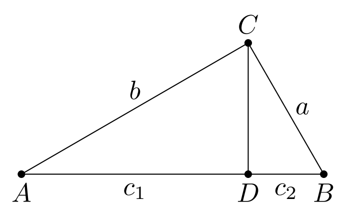

Figure 17: Another proof of the Pythagorean theorem

In Figure 17, we consider the two triangles $\triangle ADC$ and $\triangle ACB$. The two triangles are similar because $\angle ADC = \angle ACB = \frac{\pi}{2}$ and $\angle A$ are in common. So they have proportional lengths. In particular,

$$\frac{b}{c} = \frac{|AC|}{|AB|} = \frac{|AD|}{|AC|} = \frac{c_1}{b}$$

and so

$$b^2 = cc_1.$$

Likewise, if we look at $\triangle BDC$ and $\triangle BCA$, they are similar. So by the same reasoning, we get

$$a^2 = cc_2.$$

If we add the two equations, we get

$$a^2 + b^2 = c(c_1 + c_2) = c^2.$$

6.2 Proof of Thales's theorem

Let us now prove Thales's theorem.

Theorem 6.2 (Thales)**.** *A line parallel to one side of a triangle cuts the other two sides proportionally.*

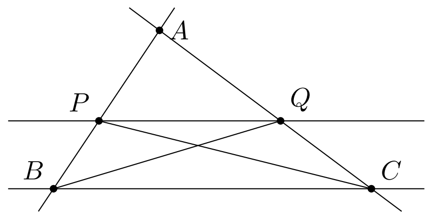

Figure 18: Thales's theorem

Proof. Consider a triangle $\triangle ABC$ and assume that PQ is parallel to BC. Because $\triangle PQB$ and $\triangle PQC$ share the same base and have the same height, the two triangles have the same area. If we compare the areas of $\triangle APQ$ and $\triangle PQB$, we can say that they have the same height with base AP and PB. So

$$\frac{|AP|}{|PB|} = \frac{\text{Area}(\triangle APQ)}{\text{Area}(\triangle PQB)}.$$

Likewise, if we compare the areas of $\triangle APQ$ and $\triangle PQC$, we get

$$\frac{|AQ|}{|QC|} = \frac{\text{Area}(\triangle APQ)}{\text{Area}(\triangle PQC)}.$$

Because the areas of $\triangle PQB$ is equal to $\triangle PQC$ are the same,

$$\frac{|AP|}{|PB|} = \frac{\text{Area}(\triangle APQ)}{\text{Area}(\triangle PQB)} = \frac{\text{Area}(\triangle APQ)}{\text{Area}(\triangle PQC)} = \frac{|AQ|}{|QC|}. \qquad \square$$

6.3 The Dedekind axiom

Now we jump about 2000 years to early 20th century. Euclid missed many of the axioms, and Hilbert attempted to give a complete rigorous set of axioms for doing Euclidean geometry. For instance, it is not trivial that multiplication commutes, i.e., $ab = ba$. This fails for matrices for instance. There is another example. If $AB \cong CD$ and $CD \cong EF$, is it always true that $AB \cong EF$? This is again not trivial, because your friend's friend is not always your friend. This should be taken as an axiom, that equality is a notion of equivalence.

Here is one important axiom. The idea is that the line is complete, that is, has no gap.

Dedekind axiom. *Let $\mathcal{L} = A \cup B$ be a line partitioned into a disjoint set, so that $A \cap B = \emptyset$. Suppose there does not exist a point in A that lies between two points in B, and vice versa. If A and B are both nonempty, there exists a point $p \in \mathcal{L}$ such that one side of p is entirely in A and the other side of p is entirely in B.*

This axiom is important because this allows us to take intersections figures. Otherwise, we can't be sure that $\sqrt{2}$ is an actual length that can be put on an arbitrary line.

7 February 5, 2018

We jump 2000 years and now we are going to talk about coordinates. In the real numbers $\mathbb{R}$, there are the rational numbers $\mathbb{Q}$ and the irrational numbers like $\sqrt{2}$. A rational number is a number of the form $m = \frac{a}{b}$ for $a, b \in \mathbb{Z}$, but we can also write this as a decimal number

$$m = \frac{a}{b} = \alpha.\beta\gamma\delta\epsilon\ldots = \alpha + \frac{\beta}{10} + \frac{\gamma}{100} + \frac{\delta}{1000} + \cdots .$$

This can be taken as the working definition of real numbers. But one problem with this is that decimal representations of numbers are not unique:

$$1 = 0.9999\ldots.$$

That is why we need a rigorous definition, similar to Hilbert's axiom.

7.1 Dedekind cuts

Definition 7.1. A **Dedekind cut** is nonempty partition of $\mathbb{Q}$ into $\mathbb{Q} = A \cup B$, such that

(i) A has no largest element,

(ii) all elements in A are less than those B.

How does this correspond to the usual notion of real numbers. By the two conditions, a Dedekind cut will necessarily look like $A = \{x \in \mathbb{Q} : x < c\}$ and $B = \{x \in \mathbb{Q} : x \geq c\}$ for some real number c. This c at the boundary of A and B is what the Dedekind cut should correspond to.

Example 7.2. The square root $\sqrt{2}$ corresponds to the Dedekind cut

$$A = \{\alpha \in \mathbb{Q} : \alpha \leq 0 \text{ or } \alpha^2 < 2\}, \quad B = \{\beta \in \mathbb{Q} : \beta^2 \geq 2 \text{ and } \beta > 0\}.$$

Even though $\mathbb{Q}$ does not contain the element $\sqrt{2}$, what defines this partition is something like $\sqrt{2}$. But we still have to check that A, B is really a Dedekind cut.

Proposition 7.3. *The set $A = \{\alpha \in \mathbb{Q} : \alpha \leq 0$ or $\alpha^2 < 2\}$ has no largest element.*

Proof. Suppose there exists a largest element $x \in A$. We want to construct a $y \in A$ with $y > x$ and $y^2 < 2$. We do this by guessing

$$y = \frac{2x+2}{x+2}.$$

This is a rational number, and you can check that $x < y$ and $y^2 < 2$. To check $y^2 < 2$, we compute

$$\begin{aligned} y^2 = \Big(\frac{2x+2}{x+2}\Big)^2 &= \frac{4x^2+8x+4}{x^2+4x+4} \\ &= \frac{2(x^2+4x+4)+2(x^2-2)}{x^2+4x+4} = 2 + \frac{2(x^2-2)}{x^2+4x+4} < 2. \end{aligned}$$

This contradicts that x is the largest element. □

Definition 7.4. We define the **real numbers** $\mathbb{R}$ as the set of Dedekind cuts on $\mathbb{Q}$.

It's interesting how people construct numbers. First we learn how to count, like $1, 2, \ldots$. Mathematically, we define

$$0 = \{\}, \quad 1 = \{0\} = \{\{\}\}, \quad 2 = \{0, 1\} = \{\{\}, \{\{\}\}\}, \ldots.$$

This is the set-theoretic approach, but we can also just define natural numbers axiomatically. This is called Peano's axioms. There exist $0, 1$, and for each k, there exists $k+1$, and something like induction. So anyways, we can define the natural numbers $\mathbb{N} = \{0, 1, 2, \ldots\}$ in this way.

Next, people came up with the concept of negative numbers. For $n \in \mathbb{N}$, we formally say that there is a number a such that $a + n = 0$, even if this doesn't make sense in $\mathbb{N}$. Then we can think of a as $-n$, so that $n + (-n) = 0$. There is also multiplication in $\mathbb{N}$, which is defined as

$$n \cdot a = \overbrace{a + \cdots + a}^{n}.$$

There exists this special element 1 such that $1 \cdot a = a \cdot 1 = a$ for all a. So we formally define a number m such that $nm = 1$, which we think of as 1 divided by n, denoted $\frac{1}{n}$. This way, we get the rational numbers $\mathbb{Q}$. From this, we can get the real numbers $\mathbb{R}$ using Dedekind cuts.

7.2 Coordinate plane

In this chapter, we are going to basically reprove Euclid's geometry. Here, we describe the plane as

$$\text{Plane} = \mathbb{R}^2 = \{(x, y) : x, y \in \mathbb{R}\}.$$

The story is that Descartes was sick and lying on his bed, and he came up with the concept of coordinates by watching flies moving around on the ceiling.

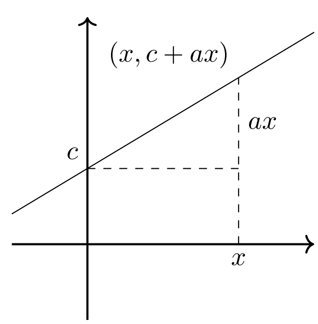

Figure 19: A line in a coordinate plane

A line is given by a slope. This is given by the equation

$$y = ax + c$$

as long as the line is not vertical. Here a is the **slope**, which is intuitively the "rise over run", and c is the y-intercept, i.e., the intersection of the line and the y-axis. But this formula cannot describe the vertical line. A general formula of a line is given by

$$ax + by + c = 0$$

for $(a, b) \neq (0, 0)$. The line is vertical if $b = 0$ and horizontal if $a = 0$.

There was an axiom (E1) that states that any two points can be joined by a straight line. Suppose $P = (x_1, y_1)$ and $Q = (x_2, y_2)$ are two points. Then the line with equation given by

$$\ell : \frac{y - y_2}{x - x_2} = \frac{y_1 - y_2}{x_1 - x_2}$$

passes through P and Q. (Here, we're assuming that $x_1 \neq x_2$, but if $x_1 = x_2$ you can simply take $\ell : x = x_1$.)

We can also prove the parallel postulate. Consider an arbitrary line $\ell : y = ax+c$. (Assume for simplicity that it's not vertical.) To construct a line parallel to ℓ passing through (x_1, y_1), you can take

$$\ell' : (y - y_1) = a(x - x_1) + c.$$

8 February 7, 2018

Last time we looked at the equation of a line in $\mathbb{R}^2$, and used this to reprove the axioms.

8.1 Distance in coordinates

Distance is going to be defined using the Pythagorean theorem. Consider two points (x_1, y_1) and (x_2, y_2) on the coordinate plane. These two points and (x_2, y_1) will form a right triangle, and so then the distance between $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$ can be written as

$$|P_1P_2| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}.$$

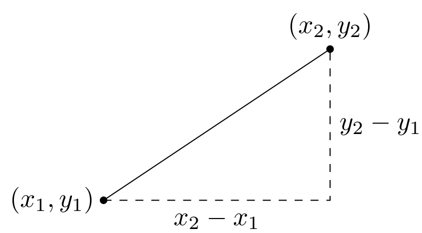

Figure 20: Distance between two points

Using this, we can write down the equation of a circle. Consider a circle of radius r centered at (c_1, c_2). Then a point (x, y) lying on the circle is equivalent to its distance from (c_1, c_2) being r, and this can be written as

$$(x - c_1)^2 + (y - c_2)^2 = r^2.$$

Note that everything we have been doing can be generalized to n-dimensional space.

Proposition 8.1. *The set of points that are equidistant from two given points is a line.*

Geometrically, we know that this line should be the perpendicular bisector of the two given points. But this can be shown algebraically.

Proof. Let (x_1, y_1) and (x_2, y_2) be the two given points. The equation for equidistance is

$$\sqrt{(x - x_1)^2 + (y - y_1)^2} = \sqrt{(x - x_2)^2 + (y - y_2)^2},$$

which becomes

$$\begin{aligned} x^2 - 2x_1x + x_1^2 + y^2 - 2y_1y + y_1^2 &= (x - x_1)^2 + (y - y_1)^2 \\ &= (x - x_2)^2 + (y - y_2)^2 \\ &= x^2 - 2x_2x + x_2^2 + y^2 - 2y_2y + y_2^2 \end{aligned}$$

after squaring both sides. Then canceling the x^2 and y^2 terms give

$$x(2x_2 - 2x_1) + y(2y_2 - 2y_1) + (x_1^2 + y_1^2 - x_2^2 - y_2^2) = 0.$$

This shows that it is a line, with $a = 2x_2 - 2x_1$, $b = 2y_2 - 2y_1$, and $c = x_1^2 + y_1^2 - x_2^2 - y_2^2$. □

In particular, the midpoint $(\frac{1}{2}(x_1 + x_2), \frac{1}{2}(y_1 + y_1))$ is on the line.

8.2 Intersection points

Consider two lines

$$\ell : ax + by + c = 0, \quad \ell' : a'x + b'y + c' = 0.$$

Assume that ℓ and ℓ' are not vertical, and not parallel. Then b, b' are both nonzero. Then we can divide the equations by b and b' to make

$$\ell : \tfrac{a}{b}x + y + \tfrac{a}{c} = 0, \quad \ell' : \tfrac{a'}{b'}x + y + \tfrac{a'}{c'} = 0.$$

Here, because the two lines are not parallel, we have $\frac{a}{b} \neq \frac{a'}{b'}$. Then we can subtract the two equations to get

$$(\tfrac{a}{b} - \tfrac{a'}{b'})x + (c - c') = 0,$$

and then we get

$$x = \frac{-(c' - c)}{\frac{a'}{b'} - \frac{a}{b}}, \quad y = \cdots .$$

If one of ℓ and ℓ' is vertical, you can also show that the two lines intersect at one point. I'll leave it to you as an exercise.

Proposition 8.2. *Two lines that are not parallel intersect at a unique point.*

Now let us look at two circles. Consider two circles centered at (x_1, y_1) and (x_2, y_2), with radii r_1 and r_2 respectively. Because the circles don't meet unless the radii are sufficeintly large, we need something like

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \leq r + s$$

to have existence of an intersection. But this is not enough, because one circle might contain the other one completely.

You can calculate the intersections points algebraically. For simplicity, assume that $(x_2, y_2) = (0, 0)$ by just shifting the coordinates. Then the two equations for the circles are

$$\begin{cases} x^2 + y^2 = r^2, \\ (x - x_1)^2 + (y - y_1)^2 = s^2. \end{cases}$$

Then

$$s^2 = (x^2 - 2x_1x + x_1^2) + (y^2 - 2y_1y + y_1^2) = r^2 - 2x_1x + x_1^2 - 2y_1y + y_1^2.$$

Assuming $x_1 \neq 0$, we can now write x as

$$x = \frac{(r^2 - s^2 + x_1^2 + y_1^2) - 2y_1y}{2x_1}$$

and plugging this into $x^2 + y^2 = r^2$, we get a quadratic equation in y. So this equation can have 0 or 1 or 2 solutions depending on the discriminant. You can explicitly write down the conditions under which the two circles have certain number of intersection points, but I'm not going to do it here.

The computations we did have an interesting consequence. Recall that we have learned how to do $+$, $-$, $\times$, $\div$, and $\sqrt{\ }$. But conversely, whatever we do in Euclidean geometry, we can write down its coordinates using these five operations. So Euclidean geometry can be thought of as a geometric way of doing these operations.

Definition 8.3. A point is called **constructible** if its coordinate can be obtained by $+$, $-$, $\times$, $\div$, and $\sqrt{\ }$.

Not all real numbers are constructible. Showing that a number is not constructible is very hard, but you can show that $\sqrt[3]{2}$ is not constructible. This can be shown using Galois theory. Numbers like π or e are also not constructible, and so are Liouville numbers such as $0.101000000100000000\dots$.

8.3 Angle in coordinates

How can we describe angles? There is a notion of a slope of a line, and if we denote by θ the angle between it and the x-axis, the slope is going to be

$$\frac{\sin\theta}{\cos\theta} = \tan\theta.$$

But given two lines ℓ_1 and ℓ_2, we want to know the angle between them, not the angle between them and the x-axis. If we denote by θ_1 and θ_2 those angles, we want to find out $\theta_2 - \theta_1$. This can be computed by the formula

$$\tan(\theta_2 - \theta_1) = \frac{\tan\theta_1 - \tan\theta_2}{1 + \tan\theta_1 \tan\theta_2}.$$

If (x_1, y_1) and (x_2, y_2) are points on ℓ_1 and ℓ_2 of distance 1 from the origin, we can also express

$$\begin{aligned}\cos(\theta_2 - \theta_1) &= x_1x_2 + y_1y_2,\\ \sin(\theta_2 - \theta_1) &= x_1y_2 - y_1x_2 = \det\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}.\end{aligned}$$

9 February 9, 2018

We have talked about constructible numbers. These are numbers that can be made from 1 using $+, -, \times, \div, \sqrt{\ }$.

Definition 9.1. An **algebraic number** is a number that is a root of a polynomial with rational coefficients.

It turns out that roots of polynomials with algebraic number coefficients are still algebraic. So in particular, constructible numbers are algebraic, because the square root can be expressed as a root of $x^2 - a = 0$.

$$\begin{aligned}\mathbb{Q} = \{\text{rational numbers}\} &\subsetneq \{\text{constructible numbers}\} \\ &\subsetneq \{\text{algebraic numbers}\} \subsetneq \{\text{real numbers}\} = \mathbb{R}\end{aligned}$$

(The standard definitions of constructible and algebraic numbers are actually made in $\mathbb{C}$.)

Real numbers (or complex numbers) that are not algebraic are called **transcendental**. Here is an example:

$$L = \sum_{k=1}^{\infty} 10^{-k!} = 10^{-1} + 10^{-2} + 10^{-6} + 10^{-24} + \cdots.$$

To see that this is not algebraic, suppose the contrary and assume that $p(L) = 0$. Separate the p into the positive coefficient part and the negative coefficient part, so that $p = p_+ - p_-$ and then $p_+(L) = p_-(L)$. Now if we expand both sides of $p_+(\sum_{k=1}^{\infty} 10^{-k})$ and $p_-(\sum_{k=1}^{\infty} 10^{-k})$ and compare the decimal expansion, you will see that they cannot be equal unless $p_+ = p_-$.

9.1 Formulas for trigonometric functions

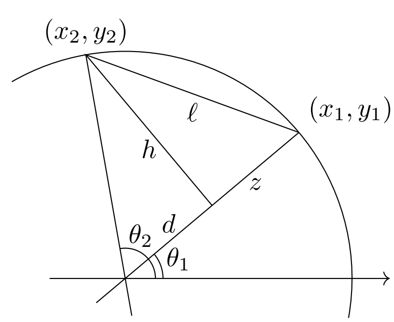

Figure 21: Computing sin and cos of an angle between lines

Let ℓ_1, ℓ_2 be two lines through the origin, and let (x_1, y_1) and (x_2, y_2) be points on the lines with both points being distance 1 from 0. If we let θ_1, θ_2

be the angle between ℓ_1, ℓ_2 and the x-axis, then the angle between ℓ_1 and ℓ_2 is $\theta_2 - \theta_1$.

To compute $\theta = \theta_2 - \theta_1$, we draw a line of perpendicular to ℓ_1 from (x_2, y_2). Then we can start computing the lengths $d = \cos\theta$ and $h = \sin\theta$. By the Pythagorean theorem, we have

$$d + z = 1, \quad d^2 + h^2 = 1, \quad h^2 + z^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2.$$

Then we can cancel out

$$h^2 + z^2 = h^2 + (1 - d)^2 = h^2 + d^2 + 1 - 2d = 2 - 2d.$$

But then

$$h^2 + z^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 = 2 - 2x_1y_1 - 2x_2y_2,$$

so we get

$$\cos(\theta_2 - \theta_1) = d = x_1x_2 + y_1y_2,$$

and likewise we can compute

$$\sin(\theta_2 - \theta_1) = h = x_1y_2 - x_2y_1.$$

9.2 Isometries of the plane

An isometry is a transformation in $\mathbb{R}^2$ that preserves lengths. These are what Euclid must had in mind when he was talking about "congruence". Mathematically, you can define

Definition 9.2. An **isometry** is a map $f : \mathbb{R}^2 \to \mathbb{R}^2$ such that for all $P, Q \in \mathbb{R}^2$ we have

$$\operatorname{dist}(P, Q) = \operatorname{dist}(f(P), f(Q)).$$

Here are some examples:

- the identity map given by $(x, y) \mapsto (x, y)$
- translations given by $(x, y) \mapsto (x + x_0, y + y_0)$
- rotations by a certain angle about a certain point
- reflection about a certain line
- any compositions of the previous three, in particular, first reflecting and translating in the direction of the line of reflection.

Here is the main theorem we are going to prove about isometries of $\mathbb{R}^2$.

Theorem 9.3 (three reflections theorem)**.** *Every isometry of $\mathbb{R}^2$ can be obtained by at most 3 reflections.*

We need the following fact.

Proposition 9.4. *An isometry is always determined* 3 *points not lying on the same line.*

Proof. Consider three points P, Q, R that are not on a line. We claim that each point $S \neq P, Q, R$ can be determined by the distances $\text{dist}(S, P)$, $\text{dist}(S, Q)$, and $\text{dist}(S, R)$. To see this assume that $S \neq S'$ are different points and $\text{dist}(S, P) = \text{dist}(S', P)$ and so on. Then P lies on the perpendicular bisector of S, S', and likewise Q, R lies on this line. This contradicts the fact that P, Q, R don't lie on a line.

Now the proposition follows from this. Suppose $f : \mathbb{R}^2 \to \mathbb{R}^2$ is an isometry, and the points $f(P), f(Q), f(R)$ are given. We want to completely determine f from this data. Consider an arbitrary point $S \neq P, Q, R$. We know $|PS|, |QS|, |RS|$, and so we know $|f(P)f(S)|, |f(Q)f(S)|, |f(R)f(S)|$. We also know $f(P), f(Q), f(S)$, and so the thing we proved in the last paragraph shows that $f(S)$ is determined by this data. □

Of course, this does not mean that we can map P, Q, R to arbitrary points $f(P), f(Q), f(R)$, and this will extend to an isometry. But the proposition tells us that there is going to be at most one such isometry.

10 February 12, 2018

We are trying to prove the "three reflection theorem", that any isometry of $\mathbb{R}^2$ is a combination of at most 3 reflections. We proved the following lemma last time.

Lemma 10.1. *Any isometry is determined by its image of* 3 *distinct points not on a line.*

Because of this lemma, it suffices to check that we can send three points to the right places.

Example 10.2. Consider the example of the isometry f that is rotation about the origin by $\frac{\pi}{2}$ clockwise. Consider $O = (0, 0)$, $A = (0, 1)$, and $B = (2, 0)$. Then

$$f(O) = (0, 0), \quad f(A) = (1, 0), \quad B = (0, -2).$$

First, we want to make $A \mapsto f(A)$, so we consider the reflection r_A about the line $y = x$. Then

$$r_A(O) = (0, 0), \quad r_A(A) = (1, 0), \quad r_A(B) = (0, 2).$$

Now O and A are send to the right place, but B is not. So we reflect again about the line $y = 0$ that passes through A and O (call this reflection r_B). Then

$$r_B \circ r_A(O) = (0, 0), \quad r_B \circ r_A(A) = (1, 0), \quad r_B \circ r_A(B) = (0, -2).$$

Then f must equal to $r_B \circ r_A$, because they are isometries and agree at three points. So in this example, we can decompose f into two reflections.

Proof of Theorem 9.3. In general, the strategy is going to be the same. Pick arbitrary three points A, B, C. Consider a reflection r_A such that $r_A(A) = f(A)$. Then

$$|f(A)r_A(B)| = |r_A(A)r_A(B)| = |AB| = |f(A)f(B)|$$

so the perpendicular bisector of $f(B)r_A(B)$ passes through $f(A)$. Consider the reflection r_B about this line. Then so far,

$$r_B \circ r_A(A) = r_B(f(A)) = f(A), \quad r_B \circ r_A(B) = f(B).$$

Now $|f(C)f(A)| = |r_B(r_A(C))f(A)|$ and $|f(C)f(B)| = |r_B(r_A(C))f(B)|$. This means that either $r_B(r_A(C)) = f(C)$, in which case we are done, or $r_B(r_A(C))$ is the reflection of $f(C)$ about the line $f(A)f(B)$, in which case we can set r_C as the reflection about this line. In the first case, we get

$$r_B \circ r_A = f,$$

and in the second case, we get

$$r_C \circ r_B \circ r_A = f.$$

So f can be written as a composition of at most 3 reflections. □

The set of isometries is naturally a group. This just means that we can compose isometries and get an isometry, there is an identity isometry, and that inverses of isometries are isometries.

10.1 Vector spaces

A **vector space** is a set of vectors with the structure of

- adding and subtracting vectors $\mathbf{v}, \mathrm{w}$ like $\mathbf{v}+\mathbf{w}, \mathbf{v}-\mathbf{w}$
- a chosen zero vector $\mathbf{0}$ such that $\mathbf{0}+\mathbf{v}=\mathbf{v}+\mathbf{0}=\mathbf{0}$
- multiplication by scalar, $\alpha\mathbf{v}$ for $\alpha \in \mathbb{R}$

satisfying the properties

- distributive laws $\alpha(\mathbf{u}+\mathbf{v})=\alpha\mathbf{u}+\alpha\mathbf{v}$ and $(\alpha+\beta)\mathbf{v}=\alpha\mathbf{v}+\beta\mathbf{v}$
- $\alpha(\beta\mathbf{v})=(\alpha\beta)\mathbf{v}$.

The point of this abstraction is to think about vectors not as elements of $\mathbb{R}^n$ but as something with structure with addition and scalar multiplication. With this viewpoint, you can think of matrices as vectors, or functions as vectors as long as you can multiply them by scalars and add them together.

Let's see what addition and scalar multiplication corresponds in 2-dimensional Euclidean geometry. Scalar multiplication is just **dilation**, that is, having the same direction but different length. Addition is given by the parallelogram rule.

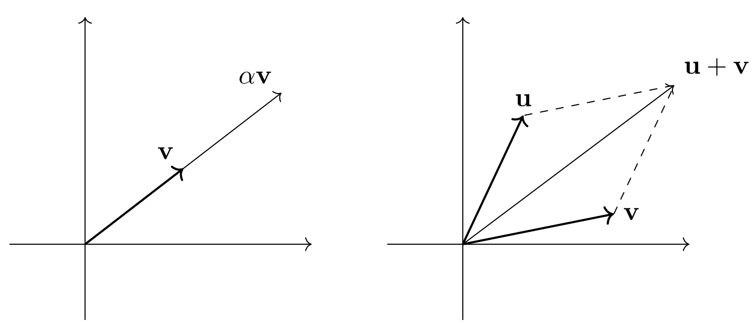

Figure 22: Dilation and addition in $\mathbb{R}^2$

Definition 10.3. We say that $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$ are **linearly dependent** if there exist $(\alpha, \beta) \neq (0,0)$ such that $\alpha\mathbf{u}+\beta\mathbf{v}=0$. This just means that $\mathbf{u}$ and $\mathbf{v}$ lie on a same line. We say that $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$ are **linearly independent** if they are not linearly independent.

Definition 10.4. We say that $\mathbf{w}$ is of **direction $\mathbf{u}$** from $\mathbf{v}$ if there exists an $\alpha \in \mathbb{R}$ such that

$$\mathbf{w}-\mathbf{v}=\alpha\mathbf{u}.$$

This means that the line connecting $\mathbf{v}$ and $\mathbf{w}$ being parallel to $\mathbf{u}$.

Definition 10.5. Using this idea, we can define that a line from $\mathbf{u}$ to $\mathbf{v}$ is **parallel** to the line from $\mathbf{s}$ to $\mathbf{t}$ if

$$\mathbf{v}-\mathbf{u}=\alpha(\mathbf{t}-\mathbf{s})$$

for some $\alpha \neq 0$.

A **line** going between $\mathbf{u}$ and $\mathbf{v}$ can be parametrized by the equation

$$\gamma : \mathbb{R} \to \mathbb{R}^2; \quad t \mapsto t(\mathbf{v} - \mathbf{u}) + \mathbf{u}.$$

Then $\gamma(0) = \mathbf{u}$ and $\gamma(1) = \mathbf{v}$ and at time t, it goes t steps in the direction of $\mathbf{v} - \mathbf{u}$.

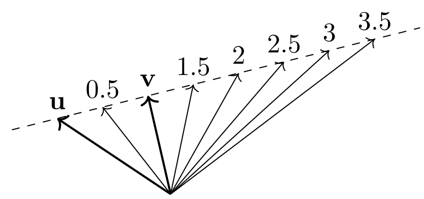

Figure 23: The line $\gamma(t) = t(\mathbf{v} - \mathbf{u}) + \mathbf{u}$ for various t

10.2 Thales's theorem again

We can now state and prove the vector formulation of Thales's theorem.

Theorem 10.6 (Thales's theorem)**.** *Let* $\mathbf{s}$ *and* $\mathbf{v}$ *be nonzero on one line through* $\mathbf{0}$*, and* $\mathbf{t}$ *and* $\mathbf{w}$ *be nonzero on another line through* $\mathbf{0}$*. If* $\mathbf{w} - \mathbf{v}$ *is parallel to* $\mathbf{t} - \mathbf{s}$*, then there exists a* $\alpha \neq 0$ *such that*

$$\mathbf{v} = \alpha \mathbf{s}, \quad \mathbf{w} = \alpha \mathbf{t}.$$

(This is like saying that $\frac{|\mathbf{v}|}{|\mathbf{s}|} = \frac{|\mathbf{w}|}{|\mathbf{t}|}$*.)*

Proof. Because $\mathbf{v}$ and $\mathbf{s}$ lie on a line, we have $\mathbf{v} = \alpha \mathbf{s}$ for some $\alpha \neq 0$. Likewise, we have $\mathbf{w} = \beta \mathbf{t}$ for some $\beta \neq 0$. But by the parallel condition, we have

$$(\mathbf{w} - \mathbf{v}) = \gamma(\mathbf{t} - \mathbf{s})$$

for some γ. By rearranging, we get

$$(\beta - \gamma)\mathbf{t} + (\gamma - \alpha)\mathbf{s} = 0.$$

But because $\mathbf{t}$ and $\mathbf{s}$ lie an distinct lines, they are linearly independent. This shows that $\beta - \gamma = \gamma - \alpha = 0$, and so $\alpha = \beta = \gamma$. $\square$

11 February 14, 2018

We have started talking about vectors. For $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$, we say that they are linearly dependent if $\alpha\mathbf{u} + \beta\mathbf{v} = 0$ implies $\alpha = \beta = 0$.

Theorem 11.1 (Thales)**.** *Let* $\mathbf{s}, \mathbf{v}$ *be nonzero on a line through* $\mathbf{0}$ *and* $\mathbf{t}, \mathbf{w}$ *be nonzero on a different line through* $\mathbf{0}$*. If* $\mathbf{w} - \mathbf{v}$ *is parallel to* $\mathbf{t} - \mathbf{s}$*, then* $\mathbf{v} = \alpha\mathbf{s}$ *and* $\mathbf{w} = \alpha\mathbf{t}$ *for some* $\alpha \neq 0$*.*

11.1 Pappus's theorem

The following is theorem can be proved using vectors. We are going to see this again in the context of projective geometry.

Theorem 11.2 (Pappus)**.** *Let* $\mathbf{r}, \mathbf{t}, \mathbf{v}$ *be nonzero on a line through* $\mathbf{0}$*, and* $\mathbf{s}, \mathbf{u}, \mathbf{w}$ *be nonzero on another line through* $\mathbf{0}$*. If* $\mathbf{u} - \mathbf{v}$ *is parallel to* $\mathbf{s} - \mathbf{r}$ *and* $\mathbf{t} - \mathbf{s}$ *is parallel to* $\mathbf{v} - \mathbf{w}$*, then* $\mathbf{u} - \mathbf{t}$ *is parallel to* $\mathbf{w} - \mathbf{r}$*.*

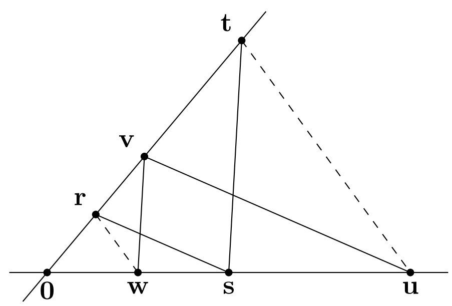

Figure 24: Pappus's theorem

Proof. By Thales's theorem, we have

$$\mathbf{u} = \alpha\mathbf{s}, \quad \mathbf{v} = \alpha\mathbf{r}$$

for some $\alpha \neq 0$. Likewise, we have

$$\mathbf{s} = \beta\mathbf{w}, \quad \mathbf{t} = \beta\mathbf{v}.$$

So

$$\mathbf{u} - \mathbf{t} = \alpha\mathbf{s} - \beta\mathbf{v} = \alpha(\beta\mathbf{w}) - \beta(\alpha\mathbf{r}) = \alpha\beta(\mathbf{w} - \mathbf{r}).$$

So $\mathbf{u} - \mathbf{t}$ is parallel to $\mathbf{w} - \mathbf{r}$. $\square$

11.2 Centroid of a triangle

What is the midpoint of $\mathbf{u}+\mathbf{v}$? You can see this from the parallelogram rule. The midpoint of $\mathbf{u}$ and $\mathbf{v}$ is going to be the midpoint of $\mathbf{0}$ and $\mathbf{u}+\mathbf{v}$, which is

$$\frac{\mathbf{u}+\mathbf{v}}{2}.$$

In general, the **barycenter** or **center of mass** is

$$\frac{1}{n}(\mathbf{u}_1+\mathbf{u}_2+\cdots+\mathbf{u}_n).$$

Definition 11.3. A **median** is a line from a vertex of a triangle to the midpoint of the opposite side.

Theorem 11.4. *The three medians of any triangle pass through the same point, which is the called the* ***centroid*** *of the triangle.*

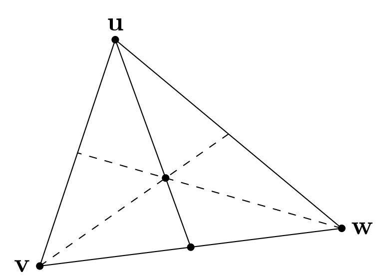

Figure 25: Centroid of a triangle

Proof. Let $\mathbf{u},\mathbf{v},\mathbf{w}$ be the three vertices of the triangle. The line ℓ passing through $\mathbf{u}$ and $\frac{1}{2}(\mathbf{v}+\mathbf{w})$ is going to be

$$\mathbf{u}+t\Big(\frac{1}{2}(\mathbf{v}+\mathbf{w})-\mathbf{u}\Big).$$

If we guess $t=\frac{2}{3}$, we see that the line ℓ passes through the point

$$\mathbf{u}+\frac{2}{3}\Big(\frac{1}{2}(\mathbf{v}+\mathbf{w})-\mathbf{u}\Big)=\frac{1}{3}(\mathbf{u}+\mathbf{v}+\mathbf{w}).$$

We now observe that this is symmetric in $\mathbf{u},\mathbf{v},\mathbf{w}$, so that all three medians should pass through this point. ☐

11.3 Inner product and cosine

Definition 11.5. For $\mathbf{u}$ and $\mathbf{v}$ with angle θ between them, we define the **inner product** as

$$\mathbf{u}\cdot\mathbf{v}=|\mathbf{u}||\mathbf{v}|\cos\theta.$$

The explicit formula is given by

$$(u_1, u_2) \cdot (v_1, v_2) = u_1 v_1 + u_2 + v_2.$$

Proposition 11.6. *If* $\mathbf{u}$ *and* $\mathbf{v}$ *are vectors with angle* θ *between them, then*

$$|\mathbf{u} - \mathbf{v}|^2 = |\mathbf{u}|^2 + |\mathbf{v}|^2 - 2|\mathbf{u}||\mathbf{v}| \cos\theta.$$

Proof. You can expand the left hand side after writing it as $(\mathbf{u}-\mathbf{v})\cdot(\mathbf{u}-\mathbf{v})$. □

Proposition 11.7 (triangle inequality)**.** *We always have* $|\mathbf{u} + \mathbf{v}| \leq |\mathbf{u}| + |\mathbf{v}|$.

Proof. If we square both sides, expand them, and cancel things we can, it boils down to

$$\mathbf{u} \cdot \mathbf{v} \leq |\mathbf{u}||\mathbf{v}|.$$

This is Cauchy–Schwartz, or you can think of it as following from $\cos\theta \leq 1$. □

Theorem 11.8. *In any triangle, the three perpendiculars from the vertices to opposite sides (called **altitiudes**) meet at a common point.*

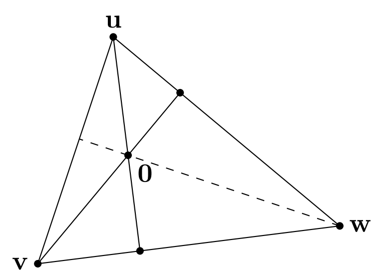

Figure 26: Orthocenter of a triangle

Proof. We may assume that the two altitudes from $\mathbf{u}$ and $\mathbf{v}$ meet at the origin $\mathbf{0}$. This because we can translate the whole picture and nothing changes. Then we have that $\mathbf{u}$ is perpendicular to $\mathbf{w} - \mathbf{v}$, and $\mathbf{v}$ is perpendicular to $\mathbf{u} - \mathbf{w}$. We can write this as

$$\mathbf{u} \cdot (\mathbf{w} - \mathbf{v}) = 0, \quad \mathbf{v} \cdot (\mathbf{u} - \mathbf{w}) = 0.$$

This can also be written as

$$\mathbf{u} \cdot \mathbf{w} - \mathbf{u} \cdot \mathbf{v} = 0, \quad \mathbf{v} \cdot \mathbf{u} - \mathbf{v} \cdot \mathbf{w} = 0.$$

Adding the two gives

$$0 = \mathbf{u} \cdot \mathbf{w} - \mathbf{v} \cdot \mathbf{w} = (\mathbf{u} - \mathbf{v}) \cdot \mathbf{w}.$$

Therefore $\mathbf{u} - \mathbf{v}$ is perpendicular to $\mathbf{w}$, and this means that the other altitude also passes through $\mathbf{0}$. □

If we look at $\mathbb{R}$, we have $+, -, \times, \div$. But in $\mathbb{R}^2$ we have $+$ and $-$ but not $\times$ or $\div$. This motivates us to define the complex numbers $\mathbb{C}$. This is sort of the same as $\mathbb{R}^2$, but write

$$(a, b) \quad \longleftrightarrow \quad a + bi.$$

Here, i is considered as a formal symbol satisfying $i^2 = -1$. So for instance,

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i.$$

This complex setup will allow us to write rotation in a simple way.

12 February 16, 2018

Today we are going to talk about complex numbers. There is $\mathbb{R}$ and $\mathbb{C}$, which is $\mathbb{R}^2$ with some algebra structure. It turns out that the next (normed) algebra structure that can be given is $\mathbb{H} \cong \mathbb{R}^4$, and then $\mathbb{O} \cong \mathbb{R}^8$. Then $\mathbb{R}^{16}$ does not have a good algebra structure.

12.1 Multiplication in the complex numbers

Multiplication in $\mathbb{C}$ is like rotation in $\mathbb{R}^2$. Take any $a+ib$, can consider the map

$$(a+ib)(-) : \mathbb{R}^2 \cong \mathbb{C} \to \mathbb{C} \cong \mathbb{R}^2; \quad z \mapsto (a+ib)z.$$

If we consider this as a linear map $\mathbb{R}^2 \to \mathbb{R}^2$, we can think of it as

$$\begin{pmatrix} c \\ d \end{pmatrix} \mapsto \begin{pmatrix} ac - bd \\ bc + ad \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix}.$$

Let us consider $u = \cos\theta + i\sin\theta$ be a complex number of norm 1. Then the multiplication corresponding to u is

$$u(-) : \mathbb{R}^2 \to \mathbb{R}^2; \quad \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

and we see that this is rotation by θ. That is, uz is z rotated by θ. More generally, $(ru)z$ is z rotated by θ and dilated by r.

We define

$$e^{i\theta} = \cos\theta + i\sin\theta.$$

The good thing about this definition is that we have

$$e^{i\theta} \cdot e^{i\eta} = (\cos\theta + i\sin\theta)(\cos\eta + i\sin\eta) = \cos(\theta+\eta) + i\sin(\theta+\eta) = e^{i(\theta+\eta)}.$$

Conversely, you can read off the addition formulas from this.

12.2 Projective plane

The motivation for studying projective geometry came from art. If there is a square tiling on the floor, how can you draw this? We first draw the horizon, and then one tile. Next, we draw the diagonal, mark the point where it meets the horizon, and then draw the next tiles so that the diagonal passes through the original tile.

There are many different ways to define a projective plane. Let us first look at the axiomatic definition. First observe that

- straight lies are always straight in the projective plane,
- intersection of lines are intersection,
- parallel lines meet on the 'horizon'.

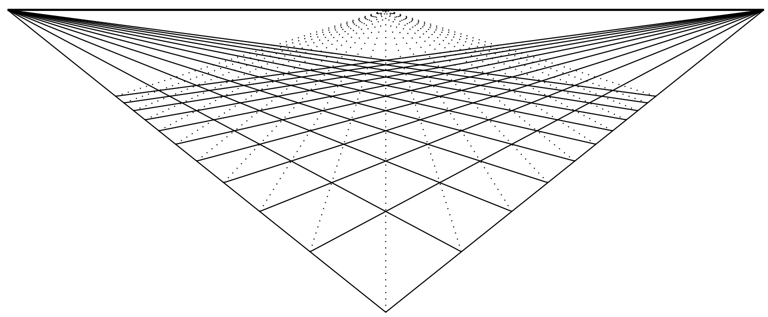

Figure 27: Tiling on a plane

So here are the axioms for the **projective plane**.

(PP1) Any two 'points' are contained in a unique 'line'.

(PP2) Any two 'lines' intersect at a single point.

(PP3) There exist four 'points' with no three of them being on a 'line'.

The second axiom might seem confusing. But we are considering parallel lines as meeting on the horizon line. The third axioms is just saying that we have an interesting geometry.

But does there exist a projective plane? It will be meaning less if there does not exist such an object after we defined it.

Example 12.1. We define the **real projective plane** $\mathbb{R}P^2$ in the following way.

(a) Points in $\mathbb{R}P^2$ are the lines through the origin in $\mathbb{R}^3$.

(b) Lines in $\mathbb{R}P^2$ are the planes through the origin in $\mathbb{R}^3$.

(c) There are four points $(1,0,0)$, $(0,1,0)$, $(0,0,1)$, $(1,1,1)$ (which should be interpreted as lines connecting these points and the origin) that do not lie on a line.

You should think of your eye as the origin, and a point you see as a line that connects you eye and that point.

13 February 21, 2018

We have given an axiomatic characterization of a projective plane. Our claim is that there exists a projective plane.

13.1 Real projective plane

Consider the following structure:

(a) points in the real projective plane $\mathbb{R}P^2$ are lines through the origin in $\mathbb{R}^3$,

(b) lines in $\mathbb{R}P^2$ are planes through the origin in $\mathbb{R}^3$,

(c) a line passes through a point when the corresponding plane contains the corresponding line in $\mathbb{R}^3$.

This is called the **real projective plane**. You can check that no three of the lines passing $(1,0,0)$, $(0,1,0)$, $(0,0,1)$, $(1,1,1)$ in $\mathbb{R}^3$ do not lie a single plane. This means that these four points in $\mathbb{R}P^2$ satisfies the last axiom (PP3). Actually, we can make this work over any field. Here we are working with $\mathbb{R}$, but you can do this over finite fields $\mathbb{F}_q$ as well.

Let us now check the first two axioms (PP1) and (PP2). Consider two given points $\mathbb{R}P^2$. They can be written as $\{t\vec{u} : t \in \mathbb{R}\}$ and $\{s\vec{v} : s \in \mathbb{R}\}$. Because the two points are distinct, the two vectors $\vec{u}$ and $\vec{v}$ are linearly independent. Then there is a unique plane containing $\vec{u}$ and $\vec{v}$, which is the plane generated by $\vec{u}$ and $\vec{v}$. More concretely, we can take $\vec{a} = \vec{u} \times \vec{v}$ and then the plane we are looking for can be written as

$$\vec{a} \cdot \vec{x} = 0.$$

For (PP2), consider any two lines in $\mathbb{R}P^2$, which will correspond to two planes in $\mathbb{R}^3$ passing through the origin. We can write down these planes using the equations

$$\vec{a} \cdot \vec{x} = 0, \quad \vec{a}' \cdot \vec{x} = 0.$$

Again, the two vectors $\vec{a}, \vec{a}'$ are linearly independent. We want to find the intersection point of these two lines (in $\mathbb{R}P^2$), and this is finding the intersecting line of the two planes (in $\mathbb{R}^3$). We can find this line as

$$\{t(\vec{a} \times \vec{a}') : t \in \mathbb{R}\},$$

because this is the vector that is perpendicular to both.

So we now know that $\mathbb{R}P^2$ indeed is a projective plane. But this is a geometry class, and we want to visualize what this is. There are two ways to think about $\mathbb{R}P^2$. The first way is to project to some other plane. But for simplicity, let us first look at $\mathbb{R}P^1$. This is the set of lines passing through the origin in $\mathbb{R}^2$. To get a good description of this space, first pick a line $\ell : y = 1$ in $\mathbb{R}^2$. Then we can identify a point in $\mathbb{R}P^1$, which is a line in $\mathbb{R}^2$, with its intersection with ℓ. But if you think about this, there is one line (in $\mathbb{R}^2$) that does not meet ℓ, namely the x-axis. So there is one extra point attached to just a line:

$$\mathbb{R}P^1 \cong \mathbb{R} \cup \{\infty\}.$$

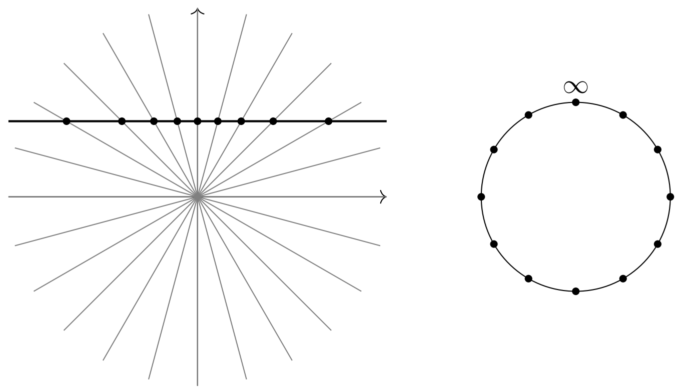

Figure 28: Visualizing $\mathbb{R}P^1$

If you think about this, the positive ∞ is the same as the negative ∞. So topologically, you can also visualize this as a circle S^1, where the two endpoints of a infinitely long line are somehow identified. Another way we can think about this is to project on onto the unit circle. Then every line through the origin meets the circle at two opposite points. So $\mathbb{R}P^1$ is a circle with the two opposite points glued together, or a circle wrapped around itself twice if you'd like.

Now let us try to visualize $\mathbb{R}P^2$. Here, we can project everything onto the plane $z = 1$. Here, every line is going to meet this plane exactly at one point, except for those line that are parallel to the plane $z = 1$. These are precisely the lines lying on the xy-plane. But if you think about it, the lines passing through the origin on the xy-plane just looks like $\mathbb{R}P^1$! So

$$\mathbb{R}P^2 = \mathbb{R}^2 \cup \mathbb{R}P^1.$$

So you can think of $\mathbb{R}P^2$ as a "hat" attached to $\mathbb{R}P^1$. But this is not just disc attached to a circle. The disc should be attached to $\mathbb{R}P^1$, which is a circle wrapped around itself twice. So $\mathbb{R}P^2$ is more like a "hat" attached to a Möbius band. Like as for $\mathbb{R}P^1$, we can think of $\mathbb{R}P^2$ as a sphere S^2 with two opposite points glued together.

14 February 23, 2018

We set up the axiom for projective space. One example we gave was $\mathbb{R}P^2$. In this example, a point in $\mathbb{R}P^2$ was a line in $\mathbb{R}^3$ through the origin. To visualize this space, we projected $\mathbb{R}^3$ to some $\mathbb{R}^2$. Using this, we were able to recognize the space as $\mathbb{R}^2$ with a $\mathbb{R}P^1$ attached around the infinity. This "line of infinity" is like the horizon. Today we are going to look at a third way of thinking about projective space, by homogeneous coordinates. By before talking about homogeneous coordinates, we need to see how we identify stuff with other stuff.

14.1 Equivalence classes

Definition 14.1. Let S be a set. A **partition** is a collection of subsets sets $\{S_\alpha \subseteq S\}$ such that $S_\alpha \cap S_\beta = \emptyset$ for $\alpha \neq \beta$ and $\bigcup_\alpha S_\alpha = S$.

Example 14.2. Consider $S = \{A, B, C, D, E, F, G, H\}$. The sets

$$S_1 = \{A, C, G\}, \quad S_2 = \{B, H\}, \quad S_3 = \{D, F\}, \quad S_4 = \{E\}$$

form a partition of S.

Here is another way to encode the same data. We can think of this partition as a map

$$f : S \to T = \{S_1, S_2, S_3, S_4\}, \quad A, C, G \mapsto S_1, \quad B, H \mapsto S_2, \quad D, F \mapsto S_3, \quad E \mapsto S_4.$$

From this map f, we can recover the partition by taking the inverse image of elements. For instance, $\{A, C, G\} = f^{-1}(\{S_1\})$. Formally, we can say that $\{f^{-1}(\{t\}) : t \in T\}$ is a partition of S.

Here is yet another way of making a partition. If $A, C, G \in S_1$, we say that A, C, G are all "equivalent".

Definition 14.3. A **relation** on a set S is a subset $R \subseteq S \times S$. If $(x, y) \in R$, we say that $x \sim y$.

Definition 14.4. An **equivalence relation** on a set S is a relation $R \subseteq S \times S$ such that

(i) $x \sim x$ for all $x \in S$,

(ii) if $x \sim y$ then $y \sim x$,

(iii) if $x \sim y$ and $y \sim z$ then $x \sim z$.

Here, we are thinking of $x \sim y$ to mean that the two objects x and y are equivalent. Given a partition, we can set up an equivalence relation that says "x and y are equivalent if they are in the same partition", or formally, $x \sim y$ if and only if $x, y \in S_\alpha$ for some α. The converse is also possible.

Definition 14.5. Let S be a set with an equivalence relation. We say that a (nonempty) $S' \subseteq S$ is an **equivalence class** if for each $x \in S'$, we have $y \in S'$ if and only if $y \sim x$.

These S' will give us a partition of S. Each $x \in S$ will be in a unique equivalence class, so we write this equivalence class by $[x]$. Then

$$[x] = \{s \in S : x \sim s\}.$$

It is possible that $[x] = [y]$ if $x \sim y$.

14.2 Homogeneous coordinates

Let us now go back to projective space. We have defined the points of $\mathbb{R}P^3$ (or in general $\mathbb{R}P^n$) to be a line passing through the origin in $\mathbb{R}^4$ (or $\mathbb{R}^{n+1}$). A line looks like

$$\{\lambda\vec{u} = \lambda(x, y, z, w) : \lambda \in \mathbb{R}\}.$$

So if we take something like $3(x, y, z, w)$ instead of (x, y, z, w) the corresponding line

$$\{\lambda(3(x, y, z, w)) : \lambda \in \mathbb{R}\}$$

is the same line. Therefore we can identify $\vec{x} \sim \lambda\vec{x}$ for $\lambda \in \mathbb{R} - \{0\}$. That is, under this equivalence class we can regard

$$[\vec{x}] = \{\lambda\vec{x} : \lambda \in \mathbb{R} - \{0\}\}$$

as a line through the origin in $\mathbb{R}^4$, or equivalently as a point in $\mathbb{R}P^3$.

So for $\vec{a} = (a_0, a_1, a_2, a_3) \in \mathbb{R}^4 - \{0\}$, we define the equivalence class under multiplication

$$[\vec{a}] = [a_0, a_1, a_2, a_3]$$

as points in $\mathbb{R}P^3$. Because these are equivalence classes, we have $[1, 1, 1, 1] = [2, 2, 2, 2]$ in $\mathbb{R}P^3$ for example. This is called the **homogeneous coordinates**.

Last time, we identified a line in $\mathbb{R}^3$ with its intersection with the plane $z = 1$. Then a line $t(a_0, a_1, a_2)$ is identified with the point

$$\left(\frac{a_0}{a_2}, \frac{a_1}{a_2}, 1\right).$$

This indeed gives the same point, because

$$\left[\frac{a_0}{a_2}, \frac{a_1}{a_2}, 1\right] = [a_0, a_1, a_2]$$

if $a_2 \neq 0$. If $a_2 = 0$, then either a_0 or a_1 should be nonzero, so $[a_0, a_1]$ can be thought of as in $\mathbb{R}P^1$. This again gives a decomposition of $\mathbb{R}P^2$ into $\mathbb{R}^2$ and $\mathbb{R}P^1$.

On other remark I want to make is that for $[a_0, a_1, a_2, a_3] \in \mathbb{R}P^3$, there is always some i such that $a_i \neq 0$. So this point can be written as

$$[1, a_1, a_2, a_3] \text{ or } [a_0, 1, a_2, a_3] \text{ or } [a_0, a_1, 1, a_3] \text{ or } [a_0, a_1, a_2, 1].$$

Therefore $\mathbb{R}P^3$ is covered by 4 copies of $\mathbb{R}^3$.

14.3 Lines in the projective plane

So we now understand a fair amount about points in $\mathbb{R}P^2$. But what about lines? We had this axiom that any two points can be connected by a unique line. For instance, if $P, Q, R \in \mathbb{R}P^2$ lies on a line in some projection to $\mathbb{R}^2$, do they also lie on a line in any other projection to $\mathbb{R}^2$? Let me ask a more general question.

Question. *If three (non-origin) points in $\mathbb{R}^3$ are on a line, then do their projections (onto $z = 1$) also lie on a line?*

The answer is yes, if you ignore the technical issue when the three points lie on a line passing through the origin, and so that the three points have the same image. Let $\vec{a} + t\vec{b}$ be a line passing through the three points, and assume that this line does not pass through the origin. This means that $\vec{a}, \vec{b} \neq 0$ and $\vec{a}$ and $\vec{b}$ are not parallel. Then if we projective this to $z = 1$, it will become

$$\Big[\frac{a_1 + tb_1}{a_3 + tb_3}, \frac{a_2 + tb_2}{a_3 + tb_3}, 1\Big].$$

We will show next time the curve this cuts on on $z = 1$ is really a line.

15 February 26, 2018

A point is projective space $\mathbb{P}$ should be thought of as a line passing through 0 in $\mathbb{F}^{n+1}$, where $\mathbb{F}$ is a field in general. We looked at three ways of thinking about projective space:

- the axiomatic approach,
- projecting to a line/plane/hyperplane $\mathbb{F}^n$,
- homogeneous coordinates $[a_0, \dots, a_n]$.

If $a_0 \neq 0$, we can write

$$[a_0, \dots, a_n] = \left[1, \frac{a_1}{a_0}, \dots, \frac{a_n}{a_0}\right]$$

and so this can be considered as a chart on $\mathbb{R}P^n$. Roughly, we are say that we are covering $\mathbb{R}P^n$ by some number of copies of $\mathbb{R}^n$.

We were talking last time about lines on $\mathbb{R}P^2$. The easiest way to define lines in $\mathbb{R}P^2$ was first projecting to the plane $z = 1$ and then declare that lines on the $z = 1$ plane correspond to lines in $\mathbb{R}P^2$. But this is not very satisfying because we have arbitrarily chosen a plane $z = 1$. So we can ask the following question.

Question. *If $\mathcal{L}$ is a line in $\mathbb{R}^3$, and we project $\mathcal{L}$ to the plane $z = 1$, is it still a line?*

Proposition 15.1. *If $\mathcal{L}$ does not pass through the origin, and does not lie on the xy-plane, then the projection of $\mathcal{L}$ onto $z = 1$ is also a line.*

Proof. We said that we can represent $\mathcal{L}$ by $\vec{a} + t\vec{b}$. By the condition that $\mathcal{L}$ does not pass through the origin, we have that $\vec{a}, \vec{b} \neq 0$ and $\vec{a}$ and $\vec{b}$ are not parallel. Now if we project the point $\vec{a} + t\vec{b}$ to $z = 1$, we get

$$\Big(\frac{a_1 + tb_1}{a_3 + tb_3}, \frac{a_2 + tb_2}{a_3 + tb_3}\Big).\Big(\frac{a_1 + tb_1}{a_3 + tb_3}, \frac{a_2 + tb_2}{a_3 + tb_3}\Big).$$

The claim is that this is a line. This is because we can write

$$\Big(\frac{a_1 + tb_1}{a_3 + tb_3}, \frac{a_2 + tb_2}{a_3 + tb_3}\Big) = \Big(\frac{a_1}{a_3}, \frac{b_1}{b_3}\Big) + \frac{t}{a_3 + tb_3}\Big(\frac{a_3b_1 - a_1b_3}{a_3}, \frac{a_3b_2 - a_2b_3}{a_3}\Big).$$

Here, $(a_3b_1 - a_1b_3, a_3b_2 - a_2b_3)$ is nonzero because $\mathcal{L}$ does not lie on the xy-plane. $\square$

15.1 Linear fractional transformation

But here, note that the projection operation from $\mathcal{L}$ to $z = 1$ is not linear. You can see this from the $\mathbb{R}P^1$ picture as well. If you project the line $y = 1$ to $x = 1$, there is a lot of distortion. But still, such a transformation always take the form of

$$t \mapsto \frac{\alpha t + \beta}{\gamma t + \delta}.$$

Such a transform is called a linear fractional transform. We will see later that such transformations preserve the cross ratio.

But first, let us check that projections actually are linear fractional transforms. Suppose we have two lines

$$\vec{a} + t\vec{b}, \quad \vec{c} + t\vec{d},$$

not passing through the origin. Then $\vec{a}$ and $\vec{b}$ form a basis of $\mathbb{R}^2$, and likewise $\vec{c}$ and $\vec{d}$ form a basis of $\mathbb{R}^2$. So we can write

$$\vec{a} = \delta\vec{c} + \beta\vec{d}, \quad \vec{b} = \gamma\vec{c} + \alpha\vec{d},$$

where $\alpha\delta - \beta\gamma \neq 0$.

Now projecting the point $\vec{a} + t\vec{b}$ to the line $\{\vec{c} + t\vec{d}\}$ is the same as finding r and s such that

$$\vec{c} + s\vec{d} = r(\vec{a} + t\vec{b}).$$

If we plug the basis representation of $\vec{a}$ and $\vec{b}$ in tersm of $\vec{c}$ and $\vec{d}$, we get

$$\vec{c} + s\vec{d} = r(\delta\vec{c} + \beta\vec{d} + t(\gamma\vec{c} + \alpha\vec{d})),$$

and this implies that

$$1 = r(\delta + t\gamma), \quad s = r(\beta + t\alpha).$$

This shows that

$$s = \frac{s}{1} = \frac{\alpha t + \beta}{\gamma t + \delta}.$$

Definition 15.2. A function is called a **linear fractional transformation** if it is of the form

$$f(x) = \frac{ax + b}{cx + d}$$

for $ad - bc \neq 0$.

Well, generally we don't like 0 being in the denominator. So for $x = -\frac{d}{c}$, what is happens to this function? But in $\mathbb{R}P^1 = \mathbb{R} \cup \{\infty\}$, infinity is an actual point. So we can think of f as a function $\mathbb{R} \to \mathbb{R}P^1$. Even if $x = \infty$, we can define $f(\infty) = \frac{a}{c}$, and then we can think of f as a function $\mathbb{R}P^1 \to \mathbb{R}P^1$. The condition $ad - bc \neq 0$ just makes sure that f is not a constant.

Proposition 15.3. *Any linear fractional transformation is obtained by composing*

$$x \mapsto x + \alpha, \quad x + \alpha x, \quad x \mapsto \frac{1}{x}.$$

Proof. Suppose we want to decompose

$$x \mapsto \frac{ax+b}{cx+d}$$

into the simple linear fractional transformations. We can do this by

$$x \mapsto cx \mapsto cx+d \mapsto \frac{1}{cx+d} \mapsto \frac{1}{cx+d}\Big(\frac{bc-ad}{c}\Big)$$
$$\mapsto \frac{1}{cx+d}\Big(\frac{bc-ad}{c}\Big) + \frac{a}{c} = \frac{ax+b}{cx+d}.$$

□

Proposition 15.4. *Every linear fractional transformation has an inverse, and has the form of*

$$f^{-1}(x) = \tfrac{dx-b}{-cx+a}.$$

Proof. You can directly check that they are mutually inverse. Or we can check this at three points. We have

$$f : -\frac{d}{c} \mapsto \infty, \quad \infty \mapsto \frac{a}{c}, \quad -\frac{b}{a} \mapsto 0,$$

and

$$f^{-1} : \infty \mapsto -\frac{d}{c}, \quad \frac{a}{c} \mapsto \infty, \quad 0 \mapsto -\frac{b}{a}.$$

So they are inverses. □

16 February 28, 2018

We were talking about linear fractional transformations, those functions of the form

$$f : \mathbb{R}P^1 \to \mathbb{R}P^1; \quad f(x) = \frac{ax+b}{cx+d}; \quad ad - bc \neq 0.$$

Here, we think $\mathbb{R}P^1 = \mathbb{R} \cup \{\infty\}$ and we say that $f(-\frac{d}{c}) = \infty$. We showed that every linear fractional transformation can be written as a composition of

$$x \mapsto x + \alpha, \quad x \mapsto \alpha x, \quad x \mapsto \frac{1}{x}.$$

We also showed that there is an inverse

$$f^{-1}(x) = \frac{dx - b}{-cx + a}.$$

16.1 More properties of linear fractional transformations

Proposition 16.1. *Composites of linear fractional transformations again a linear fractional transformation.*

Proof. It is enough to check that the composite of two linear fractional transformations is a linear fractional transformation. Let us do this. Let

$$f(x) = \frac{ax+b}{cx+d}, \quad g(x) = \frac{a'x+b'}{c'x+d'}.$$

Then

$$f(g(x)) = \frac{a\frac{a'x+b'}{c'x+d'} + b}{c\frac{a'x+b'}{c'x+d'} + d} = \frac{(aa'+bc')x + (ab'+bd')}{(ca'+dc')x + (db'+dd')}.$$

So this is a linear fractional transformation. (You can check that $\alpha\delta - \beta\gamma \neq 0$.) $\square$

Proposition 16.2. *Any linear fractional transformation is determined uniquely by where it sends any three points. In particular, if $f(p_i) = g(p_i)$ for $i = 1, 2, 3$ and p_1, p_2, p_3 are distinct, then $f = g$.*

Why I talk about uniqueness here, I am really considering

$$f(x) = \frac{ax+b}{cx+d} \text{ and } \tilde{f}(x) = \frac{(ta)x + (tb)}{(tc)x + (td)}$$

as the same fractional linear transformations.

Proof. Let us show that given $p, q, r \in \mathbb{R}P^1$ and $p', q', r' \in \mathbb{R}P^1$ then there exists a linear fractional transformation $f : \mathbb{R}P^1 \to \mathbb{R}P^1$ such that $f(p) = p'$ and $f(q) = q'$ and $f(r) = r'$. Because we can compose linear transformations, and take inverses, we may assume that $p' = 0$, $q' = 1$, $r' = \infty$. (Then in the general case, we can do something like find f with $a \mapsto 0$, $b \mapsto 1$, $c \mapsto \infty$, find g

with $a' \mapsto 0$, $b' \mapsto 1$, $c' \mapsto \infty$ and then take $g^{-1} \circ f$.) Because we want $f(p) = 0$ and $f(r) = \infty$, we want f to take the form of

$$f(x) = \frac{a(x-p)}{c(x-r)}.$$

Then for $f(q) = 1$, we can set

$$f(x) = \frac{(q-r)(x-p)}{(q-p)(x-r)}.$$

Note that this also shows uniqueness, because the coefficients a, b, c, d were completely determined up to scalar multiplication. □

16.2 Cross ratio

There are some similarities between linear fractional transformations and isometries. Isometries on $\mathbb{R}^2$ preserves distance $d(x, y)$ for $x, y \in \mathbb{R}^2$. Linear fractional transformations preserve the cross ratio.

Definition 16.3. For $p, q, r, s \in \mathbb{R}P^1$, we define the **cross ratio** as

$$[p, q; r, s] = \frac{(r-p)(s-q)}{(s-p)(r-q)}.$$

Proposition 16.4. *Linear fractional transformations preserve the cross ratio.*

Proof. You only need to check this for $x \mapsto x + \alpha$, $x \mapsto \alpha x$, and $x \mapsto \frac{1}{x}$, because all linear fractional transformations are built out of these three. □

You can think of this as measuring "distance" with respect to three points. Given any three points, then the value of the cross ratio is uniquely determined by fourth point on $\mathbb{R}P^1$. We can check this directly. The equation

$$\frac{(r-p)(x-q)}{(x-p)(r-q)} = y$$

is equivalent to

$$x = \frac{q(r-p) + yp(r-q)}{(r-p) - y(r-q)}.$$

Proposition 16.5. *An injective $f : \mathbb{R}P^1 \to \mathbb{R}P^1$ that preserves all cross ratios is a linear fractional transformation.*

Proof. First we show that f is surjective. Take any $u, v, w \in \mathbb{R}P^1$. For an arbitrary $x \in \mathbb{R}P^1$ not equal to $f(u), f(v), f(w)$, we consider

$$y = [f(u), f(v); f(w), x]$$

and find a point s such that $y = [u, v; w, s]$. (This can be done by the previous discussion.) Then we should have

$$[f(u), f(v); f(w), x] = y = [u, v; w, s] = [f(u), f(v); f(w), f(s)].$$

Then $f(s) = x$, and so x is in the image. Because we can do this for all x, the function f is surjective.

Now take p, q, r to be the points that f sends to $0, 1, \infty$. Then for any $x \in \mathbb{R}P^1$ not equal to p, q, r, we have

$$[r, p; q, x] = [f(r), f(p); f(q), f(x)] = [\infty, 0; 1, f(x)] = f(x).$$

This shows that f is a linear fractional transformation. □

17 March 2, 2018

A linear fractional transformation is one that looks like $f(x) = \frac{ax+b}{cx+d}$ for $ad-bc \neq 0$, and this preserves the cross ratio

$$[p, q; r, s] = \frac{(r-p)(s-q)}{(s-p)(r-q)}.$$

17.1 Characterization of cross ratio as the unique invariant

They have the following properties:

(1) A linear fractional transformation is determined by its values on any 3 points.

(2) A linear fractional transformation preserves the cross ratio.

(3) Given three points of an arbitrary value of a cross ratio, there exists exactly one point which gives the value of the cross ratio.

(4) Linear factional transformations form a group under composition. That is,

  (a) composites of linear fractional transformations are linear fractional transformations,

  (b) the identity map is a linear fractional transformation,

  (c) a linear fractional transformation is bijective (as a map $\mathbb{R}P^1 \to \mathbb{R}P^1$), and its inverse is again a linear fractional transformation.

(5) Given three distinct points p, q, r and three distinct p', q', r', there exists a unique linear fractional transformation that takes $p \mapsto p'$, $q \mapsto q'$, $r \mapsto r'$. (This contains (1).)

(6) Any invariant of 4 points is a function of the cross ratio. That is, if there is a function $I(p, q, r, s)$ such that

$$I(p, q, r, s) = I(f(p), f(q), f(r), f(s))$$

for all points p, q, r, s and linear fractional transformations f, then I should be of the form

$$I(p, q, r, s) = I'([p, q, r, s]).$$

Let us prove the last property.

Proof. For such an invariant I, let us define

$$I'(x) = I(\infty, 0; 1, x).$$

Then for any four points p, q, r, s, there exists a fractional linear transformation f such that $f(p) = \infty$, $f(q) = 0$, $f(r) = 1$. Then

$$[p, q; r, s] = [f(p), f(q); f(r), f(s)] = [\infty, 0; 1, f(s)] = f(s).$$

Then

$$I(p,q,r,s) = I(f(p), f(q), f(r), f(s)) = I(\infty, 0, 1, f(s)) = I'(f(s)) = I'([p,q;r,s]).$$

$\square$

17.2 Algebra on the projective plane

On the Euclidean plane, we had the straightedge and the compass. On the projective plane, we have the cross ratio, and we can do algebra with this. We are going to work over a **field**, a set with addition and multiplication having the properties

(a) (commutative) $a + b = b + a$, $ab = ba$,

(b) (associative) $a + (b + c) = (a + b) + c$, $(ab)c = a(bc)$,

(c) (identity) $0 + a = a + 0 = a$, $a1 = 1a = a$,

(d) (inverse) $a + (-a) = (-a) + a = 0$, $aa^{-1} = a^{-1}a = 1$,

(e) (distributive) $a(b + c) = ab + ac$.

We are going to look at Pappus's and Desargues's theorems next time.

18 March 5, 2018

The problem set is now due March 9, Friday. We finished talking about projective space in general, and now we are going to focus on projective planes. Here, we will learn how to build a field out of a projective plane.

18.1 Pappus's theorem

Recall the vector Pappus's theorem.

Theorem 18.1. *If* $\mathbf{r}, \mathbf{s}, \mathbf{t}, \mathbf{u}, \mathbf{v}, \mathbf{w}$ *lie alternatively on two lines passing through* 0*, the vectors* $\mathbf{r} - \mathbf{s}$ *and* $\mathbf{v} - \mathbf{u}$ *are parallel, and* $\mathbf{v} - \mathbf{w}$ *and* $\mathbf{t} - \mathbf{s}$ *are parallel, then* $\mathbf{r} - \mathbf{w}$ *and* $\mathbf{t} - \mathbf{u}$ *are parallel.*

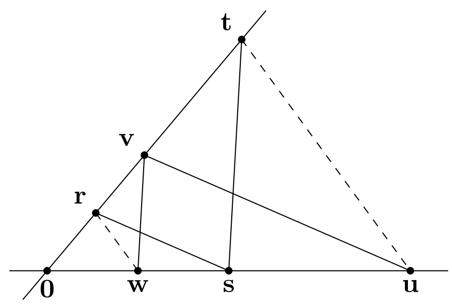

Figure 29: Vector Pappus's theorem

Here, note that parallel lines are lines that meet at the horizon, in the projective plane. So we are going to consider the following analogue of the Pappus's theorem.

Theorem 18.2 (projective Pappus's theorem)**.** *Consider six points, lying alternatively on two straight lines. Then the three points formed by intersecting two opposite sides of the hexagon lie on a single line.*

Proof. We just consider the line passing through S and U as the horizon. Then the theorem just reduces to the vector version of Pappus's theorem. (This is because lines meeting at the horizon are actually parallel.) □

18.2 Desargues's theorem

Likewise, we can generalize the vector version of Desargues's theorem to the projective version. Recall:

Theorem 18.3 (Desargues's theorem)**.** *Suppose* A, B, C *and* A', B', C' *lie on concurrent lines* $\mathcal{L}, \mathcal{M}, \mathcal{N}$*. (Equivalently, we say that* $\triangle ABC$ *and* $\triangle A'B'C'$ *are* ***perspective****.) If* AB *is parallel to* $A'B'$ *and* BC *is parallel to* $B'C'$*, then* AC *is parallel to* $A'C'$*.*

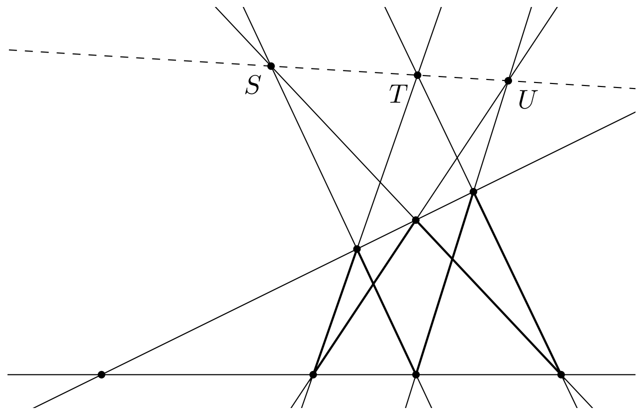

Figure 30: Projective Pappus's theorem

Here is the projective version.

Theorem 18.4 (projective Desargues's theorem)**.** *If two triangles are perspective from a point, then three intersections coming from three pairs of corresponding sides lie on a line.*

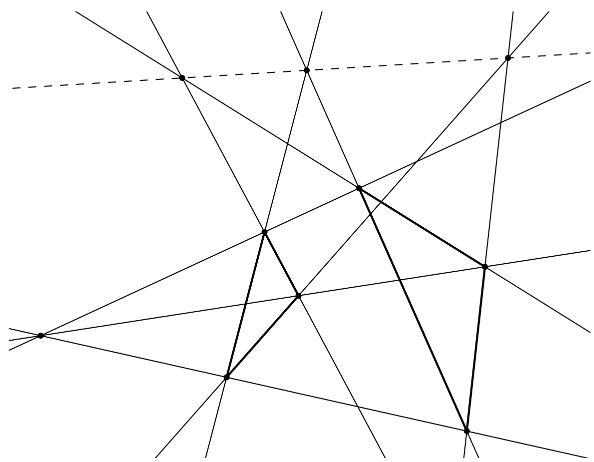

Figure 31: Projective Desargues's theorem

Proof. The proof is the same. If we send this line to the horizon, then this just reduces the vector (or classical) version of Desargues's theorem. □

As I have explained, our goal is to build a multiplication system from this projective plane $\mathbb{R}P^2$. We will make use of these two theorems (Pappus and Desargues) to define the operations. There are projective planes where the Pappus and Desargues theorems are not true, such as the Moulton plane.

Theorem 18.5 (little Desargues's theorem). *If two triangles are in perspective from a point P, and if the two pairs of corresponding sides meet on a line $\mathcal{L}$ through P, then the third pair of corresponding sides also meet on $\mathcal{L}$.*

Theorem 18.6 (converse of Desargues's theorem). *If the corresponding sides of two triangles meet on a line, then the triangles are perspective from a point.*

Theorem 18.7 (scissors theorem). *If $ABCD$ and $A'B'C'D'$ are quadrilaterals with vertices alternatively on two lines, AB is parallel to $A'B'$, BC is parallel to $B'C'$, AD is parallel to $A'D'$, then CD is parallel to $C'D'$.*

Next time, we will construct addition and multiply, and we will use these theorems to verify the field axioms.

19 March 7, 2018

Desargue's theorem states that given two triangles in perspective, the corresponding sides meet on a line. We now want to prove the converse:

Theorem 19.1 (converse Desargues's theorem)**.** *If corresponding sides of two triangles meet on a line, then the two triangles are in perspective.*

Proof. Suppose two triangles ABC and $A'B'C'$ are the two triangles whose corresponding sides meet on the line ℓ. Let P be the intersection of AA' and BB'. (We don't know yet that CC' also passes through P.) Let C'' be the intersection of PC and $B'C'$. Then $\triangle ABC$ and $\triangle A'B'C''$ are in perspective, so

$$AB \cap A'B', \quad BC \cap (B'C'' = B'C'), \quad AC \cap A'C''$$

lie on the same line, by Desargues. But the first two points lie on ℓ, so the last point $AC \cap A'C''$ is the intersection of ℓ and AC. But by assumption, $AC \cap A'C'$ is the intersection of ℓ and AC. Therefore $A'C' = A'C''$ and thus $C' = C''$. □

Theorem 19.2 (scissors theorem)**.** *Suppose A, C, A', C' lie on a line and B, D, B', D' also lie on a line. If $AB \parallel A'B'$ and $BC \parallel B'C'$ and $BC \parallel B'C'$ then $CD \parallel C'D'$.*

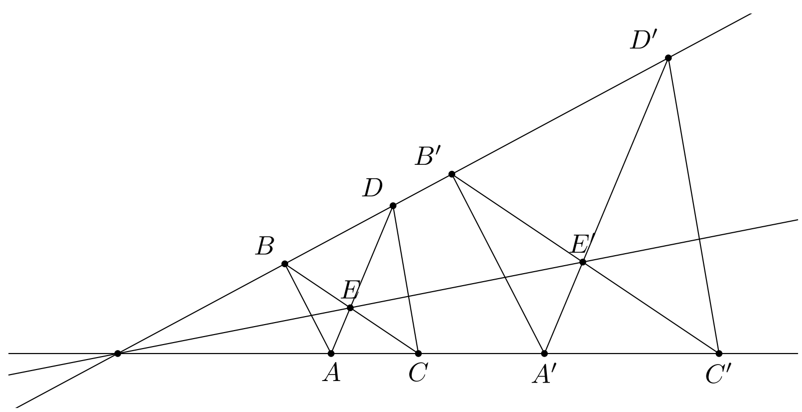

Figure 32: Scissors theorem

Proof. Let $E = AD \cap BC$ and $E' = A'D' \cap B'C'$. Note that the corresponding sides of $\triangle ABE$ and $\triangle A'B'E'$ meet on the horizon, which is a line. So by converse Desargues, we see that the two triangles are in perspective. That is,

$$AA', \quad BB', \quad EE'$$

meet at a point. So $\triangle CDE$ and $\triangle C'D'E'$ are in perspective. By Desargues, the corresponding sides should meet on a line, but CE and $C'E'$ meet on the horizon and DE and $D'E'$ meet on the horizon as well. Therefore CD and $C'D'$ should meet on the horizon as well, which means that they are parallel. □

19.1 Recovering the field

Definition 19.3. A **field** is a set S along with two operations $+, \cdot : S \times S \to S$ with two special elements $0, 1 \in S$ satisfying a bunch of properties

(1) $a + b = b + a$ and $a \cdot b = b \cdot a$ for all $a, b \in S$,

(2) $a + (b + c) = (a + b) + c$ and $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ for all $a, b, c \in S$,

(3) $a + 0 = a$ and $a \cdot 1 = a$ for all $a \in S$,

(4) for each $a \in S$ there is an $(-a) \in S$ such that $a + (-a) = 0$,

(5) for each nonzero $a \in S$ there is an $a^{-1} \in S$ such that $a \cdot a^{-1} = 1$,

(6) $a \cdot (b + c) = a \cdot b + a \cdot c$.

The goal is to define, for two points in $\mathbb{R}P^2$, a field structure on the line joining the two points (excluding the horizon) such that the two points correspond to 0 and 1.

Let's first start with addition: given two points a and b, we can define $a + b$ in the following way.

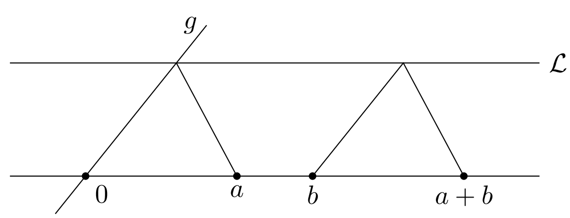

Figure 33: Addition in $\mathbb{R}P^2$

In defining $a + b$, we made arbitrary choices of $\mathcal{L}$ and g. So in order for this definition to be well-defined, we need this to be independent of the these two choices. For $a + b$ to be independent of $\mathcal{L}$, we can choose two lines $\mathcal{L}$ and $\mathcal{L}'$ and show that they give the same point $a + b$.

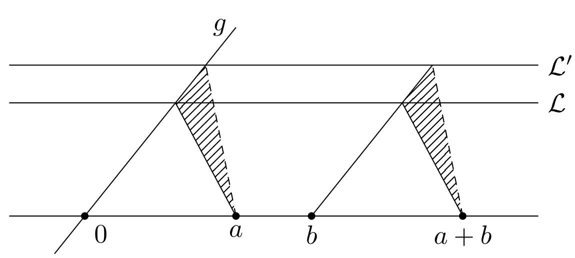

Figure 34: Independence of $\mathcal{L}$

Here, if we use converse Desargues's theorem on the two shaded triangles, we show that the other two dashed lines are parallel. We also have to check that $a + b$ is independent of the choice of g, and this can be checked using the scissor theorem.

Let us now define multiplicaiton. Here, we draw an arbitrary line g through 0, and pick a point p on g. Draw a line from p to 1 and to a, and copy this picture to b on a parallel manner.

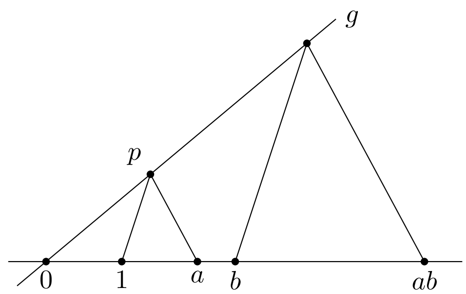

Figure 35: Multiplication in $\mathbb{R}P^2$

Again, we need to check that this is independent of the choice of p and g. If we move p, we

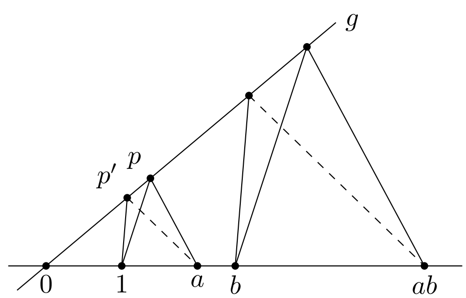

Figure 36: Independence of p

Here, we need to show that the dashed lines are parallel, and this is just the scissors theorem. If we want to prove that multiplication is independent of the choice of g, we will be using Desaruges's theorem. (You can check this by yourself.)

20 March 9, 2018

We can check all the field axioms by similar arguments, using Pappus's and Desargues's theorems.

20.1 Groups

Recall that an isometry on $\mathbb{R}^2$ is a map $f : \mathbb{R}^2 \to \mathbb{R}^2$ preserving distances. We have seen that

- compositions of isometries are isometries,
- reflections (by some line) are isometries,
- every isometry can be written as at most three reflections.

Proposition 20.1. *Any isometry f has an inverse f^{-1}, which is again an isometry.*

Proof. If we can write $f = r_1 r_2 r_3$ for r_1, r_2, r_3 reflections, for instance, we have $f^{-1} = r_3 r_2 r_1$ because

$$f \circ f^{-1} = (r_1 r_2 r_3)(r_3 r_2 r_1) = r_1 r_2 r_2 r_1 = r_1 r_1 = \mathrm{id}$$

Similarly, we can check $f^{-1} f = \mathrm{id}$. □

We can collect these properties formally and define a group.

Definition 20.2. A **group** is a set G with an operation $\cdot : G \times G \to G$ with a distinguished element $1 \in G$ satisfying

- $1 \cdot g = g \cdot 1 = g$ for all $g \in G$,
- for all $g \in G$, there exists an element $g^{-1} \in G$ such that $g \cdot g^{-1} = g^{-1} \cdot g = 1$,
- $g_1(g_2 g_3) = (g_1 g_2) g_3$.

Note that $g_1 g_2$ need not be equal to $g_2 g_1$. You should think of these as some collection of "functions preserving some structure".

Example 20.3. Isometries of $\mathbb{R}^2$, denoted by $\mathrm{Isom}(\mathbb{R}^2)$, form a group under composition.

Definition 20.4. A group G is called **abelian** if it satisfies $g_1 g_2 = g_2 g_1$ for all $g_1, g_2 \in G$.

Example 20.5. The isometry group $\mathrm{Isom}(\mathbb{R}^2)$ is *not* abelian. If we take g_1 to be counterclockwise rotation by $90°$ about 0, and g_2 to be reflection about the x-axis, then you can check $g_1 g_2 \neq g_2 g_1$ easily.

Lemma 20.6. *For any group G and $g \in G$, the inverse g^{-1} is unique.*

Proof. Suppose we have two inverses $(g^{-1})'$ and g^{-1}, both satisfying

$$g(g^{-1}) = (g^{-1})g = 1, \quad g(g^{-1})' = (g^{-1})'g = 1.$$

Then we have

$$g^{-1} = g^{-1}(g(g^{-1})') = ((g^{-1})g)(g^{-1})' = (g^{-1})'.$$

That is, $g^{-1} = (g^{-1})'$. $\square$

Example 20.7. The set of $n \times n$ matrices $M_n(\mathbb{R})$ forms a group under addition $+$, with the zero matrix $O \in M_n(\mathbb{R})$ being the identity. This group is abelian.

Example 20.8. The set of invertible $n \times n$ matrices $\mathrm{GL}_n(\mathbb{R})$ forms a group under multiplication $\times$, with the identity matrix $I \in \mathrm{GL}_n(\mathbb{R})$ being the identity. This group is not abelian if $n \geq 2$. These are invertible maps from $\mathbb{R}^n$ to $\mathbb{R}^n$ preserving the vector space structure.

Definition 20.9. We say that an element $g \in G$ has **order** m if $g \cdot g \cdot \cdots \cdot g = g^m = 1$ but $g^k \neq 1$ for all $0 < k < m$. So this is the smallest number of times you need to multiply g with itself to get the identity element. If such an m does not exist, we say that it has infinite order.

Example 20.10. Consider $\mathbb{Z}/n\mathbb{Z} = \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}$ with addition being the remainder of addition divided by n, so that $\overline{1} + \overline{n-1} = \overline{0}$ for instance. This is an abelian group, and the order of $\overline{1}$ is equal to n.

Lemma 20.11. *If $g \in G$ has order $m < \infty$, then $1, g, g^2, \ldots, g^{m-1}$ are distinct, and $\{1, g, \ldots, g^{m-1}\}$ form a group under multiplication inherited from G.*

Proof. It is closed under multiplication, and is clearly associative. One tricky thing is that it has inverses, but we check $g^{m-k} = (g^k)^{-1}$ because

$$g^k \cdot g^{m-k} = g^{m-k} \cdot g^k = g^m = 1.$$

Also, the identity is there. $\square$

Example 20.12. Consider the set of function $f : \mathbb{R}^n \to \mathbb{R}^n$ taking the form of

$$f(u) = Au + v,$$

where $A \in \mathrm{GL}_n(\mathbb{R}^n)$ is an invertible matrix and $v \in \mathbb{R}^n$ is an arbitrary vector. These are called **affine transformations** and send lines to lines, although they do not have to fix the origin. The claim is that

$$\{\text{affine transformations } \mathbb{R}^n \to \mathbb{R}^n\}$$

form a group.

20.2 Algebraic structures on $\mathbb{R}^n$

The real numbers $\mathbb{R}$ have structures of addition and multiplication. That is, we can add two things in $\mathbb{R}$, but we can also multiply things in $\mathbb{R}$. This is not true in general for $\mathbb{R}^n$. We can add things in $\mathbb{R}^n$ together (adding vectors together), but there is no obvious way of multiplying vectors. (We can define entrywise multiplication $(a_1, \ldots, a_n) \cdot (b_1, \ldots, b_n) = (a_1 b_1, \ldots, a_n b_n)$ but this is bad because things like $(1, 0) \cdot (0, 1) = (0, 0) = 0$ can happen.)

In the case of $n = 2$ we can identify $\mathbb{R}^2 \cong \mathbb{C}$ and use complex multiplication. Then we are defining

$$(a, b) \cdot (c, d) = (ac - bd, ad + bc).$$

The other cases where we can define a suitable "well-behaved" multiplication on $\mathbb{R}^n$ are $n = 4$ and $n = 8$. For $\mathbb{R}^4$, we call this the **quaternions**. Here,

$$\mathbb{R}^4 \cong \mathbb{H}; \quad (a, b, c, d) \leftrightarrow a + bi + cj + dk.$$

We define multiplication as

$$ij = k = -ji, \quad jk = i = -kj, \quad ki = j = -ik, \quad i^2 = j^2 = k^2 = -1.$$

This multiplication can also be represented by 2×2 complex matrices as

$$1 \leftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad i \leftrightarrow \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}, \quad j \leftrightarrow \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad k \leftrightarrow \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

21 March 19, 2018

In this chapter, we are going to talk about transformations, which are implicitly groups.

21.1 Isometries as a group

Definition 21.1. A **group** G is a set equipped with a operation $\cdot : G \times G \to G$ written as $(g_1, g_2) \mapsto g_1 g_2$ and a unit $1 \in G$ such that

(a) for all $g \in G$, $1 \cdot g = g \cdot 1 = g$,

(b) for all $g \in G$, there exists $f \in G$ such that $gf = fg = 1$, which we call $f = g^{-1}$,

(c) for all $g_1, g_2, g_3 \in G$, $(g_1 g_2) g_3 = g_1 (g_2 g_3)$.

Recall that an isometry is a map $f : \mathbb{R}^2 \to \mathbb{R}^2$ that preserves distance. The composition of any two isometries is going to be an isometry. So if we consider $G = \{\text{isometries of } \mathbb{R}^2\}$, we have a composition map $G \times G \to G$ given by $(g_1, g_2) \mapsto g_1 \circ g_2$. To say that G is a group under composition, we still need to check the three axioms (a)–(c). The hardest is (b), but this can be done using the "three reflections theorem". If f is an isometry, it can be written as the composition of at most three reflections. If $f = r_3 r_2 r_1$ for instance, we have

$$(r_1 r_2 r_3)(r_3 r_2 r_1) = r_1 r_2 r_3 r_3 r_2 r_1 = r_1 r_2 r_2 r_1 = r_1 r_1 = \mathrm{id}\,.$$

Likewise, $(r_3 r_2 r_1)(r_1 r_2 r_3) = \mathrm{id}$. So we can say that $r_1 r_2 r_3 = f^{-1}$.

Proposition 21.2. *The set* $\mathrm{Isom}(\mathbb{R}^2)$ *of isometries of* $\mathbb{R}^2$ *forms a group under composition.*

Recall that the composition of two reflections is a rotation or a translation. The composition of three reflections is a glide reflection.

Definition 21.3. We say that an isometry is **odd** if is the composition of 1 or 3 reflections. We say that an isometry is **even** if it is the composition of 2 reflections. We denote by the set of odd isometries $\mathrm{Isom}^-(\mathbb{R}^2)$ and the set of even isometries $\mathrm{Isom}^+(\mathbb{R}^2)$.

The set $\mathrm{Isom}^-(\mathbb{R}^2)$ of odd isometries is not a group, because the composition of two odd isometries is even. So we can't define the map $\mathrm{Isom}^-(\mathbb{R}^2) \times \mathrm{Isom}^-(\mathbb{R}^2) \to \mathrm{Isom}^-(\mathbb{R}^2)$ given by composition. On the other hand, the composition of two even isometries is an even isometry, so $\mathrm{Isom}^+(\mathbb{R}^2)$ is a group under composition.

21.2 Linear transformations

This can be defined generally on vector spaces.

Definition 21.4. A function $f : \mathbb{R}^n \to \mathbb{R}^n$ is called a **linear transformation** if it satisfies

$$f(\alpha\mathbf{u} + \mathbf{v}) = \alpha f(\mathbf{u}) + f(\mathbf{v})$$

for all $\alpha \in \mathbb{R}$ and $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$.

Linear transformations always maps lines to lines (or points). To see this, consider a line $\mathcal{L}$ defined by $\mathbf{a} + t\mathbf{u}$ for $t \in \mathbb{R}$. Then by the definition of a linear transformation,

$$f(\mathbf{a} + t\mathbf{u}) = f(\mathbf{a}) + tf(\mathbf{u}).$$

Because $f(\mathbf{a})$ and $f(\mathbf{u})$ are still vectors in $\mathbb{R}^n$, this is going to be a line.

A linear transformation is the same thing as a matrix. (Let's look at $n = 2$ for simplicity.) Given a linear transformation $f : \mathbb{R}^2 \to \mathbb{R}^2$, we can get the 2×2 matrix by defining

$$f((1, 0)) = (a, c), \quad f((0, 1)) = (b, d).$$

Then the matrix

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is the matrix representing the linear transformation f, because

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

Is the set $\{2 \times 2 \text{ matrices}\}$ a group under matrix multiplication? It has the unit $I = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. But it does not necessarily have an inverse. To make sure that there is an inverse matrix, we should take

$$\mathrm{GL}_n(\mathbb{R}) = \{n \times n \text{ matrices } M \text{ (having real entries) with } \det M \neq 0\}.$$

This then going to be a group under matrix multiplication.

This is when we take multiplication as the operation. But if we take addition as the operation,

$$M_n(\mathbb{R}) = \{n \times n \text{ matrices (having real entries)}\}$$

is a group. Here, the unit is the zero matrix.

We have talked about isometries before, and those isometries sending the origin to itself are linear transformations. So they can be written down as a 2×2 matrix. Rotation by the angle θ can be written as

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

for $\theta \in \mathbb{R}$. Also, reflection about the x-axis can be written out as

$$X = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

because it sends $(1,0) \mapsto (1,0)$ and $(0,1) \mapsto (0,-1)$. What about reflection about a general line ℓ (passing through the origin)? If the angle between ℓ and the x-axis is θ, we can consider this reflection about ℓ as (1) rotating the whole thing by $-\theta$, (2) reflecting about the x-axis, and (3) rotating by θ again to put things in the right position. So this matrix is

$$R_\theta X R_{-\theta} = \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}.$$

21.3 Linear fractional transformations revisited

Linear fractional transformations on $\mathbb{R}P^1$ are functions that look like

$$f(x) = \frac{ax+b}{cx+d}$$

with $ad-bc \neq 0$. It turns out that this is very much related to the matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$.

To see this, we recall that we defined $\mathbb{R}P^1 = \mathbb{R}^2/\sim$ where $\sim$ is the equivalence relation $\mathbf{x} \sim \lambda\mathbf{x}$. To define coordinates on $\mathbb{R}P^1$, we projected everything onto the line $y = 1$ and only took the x-coordinate. Then the point $s \in \mathbb{R}P^1$ corresponds to the line

$$\mathbb{R}P^1 \ni s \quad \longleftrightarrow \quad x - sy = 0 \in \{\text{lines in } \mathbb{R}^2\}.$$

A linear transformation $\mathbb{R}^2 \to \mathbb{R}^2$ sends a line through the origin to a line through the origin. The line through $(s,1)$ will then be sent to $(as+b, cs+d)$:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} s \\ 1 \end{pmatrix} = \begin{pmatrix} as+b \\ cs+d \end{pmatrix}.$$

Then this new line is going to correspond to the point $\frac{as+b}{cs+d}$. That is, this linear transformation sends $s \mapsto \frac{as+b}{cs+d}$.

Here, note that if we use the matrix

$$\begin{pmatrix} ak & bk \\ ck & dk \end{pmatrix} \text{ instead of } \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

we are going to get the same linear fractional transformation $\mathbb{R}P^1 \to \mathbb{R}P^1$. So the set of linear fractional transformations is really

$$\mathrm{PGL}_2(\mathbb{R}) = \mathrm{GL}_2(\mathbb{R})/\sim$$

where $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \sim k\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. This is a group under composition.

22 March 21, 2018

A group G is a set with an operation $\cdot : G \times G \to G$ and an element $1 \in G$ such that (i) $g \cdot 1 = 1 \cdot g = g$, (ii) for each g there exists an $f = g^{-1}$ such that $gf = fg = 1$, (iii) $(g_1 g_2) g_3 = g_1 (g_2 g_3)$.

Example 22.1. We saw that isometries of $\mathbb{R}^2$ form a group under composition. We also saw that invertible linear transformations of a vector space form a group under composition. This is the group $\mathrm{GL}_n(\mathbb{R})$ of invertible $n \times n$ matrices under multiplication. The set of matrices $M_n(\mathbb{R})$ of all $n \times n$ matrices, this is a group under addition. We also defined the projective general linear group

$$\mathrm{PGL}_2(\mathbb{R}) = \mathrm{GL}_2(\mathbb{R}) / \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \sim \lambda \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right).$$

22.1 More examples of groups

Proposition 22.2. *Every isometry in $\mathbb{R}^2$ is of the form $T(\mathbf{x}) = A\mathbf{x} + \mathbf{v}$ for some $\mathbf{v} \in \mathbb{R}^2$ and A such that $AA^T = I$.*

The difference between this and linear transformations is that we can move the origin around. Linear transformations always send the origin to the origin, but if we put in this $\mathbf{v}$ term, we can move the origin around.

Proposition 22.3. *Conversely, if $AA^T = I$, then $T(\mathbf{x}) = A\mathbf{x} + \mathbf{v}$ is always an isometry.*

Proof. This is an exercise. You can use the formulas $\|\mathbf{x}\|^2 = \langle \mathbf{x}, \mathbf{x} \rangle$ and $\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v}^T \mathbf{w}$ to prove this. □

We can then think about the orthogonal group

$$\mathrm{O}_2(\mathbb{R}) = \{A \in M_2(\mathbb{R}) : AA^T = I\}.$$

We can then think of this as the group of isometries that fixes the origin. As before we can talk about evenness and oddness in $\mathrm{O}_2(\mathbb{R})$. Note that

$$1 = \det(I) = \det(A \cdot A^T) = \det(A)\det(A^T) = \det(A)^2.$$

So $\det(A) = \pm 1$. It turns out that an isometry is even if $\det(A) = 1$ and odd if $\deg(A) = -1$. So we can go and define

$$\mathrm{SO}_2(\mathbb{R}) = \{A \in M_2(\mathbb{R}) : AA^T = I, \deg(A) = 1\}.$$

Note that $T(\mathbf{x}) = A\mathbf{x} + \mathbf{v}$ can be considered as the matrix

$$\begin{pmatrix} a & b & v_1 \\ c & d & v_2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix} = \begin{pmatrix} ax_1 + bx_2 + v_1 \\ cx_1 + dx_2 + v_2 \\ 1 \end{pmatrix}.$$

So we may regard $\mathrm{Isom}(\mathbb{R}^2)$ as setting inside $\mathrm{GL}_3(\mathbb{R})$ (or $\mathrm{PGL}_3(\mathbb{R})$ if you'd like).

22.2 Finite order isometry

Let us now talk about some finite cases.

Definition 22.4. An element $x \in G$ as **order** m if $x^m = 1$ but $x^d \neq 1$ for $0 < d < m$.

For instance, $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)^2 = I$ so this matrix has order 2 in $\mathrm{GL}_2(\mathbb{R})$. If $x \in G$ has order m, then

$$(x) = \{1, x, x^2, \dots, x^{m-1}\}$$

is going to be a group under multiplication as well. This group is going to have size m (it has order m), and it is a subset of G (it is a subgroup of G).

Consider a finite set $S = \{s_1, \dots, s_n\} \subseteq \mathbb{R}^2$ of points on the plane. Let us denote the barycenter as

$$b = \frac{s_1 + s_2 + \cdots + s_n}{n}.$$

If T is an isometry, then the barycenter of $S' = \{Ts_1, \dots, Ts_n\}$ is Tb.

We want to apply this to the case when S is the vertices of a regular n-gon.

Definition 22.5. A **regular n-gon** is an n-gon with all edges having the same lengths and all angles the same.

To construct a regular n-gon, here is what you can do. Consider two distinct points p and v_0. Let $\sigma \in \mathrm{Isom}(\mathbb{R}^2)$ be the isometry that is rotating by $\frac{2\pi}{n}$ around the point p. Then we can define $v_1 = \sigma v_0$, $v_2 = \sigma v_1 = \sigma(\sigma(v_0)) = \sigma^2 v_0$, $\dots$, $v_i = \sigma^i v_0$. Because $\sigma^n = 1$, we see that the points

$$\{v_0, v_1, \dots, v_{n-1}\} = \{v_0, \sigma v_0, \sigma^2 v_0, \dots, \sigma^{n-1} v_0\}$$

form a regular n-gon. This works because σ has order n.

23 March 23, 2018

Last time we talk about various groups of transformations. The group of isometries fixing the origin is

$$\mathrm{O}_2(\mathbb{R}) = \{A \in M_2(\mathbb{R}) : AA^T = I\},$$

and among them, the even isometries are

$$\mathrm{SO}_2(\mathbb{R}) = \{A \in \mathrm{O}_2(\mathbb{R}) : \det A = 1\} = \{A \in M_2(\mathbb{R}) : AA^T = I, \det A = 1\}.$$

Definition 23.1. An element $x \in G$ is of **order** m if $x^m = 1$ and $x^d \neq 1$ for $0 < d < m$.

In this case, the subgroup

$$\langle x \rangle = \{1, x, x^2, \dots, x^{m-1}\}$$

generated by x is a cyclic subgroup of order $m = \mathrm{order}(x)$.

Proposition 23.2. *Every finite subgroup of* $\mathrm{SO}_2(\mathbb{R})$ *is cyclic.*

To construct a regular n-gon, we fixed σ a rotation about the origin by angle $\frac{2\pi}{n}$. This isometry σ has order n, and if v_0 is any point other than the origin,

$$\sigma v_0, \sigma^2 v_0, \dots, \sigma^{n-1} v_0, \sigma^n v_0 = v_0$$

form the vertices of a regular n-gon. Now we want to think about the isometries of this n-gon.

23.1 Dihedral group

What are the isometries of the regular n-gon? First of all, σ is an isometry. This is because if we apply σ to the points

$$v_0, \sigma v_0, \dots, \sigma^{n-1} v_0,$$

we get

$$\sigma v_0, \sigma^2 v_0, \dots, \sigma^{n-1} v_0, \sigma^n v_0 = v_0.$$

So the vertices are preserved. This shows that σ^k are all isometries of the n-gon as well.

But there is another isometry, which is reflection τ about the line connecting the origin and v_0. You can check that this is an isometry, because it sends

$$\tau : v_0 \mapsto v_0, \quad \sigma v_0 \mapsto \sigma^{-1} v_0, \quad \sigma^2 v_0 \mapsto \sigma^{-2} v_0, \dots.$$

In fact, this also shows that

$$\tau \sigma^i = \sigma^{-i} \tau.$$

So we already have a group

$$D_{2n} = \{1, \sigma, \sigma^2, \dots, \sigma^{n-1}, \tau, \sigma\tau, \sigma^2\tau, \dots, \sigma^{n-1}\tau\}$$

of isometries. This is called the **dihedral group**. It is not an abelian group because $\tau\sigma \neq \sigma^{-1}\tau$. Thus, in particular, this is not a cyclic group.

Proposition 23.3. *The isometry group of the regular n-gon is D_{2n}.*

We know that all elements of D_{2n} are indeed isometries of the n-gon. The content of this theorem is that there are no other isometries, in other words, D_{2n} are all the isometries of the regular n-gon.

Proof. Take an arbitrary isometry T. Because T sends $v_0, \sigma v_0, \dots, \sigma^{n-1}v_0$ to that set, it sends the barycenter to itself (you'll check this in your homework). That is $T(0) = 0$. Also, assume that $T(v_0) = \sigma^i v_0$. Then $T' = \sigma^{-i}T$ is a composition of isometries, and so it is an isometry of the regular n-gon as well. Then $T'(0) = 0$ and $T'(v_0) = \sigma^{-i}\sigma^i v_0 = v_0$ now.

Because $T'(0)$ fixes the two points 0 and v_0, it is either the identity map, or reflection about the line connecting 0 and v_0. Then $T' = 1$ or $T' = \tau$. This means that $\sigma^{-i}T = 1$ or $\sigma^{-i}T$, and so $T = \sigma^i$ or $T = \sigma^i\tau$. Therefore T is in D_{2n}. □

23.2 Classification of finite subgroups of $\mathrm{Isom}(\mathbb{R}^2)$

So D_{2n} is a finite subgroup of $\mathrm{Isom}(\mathbb{R}^2)$. But are there any finite subgroups of $\mathrm{Isom}(\mathbb{R}^2)$? Surprisingly, we can find all of them.

Lemma 23.4. *Let $G \subseteq \mathrm{Isom}(\mathbb{R}^2)$ be a finite subgroup. Then there exists a point P fixed by all the transformations of G.*

Proof. Let $v \in \mathbb{R}^2$ be any point. Say that G has n elements and define

$$P = \frac{1}{n}\sum_{g\in G} gv = \frac{g_1v + g_2v + \cdots + g_nv}{n}$$

as the barycenter of $\{gv : g \in G\}$. (Here, it might be that $g_iv = g_jv$, but we consider them as two points.) Then we claim that P is fixed by every $h \in G$. The reason is that if $h \in G$, then hP is the barycenter of $\{hvg : g \in G\} = \{g'v : g' \in G\}$ after changing variables $g' = hg$. This shows that $hP = P$ for arbitrary $h \in G$. □

Theorem 23.5. *If $G \subseteq \mathrm{Isom}(\mathbb{R}^2)$ is a finite subgroup, then G is either cyclic or dihedral.*

Proof. First there is a point P fixed by G. Then we can just move P to the origin and assume that G just fixes 0. Then $G \subseteq \mathrm{O}_2(\mathbb{R})$. Now consider the subgroup

$$G_0 = G_0 \cap \mathrm{SO}_2(\mathbb{R}) = \{g \in G : \det g = 1\} \subseteq G.$$

This is a finite subgroup of $\mathrm{SO}_2(\mathbb{R})$, and by the Proposition we stated at the beginning, we have that G_0 is cyclic.

If $G = G_0$, then G is cyclic and we are done. If $G_0 \subsetneq G$, then let us write $|G_0| = n$ and pick and element $\tau \in G$ with $\tau \notin G_0$. Then $\det\tau = -1$ because it is not in G. This means that τ is a reflection. Both τ and $\sigma\tau$ are relfections, so $\tau^2 = (\sigma\tau)^2 = 1$. Then

$$\tau\sigma = \sigma^{-1}(\sigma\tau)(\sigma\tau)\tau^{-1} = \sigma^{-1}\tau^{-1} = \sigma^{-1}\tau.$$

So G is a dihedral group in this case. □

24 March 26, 2018

We were talking about transformations. Recall that the special orthogonal group is

$$\mathrm{SO}_2(\mathbb{R}) = \{A \in M_2(\mathbb{R}) : AA^T = I, \det A = 1\}.$$

Note that this actually can be identified as the group of complex numbers with absolute value 1. This is because every $A \in \mathrm{SO}_2(\mathbb{R})$ looks like

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

and this is like $\cos\theta + i\sin\theta$. So $\mathrm{SO}_2(\mathbb{R}) \cong S^1$. (We are going to see that $\mathrm{SO}_3(\mathbb{R}) \cong \mathbb{R}P^3$ later.)

Last time we classified finite subgroups of $\mathrm{Isom}(\mathbb{R}^2)$. Recall that we define D_{2n} as the group of isometries of a regular n-gon.

Theorem 24.1. *Every finite subgroup of* $\mathrm{SO}_2(\mathbb{R})$ *is cyclic. Every finite subgroup of* $\mathrm{Isom}(\mathbb{R}^2)$ *either is a cyclic group of rotations, or isomorphic to* D_{2n}.

24.1 Geometry on S^2

Consider the unit sphere

$$S^2 = \{(x, y, z) : x^2 + y^2 + z^2 = 1\} \subseteq \mathbb{R}^3$$

sitting inside $\mathbb{R}^3$. (Here, S^2 is S^2 because the surface is 2-dimensional.) Here, we can talk about lines in S^2. Given two points $p, q \in S^2$, we look at the plane (in $\mathbb{R}^3$) passing through $0, p, q$. This plane cuts the sphere at a circle, which we call a **great circle**. Then the two points p and q are going to be joined by an arc.

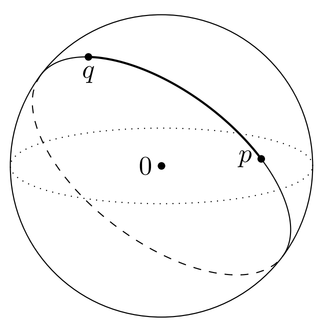

Figure 37: Great circles on a sphere

Here, this is a bit subtle because there are going to be two arcs joining the two points, on the great circle. Normally we are going to take the shorter one and define the distance $\mathrm{dist}(p, q)$ as the length of the shorter arc. Also note that if p and q lie on the exact opposite, so that $p, 0, q$ lies on a line, there are going to infinitely many lines joining p and q.

Proposition 24.2. *Any two distinct lines always intersect at exactly* 2 *points.*

Isometries of S^2 are going to be transformations that map great circles to great circles. These are also going to correspond to isometries of $\mathbb{R}^3$ that fix the origin. Examples include the reflections about the yz, xz, xy-planes

$$(x, y, z) \mapsto \begin{cases} (-x, y, z) \\ (x, -y, z) \\ (x, y, -z), \end{cases}$$

and there is also the antipodal map

$$(x, y, z) \mapsto (-x, -y, -z).$$

How do we write a general reflection about a plane? Suppose this plane is given by the equation $\vec{n} \cdot \vec{x} = 0$, where $\vec{n}$ is the unit normal vector. Then $(\vec{p} \cdot \vec{n})\vec{n}$ is the vector that connects the plane to $\vec{p}$ by the shortest length. So the reflection of $\vec{p}$ is going to be

$$\tau(\vec{p}) = \vec{p} - 2(\vec{p} \cdot \vec{n})\vec{n}.$$

You are going to check (in your homework) the three reflections theorem here as well.

Theorem 24.3 (three reflections theorem)**.** *Every isometry of S^2 can be written as a composition of at most* 3 *reflections.*

Proposition 24.4. *The composition of two reflections is a rotation.*

Proof. Let us reflect about the plane orthogonal to $\vec{a}$ and then about the plane orthogonal to $\vec{b}$. Then

$$\tau_b(\tau_a(\vec{p})) = \tau_b(\vec{p} - 2(\vec{p} \cdot \vec{a})\vec{a}) = \vec{p} - 2(\vec{p} \cdot \vec{a})\vec{a} - 2([\vec{p} - 2(\vec{p} \cdot \vec{a})\vec{a}] \cdot \vec{b})\vec{b}.$$

Here, let $\vec{c} = \vec{a} \times \vec{b}$. If $\vec{a}$ and $\vec{b}$ are linearly dependent, then the two reflections are equal and the composition is just the identity. If $\vec{a}$ and $\vec{b}$ are linearly independent, then $\vec{c} \neq 0$. Here, if the angle between $\vec{a}$ and $\vec{b}$ is θ, we claim that $\tau_b \circ \tau_a$ is actually rotation around the axis $\vec{c}$, by angle 2θ. You can calculate this out and compare it with the formula above. Alternatively, you can observe that $\tau_b(\tau_a(\vec{c})) = \vec{c}$ and use the fact that on the plane P, the transformation $\tau_b(\tau_a(\vec{c}))$ is just reflection about a line twice, and is thus a rotation. □

We can describe isometries of S^2 as matrices as well. Because these are isometries of $\mathbb{R}^3$ fixing the origin, we can describe this as

$$\mathrm{O}_3(\mathbb{R}) = \{A \in M_3(\mathbb{R}) : AA^T = I\}.$$

An isometry A is going to be a rotation if and only if $\det A = 1$, i.e., $A \in \mathrm{SO}_3(\mathbb{R})$. For instance, we can write down a rotation about the z-axis as

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

25 March 28, 2018

We were talking about isometries on the sphere. We saw that if we compose two reflections, the result is a rotation.

25.1 Orientation-preserving isometries are rotations

Proposition 25.1. *Any isometry of S^2 that fixes the origin and is represented by a determinant 1 matrix is a rotation.*

Proof. First, any isometry that fixes the origin may be considered as a linear transformation. To see this, we first note that

$$\vec{u} \cdot \vec{v} = \frac{1}{2}(|\vec{u}|^2 + |\vec{v}|^2 - |\vec{u} - \vec{v}|^2)$$

shows that any isometry preserves the inner product. Then if the transformation sends the standard basis vectors $\vec{e}_i$ to $f(\vec{e}_i)$, then it should satisfy

$$\vec{e}_i \cdot \vec{v} = f(\vec{e}_i) \cdot f(\vec{v})$$

and also $f(\vec{e}_i) \cdot f(\vec{e}_j) = \vec{e}_i \cdot \vec{e}_j = \delta_{ij}$. This means that $f(\vec{e}_i)$ form an orthonormal basis, so

$$f(\vec{v}) = \sum_{i=1}^{n} (f(\vec{v}) \cdot f(\vec{e}_i))\vec{e}_i = \sum_{i=1}^{n} (\vec{v} \cdot \vec{e}_i) f(\vec{e}_i).$$

So f really can be written as the matrix

$$f = \begin{pmatrix} | & \cdots & | \\ v_1 & \cdots & v_n \\ | & \cdots & | \end{pmatrix}.$$

Furthermore, it should satisfy $f^T f = \text{id}$, precisely because $f(\vec{e}_i) \cdot f(\vec{e}_j) = \delta_{ij}$.

Now f is a 3×3 matrix, so there are three eigenvalues. Moreover, all eigenvalues are going to have absolute value 1, because if λ is an eigenvalue with eigenvector $\vec{v}$, then

$$\overline{\vec{v}}^T \vec{v} = \overline{\vec{v}}^T \overline{f}^T f \vec{v} = (\overline{f\vec{v}})^T (f\vec{v}) = (\overline{\lambda\vec{v}})^T \overline{\lambda\vec{v}} = |\lambda|^2 \overline{\vec{v}}^T \vec{v}$$

implies that $|\lambda|^2 = 1$. Also, we have $\det A = 1$, so the multiple of the three eigenvalues should be 1. Because the determinant is the solution to $\det(f - \lambda I) = 0$, if one non-real number is an eigenvalue, its complex conjugate should also be an eigenvalue. (i) If some non-real number $\lambda_1 = \lambda$ is an eigenvalue, then $\lambda_1 = \overline{\lambda}$ is another eigenvalue, and the third should be $\lambda_3 = 1$ because the product of the three eigenvalues is 1. (ii) If all eigenvalues are real numbers, then they should be ± 1 because $|\lambda|^2 = 1$. Then the eigenvalues are either $+1, -1, -1$ or $+1, +1, +1$.

The conclusion is that 1 is an eigenvalue of f. Let $\vec{v}$ be the eigenvector, so that $A\vec{v} = \vec{v}$. This implies that f is a rotation about the axis $\vec{v}$. ∎

25.2 Quaternions

In the $\mathbb{R}^2$ case, we had seen that $\mathrm{SO}_2(\mathbb{R})$ were exactly like S^1, which is the complex numbers with absolute value 1. We are going to do a similar thing for $\mathrm{SO}_3(\mathbb{R})$.

Definition 25.2. The set of **quaternions** is

$$\mathbb{H} = \{a + b\vec{i} + c\vec{j} + d\vec{k} : a, b, c, d \in \mathbb{R}\}$$

that looks like $\mathbb{R}^4$, but with multiplication laws

$$\vec{i}^2 = \vec{j}^2 = \vec{k}^2 = -1, \quad \vec{i}\vec{j}\vec{k} = 1.$$

You can also thing of

$$q = a + b\vec{i} + c\vec{j} + d\vec{k} = \begin{pmatrix} a + ib & c + id \\ -c + id & a - ib \end{pmatrix}.$$

Note that if $\|q\| = 1$, then multiplication $v \mapsto qv$ is an isometry of $\mathbb{R}^4$, because

$$\|qv\| = \|q\|\|v\| = \|v\|.$$

Consider now the pure imaginary quaternions

$$\{q = b\vec{i} + c\vec{j} + d\vec{k} : b, c, d \in \mathbb{R}\}$$

that looks like $\mathbb{R}^3$. For each $p \in \mathbb{H}$ with $\|p\| = 1$, we consider the isometry

$$q \mapsto pqp^{-1}.$$

You can check that if q is pure imaginary, then pqp^{-1} is again a pure imaginary. So this can be thought of as an isometry of $\mathbb{R}^3$, instead of an isometry of $\mathbb{R}^4$.

If $\|p\| = 1$, we can write

$$p = \cos\frac{\theta}{2} + (l\vec{i} + m\vec{j} + n\vec{k})\sin\frac{\theta}{2}$$

where $l^2 + m^2 + n^2 = 1$. Then $q \mapsto pqp^{-1}$ is going to be a rotation (in the $(\vec{i}, \vec{j}, \vec{k})$-plane) through the axis through 0, about the axis $l\vec{i} + m\vec{j} + n\vec{k}$. Here, note that p and $-p$ represent the same rotation, becasue

$$(-p)q(-p)^{-1} = pqp^{-1}.$$

But other than this, these are going to be all the rotations, and the same rotations can only occur in this way. So we are going to get something like

$$\mathrm{SO}_3(\mathbb{R}) \cong \mathbb{R}P^3.$$

26 March 30, 2018

We were looking at the quaternions. Consider $q = a + b\vec{i} + c\vec{j} + d\vec{k}$ in $\mathbb{H}$ with $q = 1$. Then the map

$$\mathbb{H} \to \mathbb{H}; \quad p \mapsto qp$$

is an isometry. Moreover, if we look at the (i, j, k)-plane,

$$(i, j, k)\text{-plane} \to (i, j, k)\text{-plane}; \quad p \mapsto qpq^{-1}$$

is an isometry.

Note that first $-q$ defines the same isometry as q, because

$$(-q)p(-q)^{-1} = qpq^{-1}.$$

Also, if $|q| = 1$, we can write it as

$$q = \cos\frac{\theta}{2} + (l\vec{i} + m\vec{j} + n\vec{k})\sin\frac{\theta}{2}$$

with $l^2 + m^2 + n^2 = 1$.

26.1 Finite subgroups of SO(3)

For the 2-dimensional case, we looked at symmetry groups of regular polygons. Now, we are going to look at symmetry groups of regular polyhedra.

Let's look at the isometries of the tetrahedron. There first the rotation about the axis passing through one vertex. This already gives 3 isometries by rotating around. We can also move the stick around, in 4 ways. So in total, there are going to be 12 isometries. Concretely, this can be described as doing these 120° rotation, and also doing 180° rotations about the axis passing through two midpoints of the opposite edges of the tetrahedron.

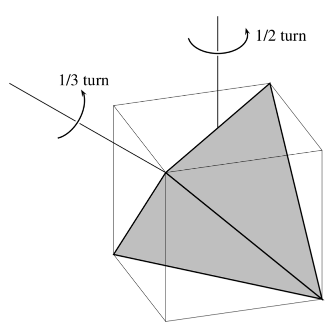

Figure 38: Isometries of the tetrahedron

We can also find the corresponding quaternions to these rotations.

- The identity isometry corresponds to ± 1.
- The 180° rotation is with $\theta = \pi$ and axis i or j or k. So they correspond to $\pm i, \pm j, \pm k$.
- The 120° rotation is with $\theta = \frac{2k\pi}{3}$ and so $\frac{\theta}{2} = \frac{k\pi}{3}$. The axis will be $\frac{1}{\sqrt{3}}(\pm i \pm j \pm k)$. So they correspond to

$$\pm\frac{1}{2} \pm \frac{i}{2} \pm \frac{j}{2} \pm \frac{k}{2}.$$

In total, we recover $1 + 3 + 8 = 12$, if we consider q the same as $-q$. But in terms of actual quaternions, the isometries of the tetrahedron is represented by 24 quaternions. If we think of these quaternions as 24 points in $\mathbb{H} \cong \mathbb{R}^4$, these will form the vertices of a 24-cell.

Also, note that isometries of $\mathbb{R}^3$ corresponds to unit quaternions modulo ± 1. So we have a bijection

$$\mathrm{SO}_3(\mathbb{R}) \cong \mathbb{R}P^3.$$

Then we can use this bijection to give $\mathbb{R}P^3$ a group structure.

27 April 2, 2018

Last week we have talked about various transformations. Now we are going to talk about the hyperbolic plane.

Consider the set
$$\mathcal{H} = \{z \in \mathbb{C} : \Im(z) > 0\} \subseteq \mathbb{C}.$$
In $\mathbb{R}P^1$, there were the linear fractional transformations

$$f(x) = \frac{ax+b}{cx+d}, \quad a, b, c, d \in \mathbb{R}.$$

It turns out that linear fractional transformations with $ad - bc > 0$ sends $\mathcal{H}$ to $\mathcal{H}$, and these are the transformations of the upper half-plane.

27.1 Stereographic projection

But let's talk about the space itself first. We learned before that

$$\mathbb{R}P^1 = \mathbb{R} \cup \{\infty\} \cong S^1.$$

You can also look at this using the stereographic projection. In $\mathbb{R}^2$, we draw a unit circle, and project the circle onto the x-axis from the point $(0, 1)$. This defines a map
$$\xi : S^1 \setminus \{(0, 1)\} \to \mathbb{R},$$
which we call the **stereographic projection**.

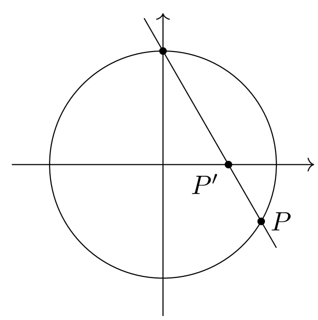

Figure 39: Stereographic projection for S^1

We can actually get an explicit formula for this projection. If $P = (x_0, y_0)$, then the equation for the line is going to be

$$y - 1 = \frac{y_0 - 1}{x_0} x.$$

So the intersection with $y = 1$ is going to be

$$x = \frac{x_0}{y_0 - 1}.$$

You can also compute the inverse of this map, and you are going to do this in the assignment.

For $\mathbb{C}P^1 = \mathbb{C} \cup \{\infty\}$, we can do the same thing. Consider a sphere $S^2 \subseteq \mathbb{R}^3$. Then again, we can project the sphere S^2 to the x_1x_2-plane $\mathbb{R}^2$ by projecting it from $(0, 0, 1)$. Then we have a projection

$$\xi : S^2 \setminus \{(0, 0, 1)\} \to \mathbb{R}^2 \cong \mathbb{C}.$$

Here, $(0, 0, 1)$ serves as the north pole, and you can get explicit formulae for both ξ and ξ^{-1}.

28 April 4, 2018

There is an Open Neighborhood Seminar on "A hands-on explanation of non-Euclidean geometry". Last time we were trying to describe the upper half-plane model. We talked about the stereographic projection

$$\xi : S^1 \setminus \{i\} \to \mathbb{R}, \quad S^2 \setminus \{N\} \to \mathbb{R}^2.$$

These allowed us two describe S^1 and S^2 as

$$S^1 = \mathbb{R} \cup \{\infty\}, \quad S^2 = \overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}.$$

Here, this $\overline{\mathbb{C}}$ is called the **Riemann sphere** because it actually looks like a sphere. Like in the 1-dimensional case, the projection map $\xi : S^2 \setminus \{N\} \to \mathbb{C}$ and its inverse map can be written down explicitly.

28.1 Upper half-plane

We can now talk about other structure on the Riemann sphere.

Definition 28.1. A **circle** on $\overline{\mathbb{C}}$ is either a circle in $\mathbb{C}$ or Euclidean line on $\mathbb{C}$ along with ∞.

For instance, the x-axis $\{z : z \in \mathbb{R}\} \cup \{\infty\}$ is a circle in $\overline{\mathbb{C}}$. We note that a circle always divides the Riemann sphere into two pieces. If we look at the circle $\{w \in \mathbb{C} : |w - z| = r\}$, the two regions are going to be

$$U_r(z) = \{w \in \mathbb{C} : |w - z| < r\}, \quad U_r(\infty) = \{w \in \mathbb{C}; |w - z| > r\}.$$

Note that both regions just look like discs. For the outer region, it is a disc as well because we are actually cutting a sphere into two parts by a circle.

The same thing works if our circle is $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$ in the Riemann sphere. This also cuts the sphere into two pieces, and one piece

$$\mathcal{H} = \{z \in \mathbb{C} : \Im(z) > 0\}$$

is the **upper half-plane**. So you can think of the boundary of $\mathcal{H}$ as $\mathbb{R} \cup \{\infty\} = \mathbb{R}P^1$.

Now consider a linear fractional transformation $f : \mathbb{R}P^1 \to \mathbb{R}P^1$, that looks like

$$f(x) = \frac{ax + b}{cx + d},$$

where $a, b, c, d \in \mathbb{R}$. We can regard this as a linear fractional transformation

$$f : \overline{\mathbb{C}} \to \overline{\mathbb{C}}; \quad f(z) = \frac{ax + b}{cz + d}.$$

Then this maps the $\mathbb{R}P^1$ inside $\overline{\mathbb{C}}$ to itself. If we impose the condition that $ad - bc > 0$, we even see that f maps $\mathcal{H}$ to $\mathcal{H}$.

We can describe inversion in this setting as well. If we want to send

$$z = re^{i\theta} \mapsto \frac{1}{r}e^{i\theta},$$

we can look at the transformation

$$z \mapsto \frac{1}{\bar{z}}.$$

29 April 6, 2018

Last time we were extending the linear fractional transformations on $\mathbb{R}P^1$ to the upper half-plane $\mathcal{H}$. We had the following maps:

- $z \mapsto z + \alpha$ for $\alpha \in \mathbb{R}$,
- $z \mapsto k\alpha$ for $k > 0$,
- $z \mapsto \frac{1}{\bar{z}}$,
- $z \mapsto -\bar{z}$.

All these transformations map $\mathcal{H}$ to $\mathcal{H}$. These generate the **Möbius transformations**. Generally, Möbius transformations mean $\overline{\mathbb{C}} \to \overline{\mathbb{C}}$, but here we're only considering maps $\mathcal{H} \to \mathcal{H}$.

29.1 Reflection about a circle

Consider a circle centered at a point C of radius r on the plane. We say that B is the **reflection** of A about the circle if A, B, C lie on a line and $|AC||BC| = r^2$. The formula can be written as

$$f : A \mapsto B = C + (A - C)\frac{r^2}{|AC|^2}$$

because we want $B - C$ to be the same direction as $A - C$ but with length $r^2/|AC|$. Here, we may consider this reflection as sending $C \mapsto \infty$ and $\infty \mapsto C$. This allows us to extend this reflection to $\overline{\mathbb{C}} \to \overline{\mathbb{C}}$.

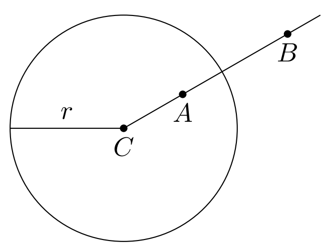

Figure 40: Reflection of a point about a circle

Note that we can generalize this to $\mathbb{R}^n$ for arbitrary n. We can define the map $f : \mathbb{R}^n \setminus \{C\} \to \mathbb{R}^n \setminus \{C\}$ in the same way (using the same formula), and then extend it to $f : S^n \to S^n$ by sending $C \mapsto \infty$ and $\infty \mapsto C$. For $n = 2$, these are going to be Möbius transformations as well.

29.2 Lines in the upper half-plane

Definition 29.1. The **lines** are going to be

(1) lines in $\mathbb{R}^2$ that are perpendicular to the real axis $\mathbb{R}$ in $\mathbb{C}$,

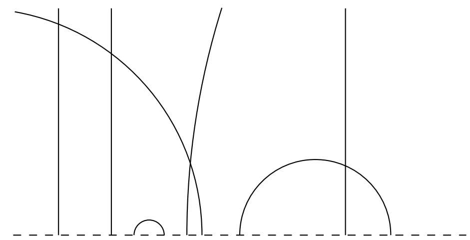

Figure 41: Lines on the upper half-plane $\mathcal{H}$

(2) semi-circles that are centered at the real axis $\mathbb{R}$.

Recall that we have computed the expression for the circle $(x-h)^2+(y-k)^2 = r^2$ in terms of $z = x + iy$ before. We saw that the equations looks like

$$z\bar{z} - \bar{z}_0 z - z_0\bar{z} + z_0\overline{z_0} = r^2$$

for some z_0. So the general formula for a line in $\mathcal{H}$ is going to be

$$A|z|^2 + B(z + \bar{z}) + C = 0$$

for some $A, B, C \in \mathbb{R}$. Note that if $A = 0$, this becomes the vertical line.

Proposition 29.2. *Möbius transformations preserve lines in* $\mathbb{C}$.

Proof. We check that the four generating transformations $z \mapsto z + \alpha$, $z \mapsto \alpha\zeta$, $z \mapsto -\bar{z}$, $z \mapsto 1/\bar{z}$ all preserve lines.

For the first one, this is just shifting along the real direction. So it clearly preserves circle and lines. For the second one, this is scaling and again preserves circles and lines. The third one is reflection about the line $\Re(z) = 0$ and so it is again clear that this preserves lines in $\mathcal{H}$. The fourth one is the tricky one. We know that the equation of a line in $\mathcal{H}$ looks like

$$A|z|^2 + B(z + \bar{z}) + C = 0.$$

If we divide this by $|z|^2 = z\bar{z}$, this is the same as

$$A + B\Big(\frac{1}{z} + \frac{1}{\bar{z}}\Big) + C\frac{1}{|z|^2} = 0.$$

So for $w = 1/\bar{z}$, the equation can be expressed as

$$A + B(w + \bar{w}) + C|w|^2 = 0.$$

This is a line in w, and this means that the transformation $z \mapsto w = 1/\bar{z}$ sends lines to lines. □

Definition 29.3. Two hyperbolic lines (i.e., lines in $\mathcal{H}$) are **parallel** if two are disjoint.

Here, we can observe in Figure 41, there are a bit too many parallel lines to a given parallel lines. In Euclidean geometry, the parallel postulate stated that given any point P and a line ℓ not passing through P, there exists a unique line passing through P and parallel to ℓ. In the upper half-plane $\mathcal{H}$, we are going to see that there are going to be infinitely many parallel lines parallel to ℓ passing through P.

30 April 11, 2018

Last time we talked about lines in $\mathcal{H}$. There were two types of lines, vertical lines and semi-circles that are centered on the real line. We can extend any line infinitely in both directions, because we should think of semi-circles as infinite lines.

30.1 Ultraparallel lines

Definition 30.1. We say that two lines ℓ_1 and ℓ_2 are **parallel** if they are disjoint. We say that they are **ultraparallel** if the boundaries of the two lines at ∞ are disjoint.

Proposition 30.2. *Let ℓ_1 and ℓ_2 be parallel lines. Then ℓ_1 and ℓ_2 are ultraparallel if and only if there exists a unique hyperbolic line perpendicular to both ℓ_1 and ℓ_2.*

Proof. We just divide into many cases. Suppose that ℓ_1 and ℓ_2 are ultraparallel and let us show that there exist a unique line perpendicular to both. If ℓ_1 is a vertical line, then ℓ_2 should be a semi-circle disjoint from ℓ_1. Then the line perpendicular to ℓ_1 should be a semi-circle centered at where ℓ_1 meets the real axis, and from this we see that there is a unique such semi-circle that is also perpendicular to ℓ_2. If ℓ_1 and ℓ_2 are both semi-circles, let ℓ_i be centered at c_i with radius r_i. If $c_1 \neq c_2$, we need a semi-circle centered at c with radius r, satisfying $|c - c_i|^2 = r^2 + r_i^2$ and so $2c(c_2 - c_1) = r_1^2 - r_2^2$. This can be solved uniquely, and $r > 0$ follows from $|c_1 - c_2| > r_1 + r_2$ or $|c_1 - c_2| < |r_1 - r_2|$. If $c_1 = c_2$, then the vertical line $\Re(z) = c_1$ is perpendicular to both ℓ_1 and ℓ_2.

Conversely, if some line is parallel to both ℓ_1 and ℓ_2, then we can check that they are ultraparallel. □

30.2 Möbius transformations are conformal

We have seen that Möbius transformations preserve angles. Here, we are going to show that they preserves angles between lines as well. But what is an angle?

Definition 30.3. Given two lines ℓ_1 and ℓ_2 intersecting at P, the **angle** between ℓ_1 and ℓ_2 is a angle between the tangent lines to ℓ_1, ℓ_2 at P.

Proposition 30.4. *Möbius transformations preserve angles between lines.*

Proof. It suffices to check this for the generating transformations $z \mapsto z + r$, $z \mapsto kz$, $z \mapsto -\bar{z}$, and $z \mapsto 1/\bar{z}$. The first three are clear, because they are translation, scaling, and reflection. To show that $z \mapsto 1/\bar{z}$ preserves angles, we instead check that $z \mapsto -1/z$ preserves angles. This is okay because if this preserves angles, we can reflect to get $z \mapsto 1/\bar{z}$.

So consider a point z and a small direction $\Delta z = \epsilon e^{i\theta}$. If we compare how z and $z + \epsilon e^{i\theta}$ are mapped under $f(z) = -\frac{1}{z}$, we have

$$\Delta f(z) = f(z + \Delta z) - f(z) = -\frac{1}{z + \epsilon e^{i\theta}} + \frac{1}{z} = \frac{\epsilon e^{i\theta}}{z(z + \epsilon e^{i\theta})} \approx \frac{\epsilon e^{i\theta}}{z^2}$$

for $\epsilon \ll 1$. So if we have another θ', we will have

$$\Delta' f(z) = \frac{\epsilon e^{i\theta'}}{z^2} = e^{i(\theta'-\theta)} \Delta f(z).$$

This means that the angle between Δz and $\Delta' z$ is $\theta' - \theta$, and this is the same as the angle between $\Delta' f(z)$ and $\Delta f(z)$. □

Möbius transformations also preserve lengths, if we use a carefully-chosen notion of length. To define length, we use cross ratio. Recall that

$$[z_1, z_2; z_3, z_4] = \frac{(z_1 - z_3)(z_2 - z_4)}{(z_1 - z_4)(z_2 - z_3)}.$$

Because Möbius transformations are linear fractional transformations, they preserve cross ratio. In particular, if the four points z_1, z_2, z_3, z_4 lie on a line, then the cross ratio is a real number. For instance if $z_k = iz_k$ for a_k real numbers, we have

$$[ia_1, ia_2; ia_3, ia_4] = \frac{(a_1 - a_3)(a_2 - a_4)}{(a_1 - a_4)(a_2 - a_3)} \in \mathbb{R}.$$

31 April 13, 2018

We checked long before that fractional linear transformations preserve cross-ratio
$$[z_1, z_2; z_3, z_4] = \frac{(z_1 - z_3)(z_2 - z_4)}{(z_1 - z_4)(z_2 - z_3)}.$$
Our claim here is the cross ratio of any 4 points on a hyperbolic line is a real number. In the case when z_1, z_2, z_3, z_4 all lie on $\ell = \{x = 0\} \cap \mathcal{H}$, then we can check this directly. Given any other line k, we can apply a fractional linear transformation to send k to ℓ. In this process, the cross-ratio is preserved, so we conclude that if $z_1, z_2, z_3, z_4 \in k$, then the cross ratio is in $\mathbb{R}$.

31.1 Hyperbolic distance

Consider the line $\ell = \{x = 0\} \cap \mathcal{H}$. On this line, consider two points pi and qi. Then we can look at
$$[pi, qi; r = 0, s = \infty] = \frac{(pi - r)(qi - s)}{(pi - s)(qi - r)} = \frac{(pi - 0)(-\infty)}{(-\infty)(qi - 0)} = \frac{p}{q}.$$
We want to say that this is something like distance in the hyperbolic plane, because the value is not going to change under Möbius transformations. But things like $x \mapsto -1/\bar{x}$ will send 0 to ∞ and ∞ to 0. Then our cross ratio becomes $\frac{q}{p}$. To fix this issue, we take
$$d(p, q) = |\log[p, q; 0, \infty]|$$
as the **distance**upper half-plane!distance in the upper half plane. Another reason for taking the log is that if we have three points pi, qi, ri, the distances add up as
$$d(pi, ri) = \left|\log \frac{r}{p}\right| = \left|\log \frac{q}{p}\right| + \left|\log \frac{r}{q}\right| = d(pi, qi) + d(qi, ri)$$
on a line.

A natural question is, what does the set of equidistanct points from a given point look like? Let us try to find out what the set of point that has distance r from i is. This is the same as finding all u such that $d(u, i) = r$. On the line $\ell = \{\Re(z) = 0\}$, the distance of ui and i is
$$d(ui, i) = |\log u|,$$
and so ie^r and ie^{-r} are the two points of distance r from i.

Now we note that any Möbius transformation that preserves the center i preserves distance from i. This means that if we map ie^r and ie^{-r} under any transformation that looks like
$$z \mapsto \frac{az - c}{cz + a}, \quad z \mapsto \frac{a\bar{z} + c}{c\bar{z} - a},$$

(where we assume $a^2 + c^2 = 1$) will lie on the equidistant set. We can write $a = \cos\theta$ and $c = \sin\theta$, and then the image is

$$iu \mapsto \frac{\sin 2\theta(u - u^{-1})}{2(u^{-1}\cos^2\theta + u\sin^2\theta)} + i\frac{1}{u^{-1}\cos^2\theta + u\sin^2\theta}.$$

We can find that the image with $u = e^{\pm r}$ is going to be

$$x^2 + (y - \cosh r)^2 = \sinh^2 r.$$

That is, the equidistant set from a point is a Euclidean circle, although it is going to be skewed.

We can make sense of congruence between triangles too.

Proposition 31.1. *On the upper half plane, the conditions SSS, SAS, ASA, AAA are all equivalent, and we call this* ***congruence*** *of triangles.*

The reason for AAA implying congruence is because the sum of the angle can be used to calculate the area of the triangle.

31.2 Tiling by triangles

In the Euclidean plane, consider a triangle with angles $\frac{\pi}{p}, \frac{\pi}{q}, \frac{\pi}{r}$ with $p \le q \le r$. Then this necessarily satisfy

$$\frac{\pi}{p} + \frac{\pi}{q} + \frac{\pi}{r} = \pi,$$

and if you work hard, you see that $(p, q, r) = (3, 3, 3), (2, 3, 6), (2, 4, 4)$ are the only possibilities. Then we can find the corresponding tiling of $\mathbb{R}^2$ so that at each point, we only see the same angles.

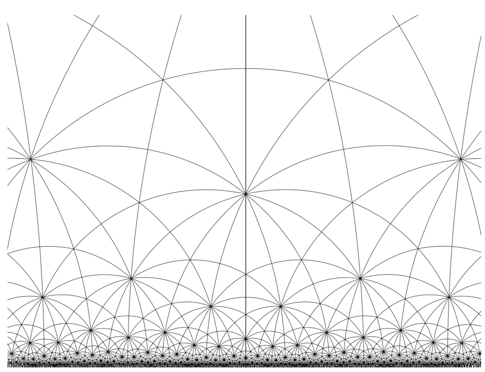

Figure 42: $(2, 3, 7)$-tiling of the upper half-plane

But in $\mathcal{H}$, we can have a triangle of angles $\frac{\pi}{p}, \frac{\pi}{q}, \frac{\pi}{r}$ as long as

$$\frac{\pi}{p} + \frac{\pi}{q} + \frac{\pi}{r} < 1.$$

So there are many more possibilities to choose. In fact, for any such (p, q, r) there is going to be such a tiling. For instance, Figure 42 shows a tiling with $(p, q, r) = (2, 3, 7)$.

32 April 16, 2018

Last time we defined distance. Today we are going to show that AAA actually gives a congruence of triangles. Because we can apply a Möbius transformation move any side to any other side, it is straightforward to see that ASA and SAS gives congruence. SSS gives congruence because two equidistant circle meet at only two points, on opposite sides on the line connecting the two centers.

32.1 Sum of angles in a triangle

We are going show that AAA implies congruence.

Proposition 32.1 (alternate angle theorem)**.** *If two lines m and n intersect ℓ at B, B', and the angle between m and ℓ and the angle between n and ℓ are equal, then the lines m and n cannot intersect.*

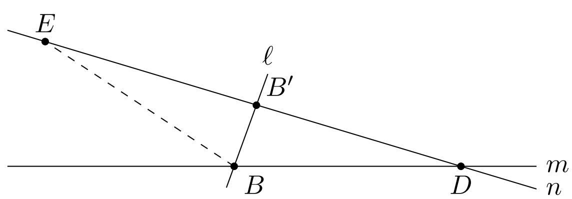

Figure 43: Alternate angle theorem

Proof. Suppose that m and n intersect at D. Pick a point E on m such that $B'D = BE$. Then by SAS, the triangles $B'BD$ and $BB'E$ are congruent. This shows that

$$\angle EBD = \angle EBB' + \angle B'BD = \angle DB'B + \angle BB'E = \pi.$$

Then E should be another intersection of m and n, which is a contradiction. □

Definition 32.2. The **defect** of a triangle ABC is defined as

$$\text{defect}(ABC) = \pi - \angle A - \angle B - \angle C.$$

Proposition 32.3. *If there exists a triangle of defect 0, then there exists a rectangle, i.e., a quadrilateral with all right angles. If a rectangle exists, then every triangle has defect 0.*

Proof. Roughly, you first construct a right triangle having defect 0, and from this construct a rectangle by putting two together. Given a rectangle, you can construct arbitrarily large rectangles, and this can be used to show that all right triangles have defect 0. Then given an arbitrary triangle you can divide it into two by an altitude and show that the triangle has defect 0. □

This is really strong because only one triangle having defect 0 shows that all triangles have defect 0. Then if there is one triangle of defect nonzero, all triangles should have defect nonzero as well.

Lemma 32.4. *There exists a triangle in the hyperbolic plane with the angle sum less than π.*

Proposition 32.5. *In hyperbolic plane, rectangles do not exist, and all triangles have angle sum less than π.*

Corollary 32.6. *In the hyperbolic plane H, all quadrilaterals have angle sum less than 2π.*

Theorem 32.7 (AAA congruence)**.** *Let ABC and DEF be two triangles with $\angle A = \angle D$, $\angle B = \angle E$, and $\angle C = \angle F$. Then $\triangle ABC \cong \triangle DEF$.*

Sketch of proof. We first make $A = D$, and then A, E, F and A, F, C be on lines. If the two triangles are not congruent, we have $E \neq B$ and $F \neq C$. In this situation, the sum of the angles of the quadrilateral $BCFE$ is 2π. □

One reason you should expect AAA congruence is because we have

$$\text{defect}(\Delta) \sim \text{area}(\Delta)$$

in the hyperbolic plane. So one the angles are fixed, the defect is fixed and so the area of the triangle is fixed.

33 April 20, 2018

When people talk about the Möbius group, they usually mean the set of transformations $\overline{\mathbb{C}} \to \overline{\mathbb{C}}$, not only those sending the upper half-plane to the upper half plane. These look like

$$m(z) = \frac{az+b}{cz+d}$$

where a, b, c, d now can be complex numbers, with $ad - bc \neq 0$.

33.1 Fixed points of a Möbius transformation

Definition 33.1. A **fixed point** of a transformation m is a point $z \in \overline{\mathbb{C}}$ such that $m(z) = z$.

Let us try to find the fixed points of a Möbius transformation $m \neq \mathrm{id}$. First consider the case when

$$m(\infty) = \frac{a}{c} = \infty.$$

Then $c = 0$ and $m(z) = \frac{a}{d}z + \frac{b}{d}$. Then this has no other fixed point if $\frac{a}{d} = 1$, and exactly one other fixed point if $\frac{a}{d} \neq 1$. Now consider the case when $m(\infty) \neq \infty$. Then we can just solve the equation

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

which also can be written as $cz^2 + (d-a)z - b = 0$. This has at most two solutions, and so at most two fixed points. So in all cases, if $m \neq \mathrm{id}$ then m can have at most two fixed points.

Theorem 33.2. *If m has 3 fixed points, then $m = \mathrm{id}$.*

Let us look at some examples. For $m(z) = \frac{2z+5}{3z-1}$, $m(\infty) = \frac{2}{3}$ and so the fixed points are the two solutions of the quadratic equation of $3z^2 - 3z - 5 = 0$. For $m(z) = 7z + 6$, the two fixed points are $z = \infty$ and $z = -1$. For $m(z) = \frac{z}{z+1}$, there is a double root at $z = 0$.

We can even classify Möbius transformations up to conjugation.

Definition 33.3. We say that two Möbius transformations are conjugate if there exists a Möbius transformation p such that $pm_1p^{-1} = m_2$.

You can check that being conjugate is an equivalence relation. So we want to figure out what the equivalence classes look like.

If $m = \mathrm{id}$, then any conjugate of $m = \mathrm{id}$ is $pmp^{-1} = pp^{-1} = \mathrm{id}$. So the equivalence class is just $\{\mathrm{id}\}$.

Now assume that m has only one fixed point x. Take a point $y \in \overline{\mathbb{C}} \setminus \{x\}$, and consider a Möbius transformation p sending

$$p : x \mapsto \infty, \quad y \mapsto 0, \quad m(y) \mapsto 1.$$

If we let $r = p \circ m \circ p^{-1}$, we have

$$p \circ m \circ p^{-1}(\infty) = p \circ m(x) = p(x) = \infty.$$

So ∞ is a fixed point and r looks like $r(z) = az + b$. Because p has one fixed point, r also has only one fixed point. So $r(z) = z + b$. But $r(0) = pmp^{-1}(0) = pm(y)) = 1$. So $r(z) = z + 1$. That is, p is conjugate to $z \mapsto z + 1$. In this case, we say that m is **parabolic**.

33.2 Poincaré disc model

Consider the disc

$$\mathcal{D} = \{z \in \mathbb{C} : |z| < 1\}.$$

There is a Möbius transformation

$$m : \mathcal{D} \to \mathcal{H}, \quad i \mapsto 0 \quad -1 \mapsto 1, \quad 1 \mapsto \infty,$$

explicitly given by

$$m(z) = \frac{z - i}{z - 1} \cdot \frac{-2}{-1 - i}.$$

This gives a bijection between $\mathcal{D}$ and $\mathcal{H}$, we can think of $\mathcal{D}$ as the same upper half-plane (or the hyperbolic plane). In this model, lines on $\mathcal{D}$ are circles that are orthogonal to the unit circle.

34 April 23, 2018

Last time we looked at the Möbius group of Möbius transformations.

34.1 Minkowski space

Minkowski space is just $\mathbb{R}^{n-1,1} = \mathbb{R}^n$ with the inner product

$$\vec{v} \cdot_L \vec{w} = v^T L w = v_1 w_1 + \cdots + v_{n-1} w_{n-1} - v_n w_n, \quad L = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & -1 \end{pmatrix}.$$

These appear in general relativity, and linear transformations preserving this inner product are called **Lorentz transformations**. This is

$$\{A : A^T L A = L\}.$$

Special relativity is doing geometry in $\mathbb{R}^{3,1}$, where the first three coordinates are like space and the last coordinate is like time.

We say that v is

- **spacelike** if $v \cdot_L v > 0$,
- **timelike** if $v \cdot_L v < 0$,
- **lightlike** if $v \cdot_L v = 0$,

In the spacetime, matter only move in timelike directions, and light moves in lightlike directions. The set

$$\{v : v \cdot_L v = -1\} = \{(x, y, z) : x^2 + y^2 - z^2 = -1\}$$

can serve as another model for the hyperbolic space.

34.2 Presentation I - Torus

This was a presentation by Shanelle. A torus is a surface that can be constructed by rotating a circle around a coplanar axis in 3-dimensional space. If we take a circle on the xz-plane of radius d, centered at $(S, 0, 0)$, we get a parametrization

$$x = (S + d\cos\theta)\cos\rho, \quad y = (S + d\cos\theta)\sin\rho, \quad z = d\sin\theta.$$

If $S > d$, then this is just the normal torus we have, which is called the ring torus, but if $S = d$, there is some degeneracy and we get a horn torus. If $S < d$, we will have some intersecting point and in this case the torus is called a spindle torus. If $S = 0$, the torus will degenerate into a sphere. You can compute the volume and surface area, and it is going to be

$$V = (2\pi S)(\pi d^2), \quad A = (2\pi S)(2\pi d).$$

34.3 Presentation II - Gabriel's horn

This was a presentation by Anne. The point of Gabriel's horn is that this has finite volume but infinite surface area. This is counter-intuitive, because you can fill the horn with paint, but you cannot not color the surface of the horn. Consider the graph of

$$y = \frac{1}{x}, \quad 1 \leq x$$

and rotate it around the x-axis in 3-dimensional space. We can calculate volume and surface area using standard calculus. If we cut it the horn to $1 \leq x \leq a$, and then the volume and area are

$$V = \pi r_a^2 = \pi \int_1^a \frac{dx}{x^2} = \pi\Big(1 - \frac{1}{a}\Big), \quad A = 2\pi \int_1^a \frac{1}{x}\sqrt{1 + \frac{1}{x^2}} dx > 2\pi \int_1^a \frac{dx}{x} = 2\pi \log a.$$

So as $a \to \infty$, we get finite volume but $A \to \infty$.

The converse is impossible, i.e., we can't have finite area but infinite volume, at least for surfaces of revolution. If we rotate $f(x)$, then

$$\limsup_{t\to\infty} f(x)^2 - f(1)^2 = \limsup_{t\to\infty} \int_1^x |2f(t)f'(t)| dt \leq \frac{A}{\pi}.$$

If we assume that A is finite, then this shows that f is bounded. Let's say that it is bounded by M. Then

$$V = \int_1^\infty \pi f^2(x) dx \leq \frac{M}{2} \int_1^\infty 2\pi f(x) dx \leq \frac{M}{2} \int_1^\infty 2\pi f(x)\sqrt{1 + f'^2(x)} dx = \frac{M}{2} A$$

implies that V is finite.

34.4 Presentation III - Octahedron

This was a presentation by Ricky. We can draw the octahedron by putting two square pyramids together. You can count that this has 12 edges, 8 faces, and 6 faces. This satisfies Euler's equation $v - e + f = 2$ where v is the number of vertices, e is the number of edges, and f is the number of faces. So the octahedron has $(v, e, f) = (6, 12, 8)$ and if you look at the cube, we have $(v, e, f) = (8, 12, 6)$. These are called dual polyhedra and we can put the octahedron inside a cube and a cube inside the octahedron.

The octahedron can be written down using the equation

$$|x| + |y| + |z| = 1.$$

The surface of the octahedron looks like the surface of a sphere, and the stereographic projection of this to the plane will look like a 3-set Venn diagram. There is a theorem that says that the symmetries of a polyhedron is the same as the symmetries of a octahedron. You can count and they are going to have the same number of symmetries.

35 April 25, 2018

35.1 Presentation IV - Möbius strip

This was a presentation by Jungyeon. The Möbius strip is a strip that does not have one "side". So you can't color it with two colors so that at each point, the opposite sides of the strip have different colors. This means that the surface is unorientable. Here is one way you can see this. If you take some figure on the surface and move it around the strip once, the figure becomes its mirror image. Its Euler characteristic is $\chi = v - e + f = 0$.

If you Möbius strip has some interesting geometry. If you take a Möbius strip and cut it along the middle line, you get a regular strip. If you take a Möbius strip, and cut it along the one-third line, you will get two strips linked together.

35.2 Presentation V - More on the Möbius strip

This was a presentation by Matt. One way you can think of non-orientability is to look at the normal vector to the surface and sliding it around the strip. You can also think of the Möbius strip as gluing two opposite sides of a rectangle in the reverse direction.

You can parametrize the Möbius strip using

$$x = [R + s\cos(\tfrac{1}{2}t)]\cos t, \quad y = [R + s\cos(\tfrac{1}{2}t)]\sin t, \quad z = s\sin(\tfrac{1}{2}t))$$

for $s \in [-w, w]$ and $t \in [0, \pi]$. You an also describe this as

$$-R^2y + x^2y + y^3 - 2Rxz - 2x^2z - 2y^2z + yz^2 = 0.$$

In astrophysics, there is this theory that there is a non-orientable wormhole, so that if you go around in space some way, you get the mirror image. People sometimes use the Möbius strip for a conveyor belt so that both sides of the belt are used and the belt lasts longer.

35.3 Presentation VI - Nine point circle

This was a presentation by Tushar. Consider any triangle ABC. The statement says that if you look at the three feet of altitudes, the midpoints of the three sides, and also the three midpoints of AH, BH, CH, these nine points lie on a single circle. (Here H is the orthocenter.)

Given a triangle ABC, we can find a triangle $A'B'C'$ such that A is the midpoint of $B'C'$ and B is the midpoint of $C'A'$ and C is the midpoint of $A'B'$. (Here, $A'BAC$ is going to be a parallelogram and so on.) So the orthocenter of ABC is going to be the circumcenter of $A'B'C'$.

By Thales's theorem, the midpoints of AH, AB, CB, CH form a rectangle. You can do this for all three sides, and we get that all the midpoints of AB, BC, CA, AH, BH, CH lies on the same circle. Then we can show that the feets of altitudes are also on the same circle.

There is also this cool line called the Euler line. It turns out the orthocenter H, circumcenter O, centroid G, and the center of the nine point circle N all lies on this line. N turns out to be the midpoint of OH.

35.4 Presentation VII - Tropical geometry

This was a presentation by Jules. Let's first define the tropical algebra.

Definition 35.1. The semi-field $T = \mathbb{R} \cup \{-\infty\}$ is defined as

$$x \oplus y = \max\{x, y\}, \quad x \odot y = x + y.$$

So if we try to draw a graph for $0 + x - x^2$ in the tropical algebra, you will see that this is actually $\max\{0, x, 2x - 1\}$ so that we only have a segmented lines. Polynomials in the algebra are things that look like

$$P(x, y) = \sum_{i,j} a_{ij} x^i y^j = \max_{i,j}(a_{i,j} + ix + jy).$$

We can think about tropical curves. For instance, if we take $0 \oplus x \oplus y = 0$, this is going to consist of three rays $x_0 = 0 \geq y_0$ and $y_0 = 0 \geq x_0$ and $x_0 = y_0 \geq 0$. Given such a curve, you can find a dual subdivision, such that vertices on the curve corresponds to regions in the dual, and lines on the curve corresponds to segments in the dual. Given a tropical curve, you can also define its ameoba, which sort of thickens the curve.

Index