MAU22302/33302 - Euclidean and Non-Euclidean Geometry

Adam Keilthy

Trinity College Dublin

Contents

0 A brief history of geometry

Geometry is a slippery beast to define precisely, but is probably best described as the study of shapes, sizes, and spatial properties. It is a field with an extremely long history, and much early mathematics was understood through the lens of geometry. Modern geometry can be broadly split into two categories: *Euclidean* geometry and *non-Euclidean* geometry, though other adjectives can be applied. For example, there is also projective geometry (often classified as non-Euclidean) in which there is no great notions of distance. We can also consider discrete geometry, which studies things such as tesselations, triangulations, and lattice polytopes. We will try to cover some of all of these areas in this course, though the focus will be on Euclidean and non-Euclidean geometries with notions of length.

Euclidean geometry is the geometry we are probably most familiar with. It is the geometry we are introduced to in our early schooling, is considered a core part of most second level maths curricula, and is the geometric framework on which we built most of our understanding of the physical world. It is based on the axioms set out by Euclid in the *Elements*.

Euclid was a Greek mathematician, alive circa 300 BCE, who wrote a collection of 13 books laying out the foundations of mathematics. The first six of these lay out a "flat" geometry, discussing properties of lines, angles, circles, and other plane figures. Books VII-X cover a good deal of number theory, including prime factorisations, common divisors, and irrational numbers, though it is all understood through a geometric language. Books XI and XII discuss solid geometry, and even feature some early ideas of calculus in relation to volumes. Book 13 is a bit more mixed, combining 2D and 3D geometry, and deriving geometric inequalities.

Some of the key defining features of Euclidean geometry include familiar results such as

- The sum of the angles in a triangle is $180°$,
- If a, b, c are the sides of a right angled triangle, with c opposite the right angle, then
$$a^2 + b^2 = c^2$$
- If C is the circumference of a circle with diameter D, then $C = \pi D$.

Despite being the most ubiquitous form of geometry, Euclidean geometry is not without its controversies. Most famously, Euclid's fifth axiom (or postulate) was widely believed to be a consequence of the first 4 until the 1800s. The *parallel postulate* was considered so artificial compared to the other postulates that a substantial amount of time was spent trying to derive it as a theorem, with very little success. Despite having since shown that the parallel postulate is independent of the others, you still find a lot of attempts to derive it, or to replace it with a more "natural" assumption!

Non-Euclidean geometry came about from accepting that the parallel postulate did not need to hold, and was formally developed first in the early 1800s.

A system of geometry in which the parallel postulate did not hold was first discussed by Gauss in a series of letters, and then was fully developed by Lobachevsky. This system of geometry was an example of what is now knowns as *hyperbolic geometry.* In hyperbolic geometry, we find a number of counter intuitive results:

- Given a line ℓ and a point P not on ℓ, there exists infinitely many lines through P that do not intersect ℓ,
- The sum of angles in a triangle is less that $180°$. In fact, there are hyperbolic triangles with arbitrarily small angle sums!
- The ratio of the circumference of a circle to its diameter is greater than π.

Examples of hyperbolic geometries are often difficult to visualise. Locally, a hyperbolic surface looks something like the surface of a saddle, with coral reefs giving a decent impression of the overall shape. A very well known visualisation of a hyperbolic surface is the Poincaré disc, made popular in M.C. Escher's '*Angels and Devils*'. Building three dimensional models took longer. There were early attempts with very delicate paper structures, but now a very popular way of building models of a hyperbolic plane is via crochet.

The next main example of non-Euclidean geometry to be formally developed was *elliptic geometry*, studied in the framework developed by Riemann in the 1850s. Technically, elliptic geometry has been around for much longer, with a very complete theory of spherical trigonometry being developed for navigation purposes. However, as geometry on a sphere or ellipsoid does not satisfy the the first four postulates of Euclidean geometry, and spheres are very easy to consider in a Euclidean framework, this area was largely neglected until Riemann. In elliptic geometry, we find

- Any two lines will intersect if extended far enough, so parallel lines do not exist,
- The sum of angles in a triangle is greater than $180°$. A well known lateral thinking problem features an Arctic explorer walking in a triangle where every angle is a right angle!
- The ratio fo the circumference of a circle to its diameter is less than π.

Also under the umbrella of non-Euclidean geometry is *projective geometry*, developed by Poncelet, largely while as a prisoner of war during the Napoleonic wars. Projective geometry has a lot of similarities to Euclidean geometry, but includes "points at infinity" and lacks any real notions of distances and angles.

Non-Euclidean geometry, despite being less familiar, has become very important in the modern world. Not only does it provide inspiration for art and architecture, but non-Euclidean frameworks were essential for the development of general relativity, and can be used to streamline computer graphics and navigation systems.

1 Classical Euclidean Geometry

Classical (or *synthetic*) Euclidean geometry is what most of us think of when we think of geometry. We consider only lengths, angles, and ratios. There is no algebra as in coordinate geometry and we rely on ideas of congruent triangles, intersecting circles, and so on. We will try to set out a development of synthetic geometry starting from Book I of the Elements.

Book I of the Elements begins with 23 definitions, 5 common notions, and 5 postulates. Many of the definitions rely on a number of topological assumptions, as well as an inherent continuity, that we will accept freely. Even with this caveat, most of the definitions would still be considered quite imprecise by modern standards. Even by Aristotle's standards, many of these definitions would have been considered unscientific, and alternatives were proposed by many of Euclid's contemporaries.

Let us start with (a sample of) the 23 definitions.

Definition 1.1. *We define the following geometric concepts*

1. *A point is that which has no part.*
2. *A line is breadthless length.*
3. *The extremities of a line are points.*
4. *A straight line is a line which lies evenly with the points on itself.*
5. *A circle is a plane figure contained by a single line such that all straight lines falling upon it from one point among those lying within the figure are equal to one another*

Note that Euclid considers all curves as lines. Today we would usually reserve the word "line" for a straight line. Indeed, Euclid would consider a circle (the boundary) to be a line. There are also a number of ambiguities or implicit assumptions in these definitions and their later usage. For example, in modern mathematics, we would consider a circle to the the set of points equidistant from a given centre point, and would refer to the interior of this figure as a disc. Euclid instead follows the more common convention of referring to both the disc and its boundary as a circle. His definition of a point does not really distinguish between a single point or a set of discrete points. There is an implicit assumption of "oneness". There is a degree of "obviousness" assumed throughout the elements which does leave many of Euclid's arguments incomplete by modern standards.

His definition of a straight line is arguably one of the most tenuous, and was contested even in his day. Plato proposed the alternative definition of a straight line:

> A straight line is a line whose middle covers its ends.

Proclus gave two alternative definitions:

> A straight line is a line which is fixed by any rotation fixing its endpoints.

> A straight line is a line stretched to the utmost.

The first two of these reflect a nice three dimensional understanding of a straight line, with Plato suggesting that a straight line is a line that looks like a point from the appropriate angle, while Proclus considers rotations in three dimensional space. The last of these is probably the closest to a modern definition of a straight line, as the curve of shortest length connecting two points.

After the definitions, we introduce Euclid's common notions - these are meant to be truths so foundational, that there can be no debate.

Common Notions 1.2.

1. *Things that are equal to the same thing are equal to one another,*
2. *If equals be added to equals, the totals are equal,*
3. *If equals be subtracted from equals, the remainders are equal,*
4. *Things which coincide with one another are equal to one another,*
5. *The whole is greater than the part.*

Only the fourth of these needs comment: as Euclid does not ever consider lengths or numbers explicitly, instead using line segments as stand-ins, this is essentially saying that if two line segments can be moved to coincide with one another, they were equal in length.

Finally we get to the postulates, which would be referred to as axioms in modern language. These are notions that, while less obvious than the common notions, are still meant to be self-evident truths and/or reasonable assumptions.

Postulates 1.3. *Let the following be postulated:*

- *To draw a straight line from any point to any point,*
- *To produce a finite straight line continuously in a straight line,*
- *To describe a circle with any centre and distance,*
- *That all right angles are equal to one another,*
- *That, if a straight line falling on two straight lines makes the interior angles on the same side less than two right angles, the two straight lines - if produced indefinitely - meet on that side on which are the angles less than two right angles.*

Remark 1.4. *Any geometric framework satisfying the first four of these postulates is called an absolute geometry. Hyperbolic geometry will be an example of an absolute geometry. Elliptic geometry will not be, as there will be a maximal circle we can draw on a sphere.*

The fifth postulate, usually called the parallel postulate, stands out. It is not hard to see why this was controversial. It is long and complex, and have the "feel" of a theorem. Many people attempted to propose alternatives, or to prove it, but with little success. Ptolemy attempted to show the alternate angles theorem held and derived the parallel postulate from this. Proclus attempts to show that two lines which do not meet must be at constant distance to one another. Lambert attempts an argument using the internal angles of a triangle and succeeds in showing the parallel postulate is incompatible with spherical geometry. Legendre arguably comes the closest to "proving" the parallel postulate, with an argument relying on similar triangles. Unfortunately, the results about similar triangles needed rely on the parallel postulate, but the argument does lead to a somewhat compelling case for use to replace the parallel postulate with an axiom about similarity.

Nowadays, we usually replace the parallel postulate with a reformulation due to Playfair.

Postulate 1.5. *Call two lines parallel if they do not intersect, even when extended indefinitely in either direction. Given a straight line and a point not contained in the line, there exists exactly one straight line through this point parallel to the first line.*

This is still more complex than the other postulates, but is at least a bit more manageable.

1.1 Proving some propositions

Let us now proceed through the first five propositions of the Elements, seeing what we can derive "only" from the postulates. The arguments presented largely agree with Euclid's, modified slightly for modern palates. In particular, we will use "line" to mean "straight line".

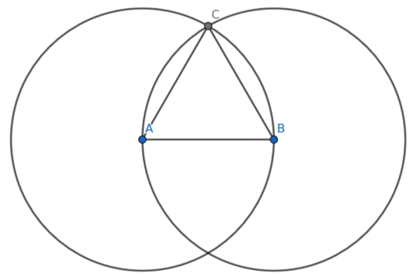
Remark 1.6. *Accompanying diagrams will be added at a later point.*

Proposition 1.7. *Any finite straight line is the base of an equilateral triangle.*

Proof. We will show this by giving a construction for the equilateral triangle. Let AB be a given finite line segment. Using Postulate 3, we construct a circle C_A centred at A of radius $|AB|$, and a circle C_B centred at B of radius $|BA| = |AB|$. These circles must intersect as a point C.

We claim ABC is an equilateral triangle. Indeed we have that $|AC| = |AB|$ by the definition of a circle. Similarly

$$|BC| = |BA| = |AB|$$



by the definition of a circle. By common notion 1, we therefore have

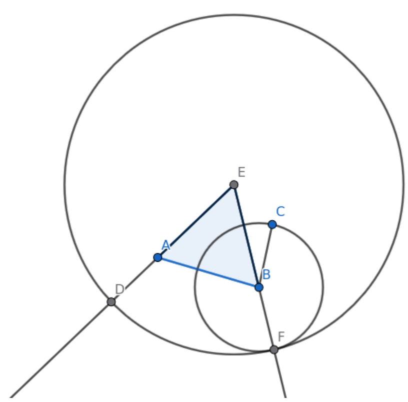
$$|AC| = |BC|$$

and so ABC is an equilateral triangle. □

Remark 1.8. *We already have our first hole in a proof in this proposition. Why do the circles C_A and C_B intersect? t's obvious, right? This issue of intersections existing is a recurring issue in much the the Elements. Usually, it is obvious, but as proofs get more complex, its important to have a justification for why this is so!*

Proposition 1.9. *Given a point A and a straight line BC, there exists a point D such that $|AD| = |BC|$.*

Proof. It suffices to construct such a point D. Using Proposition 1.7, we can construct an equilateral triangle ABE. Using Postulate 3, we construct a circle C_B centred at B of radius $|BC|$. Using Postulate 2, we extend the line EB beyond B to intersect C_B at a point F. Construct a circle C_E centred at E of radius $|EF|$. Extend EA through A to intersect C_E at D. By construction, we



have

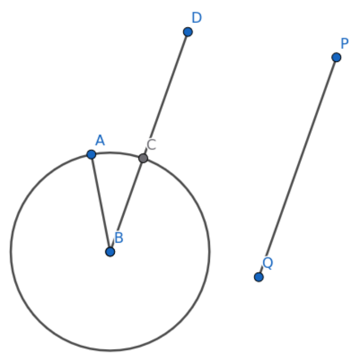
$$|EF| = |ED|, \quad |EA| = |EB|.$$

By the third common notion, we must therefore have $|AD| = |BF|$. As $|BF| = |BC|$ by construction, we therefore get $|AD| = |BC|$. □

Proposition 1.10. *Given two unequal straight lines AB and PQ with $|AB| < |PQ|$, there exists a line CD such that*

$$|AB| + |CD| = |PQ|.$$

Proof. We use Proposition 1.9 to construct a line BD such that $|BD| = |PQ|$. We then construct a circle C_B of radius $|AB|$ centred at B to intersect the line $|BD|$ at C. By construction $|AB| = |BC|$, and we clearly have



$$|PQ| = |BD| = |BC| + |CD| = |AB| + |CD|.$$

□

Proposition 1.11. *Let ABC and DEF be triangles such that $|AB| = |DE|$, $|AC| = |DF|$, and $\angle BAC = \angle EDF$. Then $|EF| = |BC|$, $\angle ABC = \angle DEF$, and $\angle ACB = \angle DFE$. That is, $ABC = DEF$ as triangles.*

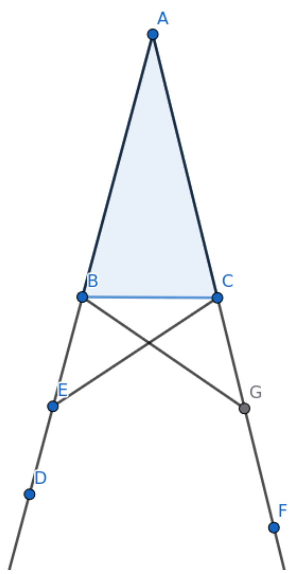
Proof. We apply the triangle ABC to DEF by translating, rotating, and reflecting ABC so that A coincides with D, AB lies along DE, and AC is on the same side of AB as DF. As $|AB| = DE$, B coincides with E (using common notion 4). As $\angle BAC = \angle EDF$, AC lies along DF, and we similarly conclude that C coincides with F. Therefore BC coincides with EF, by Postulate 1 (there is a unique straight line between two points) and so, by common notion 4, $|BC| = |EF|$. In fact, ABC coincides with DEF, from which all claims follow. □

Proposition 1.12. *In an isosceles triangle, the angles at the base are equal to one another.*

Proof. Let ABC be an isosceles triangle with $|AB| = |AC|$. Extend AB to AD and AC to AE, with $|AE| > |AD|$. Take a point F at random on BD and construct G on AE such that $|AF| = |AG|$. Construct the lines BG and CF. As $|AB| = |AC|$ and $|AF| = |AG|$, common notion 3 implies $|BF| = BG|$. We also have that

$$\angle FAC = \angle BAC = \angle GAB$$

and so by Proposition 1.11, we have that $|FC| = |GB|$, and $FAC = GAB$. Hence



$$\angle ACF = \angle ABG, \quad \angle AFC = \angle AGB.$$

Hence, we therefore have $BFC = CGB$ as triangles, by Proposition 1.11. Thus $\angle FBC = \angle GCB$. As $\angle ABF = \angle ACG$, we we subtract to find $\angle ABC = \angle ACB$. □

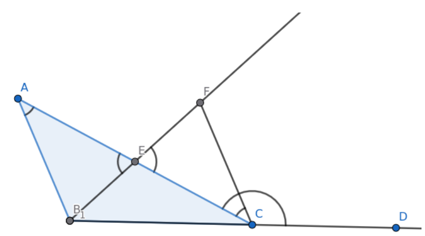
Remark 1.13. *In English education systems, this proposition is sometimes referred to as the pons asinorum or "the bridge of asses". The origins of this nickname are a bit unclear, with some people claiming it comes from the diagram looking a bit lit a bridge. However, the name came to be associated with this proposition being the first difficulty most students encounter, and so students who couldn't master it were deemed to be "asses". Understanding the fifth proposition marked you as a person of normal intelligence.*

Interestingly, this nickname also appeared in the French education system, but for the 47^{th} proposition, the pont des ânes, better known to use as Pythagorus' Theorem. This proposition has acquired many nicknames, including the bridge's throne and the Fransciscan's cowl.

Book I continues in this fashion, establishing standard results about constructions of bisectors and perpendiculars, as well as results about congruent triangles, right angled triangles, opposite angles. We will now skip ahead to the 16^{th} proposition, to sample something more complex.

Proposition 1.14. *The exterior angle in a triangle is greater than either interior angle.*

Proof. Let ABC be a triangle, and extend BC through C to a point D. Let E be the midpoint of AC. Extend BE through E to a point F such that $|BE| = |EF|$. As $|AE| = |EC|$, and opposite angles $\angle AEB$ and $\angle CEF$ are



equal, we have that $AEB = CEF$ as triangles. In particular

$$\angle BAC = \angle BAE = \angle FCE < \angle DCA$$

as claimed. □

Proposition 1.15. *The sum of two internal angles of a triangle is at most two right angles*

Proof. Let ABC be a triangle. Extend BC through C to a point D. By Proposition 1.14,

$$\angle ACD > \angle BAC$$

and hence

$$\angle BCD = \angle BCA + \angle ACD > \angle BCA + \angle BAC.$$

As BCD is a straight line, $\angle BCD$ is equal to two right angles. The claim then follows. □

Remark 1.16. *From here onwards, we will start to measure angles in radians, so that we no longer need to use two right angles as our reference point.*

Skipping further ahead, past the ability to drop a perpendicular from a point, we will consider the 27^{th} proposition.

Proposition 1.17. *27 Given two distinct points, there exists a pair of parallel lines, one through each point.*

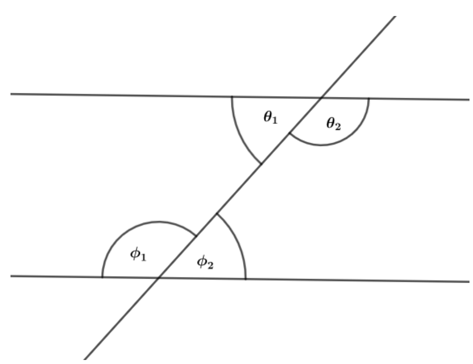
Proof. Construct the line PQ and extend beyond both points as far as is needed. Using early propositions, discussed in the tutorial, we can construct lines through P and Q respectively, perpendicular to PQ. Suppose these lines intersect in R. Then PQR is a triangle with two internal right angles, in contradiction to Proposition 1.15. Thus, these perpendiculars cannot intersect and are therefore parallel. □

Remark 1.18. *This proposition relies only on the first four postulates, and so parallel lines exist in any absolute geometry. It is also the last proposition before the parallel postulate starts to appear!*

The 28^{th} proposition is the first proposition to use the parallel postulate, but is (possibly) one of the first propositions to be proved in the Leaving Cert.

Proposition 1.19. *If a straight line is cut by two parallel lines, they cut off equal alternate interior angles.*

Proof. Let ℓ_1 and ℓ_2 be parallel lines cutting the line ℓ. Let θ_1, θ_2 be the interior angles made by the intersection of ℓ_1 with ℓ. Similarly let ϕ_1, ϕ_2 be the interior angles from the intersection of ℓ_2 with ℓ, with each ϕ_i on the same side of ℓ as the corresponding θ_i. If $\theta_1 \neq \phi_2$, then we can assume $\theta_1 > \phi_2$, without loss of generality. Hence



$$\theta_2 = \pi - \theta_1 < \pi - \phi_2 = \phi_1$$

and so

$$\theta_2 + \phi_2 < \phi_1 + \phi_2 = \pi$$

Postulate 5 then implies that ℓ_1 and ℓ_2 (suitably extended) must intersect on the same side as θ_2 and ϕ_2. This gives a contradiction, and hence we must have $\theta_1 = \phi_2$. □

2 Analytic approaches to geometry

While Riemann probably best gets to lay claim to forging modern geometry, the necessary framework dates back to Descartes and his introduction of coordinates for points on the plane. By assigning a pair (x, y) to every point in the plane in a sensible way, we can suddenly translate questions of geometry into questions of algebra. Suddenly, many geometry problems become "easier", reducing to (often complicated) algebraic manipulations. This switch to coordinates, is how we define many geometries today. However, it is not strictly necessary that we have coordinates for every point. As long as we have a way to define straight lines and circles, and a way to measure angles, we can recreate build a quite reasonable theory of geometry. In this section, we will introduce some of the framework for this, before recasting Euclidean geometry in this language and studying it's properties.

Note 2.1. *From this point onwards, a degree of familiarity with linear algebra over $\mathbb{R}$ is assumed. Please see Appendix A for a hopefully comprehensive overview of the ideas we'll need.*

To define circles, we need a notion of distance. This is captured in the idea of a metric space.

2.1 A review of metric spaces

Definition 2.2. *A metric space is a set X equipped with a function $d : X \times X \to \mathbb{R}$ such that*

1. $d(x, y) \geq 0$ *for all* $x, y \in X$*, with equality if and only if* $x = y$*,*
2. $d(x, y) = d(y, x)$ *for all* $x, y \in X$*,*
3. $d(x, z) \leq d(x, y) + d(y, z)$ *for all* $x, y, z \in X$*.*

We refer to d as a metric or a distance function.

Given a notion of distance, we can define a circle in a metric space as the set of points a fixed distance from a given centre point. We can also (almost) define a notion of a straight line. Strictly speaking, to define a straight line, we need a way to associate meaningful lengths to curves in a metric space, which veers into calculus and differentiable structures. We will return to this in more detail when discussing non-Euclidean geometries, but will give a "pseudo-definition" for now.

Definition 2.3. *A curve in a metric space (X, d) is the image of a continuous function*

$$\gamma : [0, 1] \to X.$$

The points $\gamma(0)$ and $\gamma(1)$ are called the endpoints of the curve. A curve between points x and y in X is a curve whose endpoints are x and y.

Pseudo-definition 2.4. *A straight line between points x and y in a metric space (X, d) is a curve between x and y such that, for any $a < b \in [0, 1]$ and $z \in X$, if*

$$d(\gamma(a), \gamma(b)) = d(\gamma(a), z) + d(z, \gamma(b))$$

then there exists $a < c < b$ such that $z = \gamma(c)$.

When we have a well defined notion of length of a curve, curves of minimal length between two points are called *geodesics*. These play an important role in Riemannian geometry and theoretical physics. A key notion is general relativity is that the kinematics of light is determined by light following a geodesic in space time. However, determining geodesics can often be difficult.

Remark 2.5. *It is possible to give a purely metric definition of a geodesic. A (metric) geodesic is a curve γ such that, for every $t \in (0,1)$, there exists an open interval $I \ni t$ such that*

$$d(\gamma(t_1), \gamma(t_2)) = v_\gamma |t_1, t_2|$$

for all $t_1, t_2 \in I$ and constant v_γ. However, we only need a loose notion of a geodesic, so we will not dwell on this.

Example 2.6. *Here are a number of metric spaces, along with their straight lines.*

- $(\mathbb{R}^n, d_E)$ *with distance defined by the Euclidean norm:*

$$d_E(x,y) = \|x - y\|$$

  This essentially defines Euclidean geometry. The straight lines are exactly lines, and are unique!

- *The unit circle (S^1, d_θ) with distance given by the angle between two points gives a metric space, where a straight line between x and y is the shortest arc between x and y. This is unique unless x and y are diametrically opposite. This means (S^1, d_θ) cannot be an absolute geometry.*

- $(\mathbb{R}^2, d_1)$*, where*

$$d_1\left((x,y),(z,w)\right) = |x - z| + |y - w|$$

  is the Manhattan (or taxicab) metric. This has many shortest paths between a pair of points, with any path consisting of a sequence of horizontal and vertical line segments giving a straight line.

- $(\mathbb{R}^2, d_{IE})$ *where*

$$d_{IE}((x,y),(z,w)) = d_E((x,y),(0,0)) + d_E((z,w),(0,0))$$

  is the Irish Rail metric. We cannot do geometry here, as it has no curves!

Remark 2.7. *Given a metric space (X, d), we can always define another metric by*

$$\tilde{d}(x,y) = \frac{d(x,y)}{1 + d(x,y)}$$

This new metric is bounded, so that $(X, \tilde{d})$ is a metric space of finite diameter. The metrics d and $\tilde{d}$ are equivalent. A sequence in X converges with respect to

d if and only if it converges with respect to $\tilde{d}$, and the metric spaces (X, d) and $(X, \tilde{d})$ have the same Cauchy sequences. However, the geometries associated to them could be distinct. As $\tilde{d}$ is bounded, it cannot define an absolute geometry, while d possible could.

As noted in Remark 2.7, the kinds of geometry we can define on a space depends explicitly on the metric. Homeomorphic spaces could have distinct geometric properties, and so our usual notions of equivalence for metric spaces is not sufficient.

Definition 2.8. *Let (X, d_X) and (Y, d_Y) be metric spaces. We say $f : X \to Y$ is an isometry if f is a bijection and*

$$d_Y(f(x_1), f(x_2)) = d_X(x_1, x_2)$$

for all $x_1, x_2 \in X$.

Isometries will preserve all geometric properties, and so are natural choices for thinking about equivalence of geometries. For example, Euclidean geometry should be translation invariant. Translation by a constant vector v

$$T_v : \mathbb{R}^n \to \mathbb{R}^n x \qquad \mapsto x + v$$

defines an isometry $T_v : (\mathbb{R}^n, d_E) \to (\mathbb{R}^n, d_E)$.

2.2 Euclidean space

With this discussion of isometries in mind, rather than define Euclidean space as $(\mathbb{R}^n, d_E)$, we will instead allow any isometric space.

Definition 2.9. *We say a metric space (X, d) is an n-dimensional Euclidean space if there is an isometry*

$$\Phi : (X, d) \to (\mathbb{R}^n, d_E)$$

i.e. a bijection $\Phi : X \to \mathbb{R}^n$ such that

$$d(x, y) = \|\Phi(x) - \Phi(y)\|$$

for all $x, y \in X$. We call this isometry an affine coordinate system (ACS). We will refer to any n-dimensional Euclidean space with an affine coordinate system as an n-dimensional Euclidean geometry.

Remark 2.10. *Any two n-dimensional Euclidean geometries are isometric. We write $\mathbb{E}^n$ for the isometry class of n-dimensional Euclidean geometries, and will define operations on $\mathbb{E}^n$ by choosing a particular representative.*

We can think of an ACS as defining coordinate axes in X, in particular, fixing an origin.

Example 2.11. • *A translation $T_v : \mathbb{R}^n \to \mathbb{R}^n$, $T_v(x) = x + v$ is an isometry, and so defines an affine coordinate system. We can think of this as making v the new origin.*

- *The reflection $(x, y) \mapsto (x, -y)$ is an isometry $\mathbb{R}^2 \to \mathbb{R}^2$, and defines an affine coordinate system. This changes the orientation of the y-axis.*
- *Given any affine coordinate system $\Phi : X \to \mathbb{R}^n$, the composition $T_v \circ \Phi$ defines a new ACS, shifting our chosen origin.*

2.3 Angles and lines in Euclidean geometry

By choosing an affine coordinate system, we can exploit the vector space structure of $\mathbb{R}^n$ in order to begin defining things like lines and angles in (some) $\mathbb{E}^n$. For example, recall the relationship between the (small) angle between two vectors and their inner product:

$$\langle v, w\rangle = \|v\|\|w\| \cos(\theta_{v,w}).$$

Thus, we can use this to define the angle between two lines in a Euclidean geometry.

Definition 2.12. *Given an affine coordinate system $\Phi : X \to \mathbb{R}^n$, and points $u, v, w \in X$, we define the angle $\angle uvw$ by*

$$\angle uvw = \cos^{-1}\left(\frac{\langle \Phi(u) - \Phi(v), \Phi(w) - \Phi(v)\rangle}{\|\Phi(u) - \Phi(v)\|\|\Phi(w) - \Phi(v)\|}\right).$$

For angle to be a well defined geometric concept, we need that if $\phi : X \to Y$ is an isometry of n-dimensional Euclidean geometries, we have

$$\angle uvw = \angle\phi(u)\phi(v)\phi(w)$$

for all $u, v, w \in X$. It suffices to show that it is independent of choice of ACS for a fixed n-dimensional Euclidean space X.

Lemma 2.13. *Let (X, d) be an n-dimensional Euclidean space and let $u, v, w \in X$. The angle $\angle uvw$ is independent of choice of affine coordinate system.*

Proof. Recall that if $\Phi : X \to \mathbb{R}^n$ is an ACS, then

$$\|\Phi(x) - \Phi(y)\| = d(x, y)$$

for all $x, y \in X$. As such, if we can rewrite the angle in terms of norms, then $\angle uvw$ will depend only on distance in (X, d), not on the choice of ACS. By expanding the right hand side in terms of inner products, we can easily show

$$\begin{aligned}\langle \Phi(u) - \Phi(v), \Phi(w) - \Phi(v)\rangle =& \frac{1}{2}\big(\|\Phi(u) - \Phi(v)\|^2 + \|\Phi(w) - \Phi(v)\|^2 \\ &-\|(\Phi(u) - \Phi(v)) - (\Phi(w) - \Phi(v))\|^2\big) \\ =& \frac{1}{2}(d(u, v)^2 + d(w, v)^2 - d(u, w)^2).\end{aligned}$$

and so $\angle uvw$ depends only on the distance in X, □

Remark 2.14. *This can be used to define angles in any metric space:*

$$\angle uvw = \cos^{-1}\left(\frac{d(u,v)^2 + d(w,v)^2 - d(u,w)^2}{2d(u,v)d(w,v)}\right).$$

This formula is also quite familiar to us! Rearranging and writing

$$a = d(u,v), \quad b = d(w,v), \quad c = d(u,w), \quad \theta = \angle uvw$$

we find our formula is equivalent to cosine rule for triangles:

$$c^2 = a^2 + b^2 - 2bc\cos(\theta).$$

Having defined angles, the next key component of geometry to define is straight lines. As noted in Example 2.6, straight lines in $(\mathbb{R}^n, d_E)$ are straight lines as we usually think of them. We know that every such line can be described as a set

$$\ell = \{a + \lambda v \mid \lambda \in \mathbb{R}\}$$

for some $a, v \in \mathbb{R}^n$. A reasonable definition of a line in an n-dimensional Euclidean geometry is therefore defined in terms of such lines.

Definition 2.15. *A subset $A \subset \mathbb{R}^n$ is called an affine subspace of dimension k if there exists a vector subspace $V \subset \mathbb{R}^n$ of dimension k and a vector $a \in \mathbb{R}^n$ such that*

$$A = \{a + v \mid v \in V\}.$$

Example 2.16. *An affine subspace of dimension 1, usually called an affine line, is precisely a line*

$$\ell = \{a + \lambda v \mid \lambda \in \mathbb{R}\}$$

Definition 2.17. *Given a choice of affine coordinate system $\Phi : X \to \mathbb{R}^n$, a line in (X, d) is the inverse image $\Phi^{-1}(\ell)$ of an affine line $\ell \subset \mathbb{R}^n$.*

It remains to be seen if this is a geometric definition of a line, as this currently depends explicitly on a choice of ACS. A line for a given Φ may not be a line for a different choice of ACS. We will handle this momentarily.

Definition 2.18. *The points $x_1, \ldots, x_m$ in an n-dimensional Euclidean geometry (X, d) with ACS $\Phi : X \to \mathbb{R}^n$ are called collinear if they lie on a single line.*

Remark 2.19. *Any two points are collinear. The points x and y lie on the line ℓ determined by*

$$\Phi(\ell) = \{\Phi(x) + \lambda(\Phi(y) - \Phi(x)) \mid \lambda \in \mathbb{R}\}.$$

In order to ensure that collinearity is geometric, we need to show that it is independent of a choice of ACS, depending only on distances in X.

Proposition 2.20. *Three points $v_1, v_2, v_3 \in \mathbb{R}^n$ are collinear if and only if*

$$\|v_1 - v_2\| + \|v_2 - v_3\| = \|v_1 - v_3\|$$

up to possible relabelling the points.

Proof. If the points are collinear, then there exist $a, w \in \mathbb{R}^n$ and $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ such that

$$v_i = a + \lambda_i w$$

and we compute

$$\|v_1 - v_2\| + \|v_2 - v_3\| = (|\lambda_1 - \lambda_2| + |\lambda_2 - \lambda_3|)\,\|w\|.$$

By relabelling the points if needed, we can assume

$$\lambda_1 \geq \lambda_2 \geq \lambda_3$$

and so we can remove the absolute values to obtain

$$\begin{aligned}\|v_1 - v_2\| + \|v_2 - v_3\| &= (\lambda_1 - \lambda_3)\,\|w\| \\ &= |\lambda_1 - \lambda_3|\|w\| \\ &= \|v_1 - v_3\|.\end{aligned}$$

Conversely, if the equality holds, the Exercise 8.17 implies

$$v_2 = v_1 + \lambda_0(v_3 - v_1)$$

for some real λ_0. Thus v_1, v_2, v_3 lie on the line

$$\ell = \{v_1 + t(v_3 - v_1) \mid t \in \mathbb{R}\}.$$

□

Corollary 2.21. *Collinearity in (X, d) is independent of choice of affine coordinate system.*

This already implies that lines are independent of a choice of ACS! As such, we can give a purely algebraic proof of one of Euclid's early propositions.

Lemma 2.22. *Given a non-degenerate triangle in $\mathbb{R}^n$, the longest side is opposite the largest angle.*

Proof. Let u, v, w be the vertices of the triangle. We will write

$$a = d(u, v), \quad b = d(v, w), \quad c = d(w, u).$$

It will suffice to show that if $a > b$, then $\angle uwv > \angle vuw$. We have that

$$\angle uwv = \cos^{-1}\left(\frac{b^2 + c^2 - a^2}{2bc}\right), \quad \angle vuw = \cos^{-1}\left(\frac{a^2 + c^2 - b^2}{2ac}\right).$$

As the function $\cos^{-1}$ is decreasing on $[-1, 1]$, it is enough to show

$$\frac{b^2 + c^2 - a^2}{2bc} < \frac{a^2 + c^2 - b^2}{2ac}.$$

Rearranging this, this is equivalent to

$$0 < a^2b - ab^2 - ac^2 + bc^2 + a^3 - b^3$$

We can pull out a factor of $(a - b)$ from the right hand side, and so it is enough to show

$$0 < (a - b)((a + b)^2 - c^2)$$

As $a > b$, $(a - b) > 0$. As $a + b > c$, $(a + b)^2 - c^2 > 0$. Hence this inequality holds, and the claim follows. □

We might also be interested in affine subspaces of higher dimension, and how we can discuss these spaces in a general Euclidean space. For example, we can show that, given an ACS, a non-degenerate triangle defines a plane.

Definition 2.23. *A plane in $\mathbb{R}^n$ is an affine subspace of dimension 2. A plane in a Euclidean space (X, d) is a subspace P such that $\Phi(P)$ is a plane in $\mathbb{R}^n$.*

Lemma 2.24. *A non-degenerate triangle in $\mathbb{R}^n$ is contained in a unique plane.*

Proof. Let a, b, c be the vertices of the triangle. Then a, b, c are contained in the plane

$$P = \{a + s(b - a) + t(c - a) \mid s, t \in \mathbb{R}\}$$

Clearly, the lines

$$\ell_{a,b} = \{a + s(b - a) \mid s \in \mathbb{R}\},$$
$$\ell_{a,c} = \{a + t(c - a) \mid t \in \mathbb{R}\}$$

are contained in P. The line

$$\begin{aligned}\ell_{b,c} &= \{b + \lambda(c - b) \mid \lambda \in \mathbb{R}\}\\ &= \{a + (b - a) + \lambda(c - a - b + a) \mid \lambda \in \mathbb{R}\}\\ &= \{a + (1 - \lambda)(b - a) + \lambda(c - a) \mid \lambda \in \mathbb{R}\}\end{aligned}$$

is similarly contained within P. The uniqueness of P is left to the reader. □

Corollary 2.25. *A non-degenerate triangle in a Euclidean space (X, d) with affine coordinate system Φ is contained within a unique plane*

2.4 Affine transformations and isometries

Rather than trying to show that all of our definitions of geometric ideas in $\mathbb{E}^n$ are truly geometry (independent of choice of ACS for a given representative (X, d)) by expressing every property in terms of distance, it would be easier to have some way to compare difference affine coordinate systems.

Suppose (X, d) is an n-dimensional Euclidean space with two affine coordinate systems

$$\Phi_1, \Phi_2 : X \to \mathbb{R}^n.$$

Then, the composition

$$\Phi_1 \circ \Phi_2^{-1} \mathbb{R}^n \to \mathbb{R}^n$$

is an isometry. Thus to describe all affine coordinate systems for X, it suffices to describe all isometries $\mathbb{R}^n \to \mathbb{R}^n$. Such isometries are often called *rigid motions.*

Example 2.26. *Some examples of rigid motions in $\mathbb{R}^3$ include rotations around an axis, reflection in a plane, and translation by a fixed vector. Note that all of these preserve collinearity, angles, and intersections. Indeed, these are all transformations that were understood to preserve geometric statements in classical Euclidean geometry.*

Remark 2.27. *Rigid motions are generally not linear maps. Can you suggest a few that are linear?*

Any property invariant under a rigid motion is independent of choice of affine coordinate system, and so understanding these will be our next major goal.

Note 2.28. *From this point, some basic knowledge of group theory is assumed. Primarily, the definition of a group and a number of examples. The necessary background is outlined in Appendix B.*

Lemma 2.29. *Given a metric space (X, d) the set of isometries $X \to X$ forms a group under composition*

Proof. Composition is clearly associative and the composition of two isometries is an isometries:

$$\begin{aligned} d(\Psi(\Phi(x)), \Psi(\Phi(y))) &= d(\Phi(x), \Phi(y)) \\ &= d(x, y). \end{aligned}$$

The map $I : x \mapsto x$ is an isometry and $\Phi \circ I = I \circ \Phi = \Phi$. Every isometry is a bijection, so there exists an inverse map (of sets), and this inverse is an isometry:

$$\begin{aligned} d(\Phi^{-1}(x), \Phi^{-1}(y)) &= d(\Phi(\Phi^{-1}(x)), \Phi(\Phi^{-1}(y))) \\ &= d(x, y). \end{aligned}$$

□

Definition 2.30. *Denote the groups of isometries of* (X, d) *by* $\text{Isom}(X, d)$*. If the metric* d *has been specified already, we will often omit it.*

Lemma 2.31. *Let* (X, d) *be a Euclidean space with affine coordinate system* $\Phi : X \to \mathbb{R}^n$*. Then there is an isomorphism*

$$\text{Isom}(X) \cong \text{Isom}(\mathbb{R}^n)$$

Proof. If $f \in \text{Isom}(\mathbb{R}^n)$, then it is easy to check that $\Phi^{-1} f \Phi \in \text{Isom}(X)$. Similarly, if $F \in \text{Isom}(X)$, then $\Phi F \Phi^{-1} \in \text{Isom}(\mathbb{R}^n)$. These are clearly inverse, and can be seen to be homomorphisms. □

Thus, up to isomorphism, there is a single isometry group of n-dimensional Euclidean geometry, and we therefore write $\text{Isom}(\mathbb{E}^n) = \text{Isom}(\mathbb{R}^n)$ for this group, and will abusively speak of isometries of $\mathbb{E}^n$ and their action on points of $\mathbb{E}^n$.

Lemma 2.32. *Isometries of* $\mathbb{E}^n$ *map lines to lines.*

Proof. Fix a geometry. Let ℓ be a line, A, B be distinct points on ℓ, and let C be a point between A and B. Let $\Phi \in \text{Isom}(\mathbb{E}^n)$. As A, B, C are collinear, we have that

$$d(A, B) = d(A, C) + d(C, B).$$

As Φ is an isometry,

$$\begin{aligned} d(\Phi(A), \Phi(C)) + d(\Phi(C), \Phi(B)) &= d(A, C) + d(C, B) \\ &= d(A, B) \\ &= d(\Phi(A), \Phi(B)). \end{aligned}$$

Hence $\Phi(C)$ is on the line segment $\Phi(A)\Phi(B)$.

Conversely, suppose $\tilde{C}$ is on the line segment $\Phi(A)\Phi(B)$. Analogously $C = \Phi^{-1}(\tilde{C})$ will be a point on the line segment AB. Thus Φ maps the line segment AB to the line segment $\Phi(A)\Phi(B)$.

For C not between A and B, we make an analogous argument, instead using either

$$d(A, C) = d(A, B) + d(B, C), \quad \text{or} \quad d(C, B) = d(C, A) + d(A, B)$$

as the collinearity condition. Thus, Φ will map the line segment AC (CB resp.) containing B (A resp.) to the line segment $\Phi(A)\Phi(C)$ ($\Phi(C)\Phi(B)$ resp.) containing $\Phi(B)$ ($\Phi(A)$ resp.). As these line segments must extend AB and $\Phi(A)\Phi(B)$, we conclude Φ maps ℓ to a line. □

Corollary 2.33. *Let* $\Phi, \Psi \in \text{Isom}(\mathbb{E}^n)$ *and* $x, y \in \mathbb{E}^n$ *be distinct points. If*

$$\Phi(x) = \Psi(x), \quad \text{and} \quad \Phi(y) = \Psi(y)$$

then $\Phi(z) = \Psi(z)$ *for all* z *on the line between* x *and* y*.*

Proof. Let $X = \Phi(x)$ and $Y = \Phi(y)$. Suppose first that z is on the line segment xy. From the proof of Lemma 2.32, both $\Phi(z)$ and $\Psi(z)$ must lie on the line segment XY. Without loss of generality, we assume $\Phi(z)$ is closer to X than $\Psi(z)$. By collinearity of the points $X, \Phi(z), \Psi(z), Y$, we must have

$$\begin{aligned} d(x,y) = d(X,Y) &= d(X,\Phi(z)) + d(\Phi(z),\Psi(z)) + d(\Psi(z),Y) \\ &= d(\Phi(x),\Phi(z) + d(\Phi(z),\Psi(z)) + d(\Psi(z),\Psi(y)) \\ &= d(x,z) + d(\Phi(z),\Psi(z)) + d(z,y) \\ &= d(x,y) + d(\Phi(z),\Psi(z)) \end{aligned}$$

and so $d(\Phi(z),\Psi(z)) = 0$. Hence $\Phi(z) = \Psi(z)$. A similar argument holds for z not between x and y, and the claim follows. □

This says that if two isometries agree on a pair of points, they agree on all points between them. We can in fact extend this to show that an isometry is entirely determined by the image of three (generic) points in two dimensions.

Theorem 2.34. *If $\Phi, \Psi \in \mathrm{Isom}(\mathbb{E}^2)$ agree on 3 non-collinear points, they are equal.*

Proof. Let x, y, z be the vertices of a non-degenerate triangle such that Φ and Ψ agree on x, y, z. By Corollary 2.33, Φ and Ψ agree on the edges of the triangle and their extensions. Pick a point p not contained in one of the triangle edges. As x, y, z are not collinear, there exists a line through p intersecting two distinct sides of the triangle, extended appropriately - for example, a line through p in the plane defined by x, y, z parallel to one of the edges. Without loss of generality, we assume this line intersects the lines xy and xz, extended appropriately, in the points u and v respectively. As the isometries agree on the edges, we have

$$\Phi(u) = \Psi(u), \quad \text{and} \quad \Phi(v) = \Psi(v).$$

Hence Φ and Ψ agree on all points on the line between u and v. In particular $\Phi(p) = \Phi(v)$. Hence they agree everywhere. □

Corollary 2.35. *If $\Phi \in \mathrm{Isom}(\mathbb{E}^2)$ fixes three non-collinear points, it fixes all points.*

Let us now prove our examples of isometries are indeed isometries.

Lemma 2.36. *Translations, and multiplication by $A \in \mathrm{O}_n(\mathbb{R})$ define isometries of $\mathbb{E}^n$.*

Proof. It is enough to show these define isometries of $\mathbb{R}^n$. A translation $T_v(x) = x + v$ is clearly an isometry, as it has an inverse $T_v^{-1}(y) = y - v$, and

$$d(T_v(x), T_v(y)) = \|T_v(x) - T_v(y)\| = (\|x + v - y - v\| = \|x - y\| = d(x,y).$$

For a matrix $A \in \mathrm{O}_n(\mathbb{R})$, we note that it is invertible by definition and

$$\begin{aligned} d(Ax, Ay)^2 &= \|Ax - Ay\|^2 = \|A(x-y)\|^2 \\ &= \langle A(x-y), A(x-y) \rangle \\ &= \langle A^T A(x-y), x-y \rangle \\ &= \langle x-y, x-y \rangle = \|x-y\|^2 = d(x,y)^2 \end{aligned}$$

□

Translations, and multiplication by orthogonal matrices (and their compositions) are examples of affine transformations.

Definition 2.37. *A map $f : \mathbb{R}^n \to \mathbb{R}^n$ is called an affine transformation if $f(x) = Ax + v$ for some $A \in \mathrm{GL}_n(\mathbb{R})$ and $v \in \mathbb{R}^n$.*

Exercise 2.38. *Show that the set of affine transformations forms a group under composition.*

Affine transformations are usually not linear, but can be classified by their linear-ish behaviour.

Lemma 2.39. *For a bijection $F : \mathbb{R}^n \to \mathbb{R}^n$, the following are equivalent:*

1. *F is an affine transformation*
2. *$F(ax + by) - F(0) = a(F(x) - F(0)) + b(F(y) - F(0))$ for all $a, b \in \mathbb{R}$, $x, y \in \mathbb{R}^n$.*
3. *$F((1-\lambda)x + \lambda y) = (1-\lambda)F(x) + \lambda F(y)$ for all $\lambda \in \mathbb{R}$, $x, y \in \mathbb{R}^n$.*

Proof. Assume 1) holds. The $F(x) = Ax + v$ for some A and v. Hence

$$\begin{aligned} F(ax+by) - F(0) &= A(ax+by) + v - A(0) - v \\ &= a(Ax) + b(Ay) \\ &= a(Ax + v - v) + b(Ay + v - y) = a(F(x) - F(0)) + b(F(y) - F(0)). \end{aligned}$$

Therefore 2) holds.

Assume 2) holds. Then

$$\begin{aligned} F((1-\lambda)x + \lambda y) &= F(0) + (1-\lambda)(F(x) - F(0)) + \lambda(F(y) - F(0)) \\ &= F(0) + (1-\lambda)F(x) + \lambda F(y) - F(0) \\ &= (1-\lambda)F(x) + \lambda F(y). \end{aligned}$$

Therefore 3) holds.

Assume 3) holds. Let $v = F(0)$. It would suffice to show that the map

$$G(x) := F(x) - F(0)$$

is a linear map, as it is clearly invertible if F is. We first note that

$$\begin{aligned} G(\lambda x) &= F(\lambda x + (1-\lambda)0) - F(0) \\ &= \lambda F(x) + (1-\lambda)F(0) - F(0) \\ &= \lambda\left(F(x) - F(0)\right) = \lambda G(x) \end{aligned}$$

for all $\lambda \in \mathbb{R}$. Next note that

$$\begin{aligned} G(x+y) &= F(x+y) - F(0) \\ &= F\left(\frac{1}{2}(2x) + \frac{1}{2}(2y)\right) - F(0) \\ &= \frac{1}{2}F(2x) + \frac{1}{2}F(2y) - F(0) \\ &= \frac{1}{2}\left(F(2x) - F(0) + F(2y) - F(0)\right) \\ &= \frac{1}{2}\left(G(2x) + G(2y)\right) = G(x) + G(y). \end{aligned}$$

Hence G is linear, as needed. □

With this description of affine transformations, we can accurately classify all isometries of $\mathbb{E}^n$.

Theorem 2.40. *The isometries of $\mathbb{E}^n$ are exactly the affine transformations whose matrix part lies in $\mathrm{O}_n(\mathbb{R})$.*

Proof. We have already seen that all affine transformations with orthogonal matrix part are isometries, so it suffices to show the opposite inclusion. By Lemma 2.39, it suffices to show that, for any isometry $F \in \mathrm{Isom}(\mathbb{R}^n)$, we have

$$F((1-\lambda)x + \lambda y) = (1-\lambda)F(x) + \lambda F(y)$$

for all $\lambda \in \mathbb{R}$, $x, y \in \mathbb{R}^n$. If $x = y$, this is trivial. Similarly, if $\lambda = 0$ or $\lambda = 1$. Otherwise, let $z = (1-\lambda)x + \lambda y$. If $0 < \lambda < 1$, then z is on the line segment xy and so, from the proof of Lemma 2.32, $F(z)$ is on the line segment $F(x)F(y)$. Hence, there exists $0 < \mu < 1$ such that

$$F(z) = (1-\mu)F(x) + \mu F(y).$$

We can compute

$$\|x - z\| = \|\lambda x - \lambda y\| = \lambda\|x - y\|$$

and

$$\|F(x) - F(z)\| = \|\mu F(x) - \mu F(y)\| = \mu\|F(x) - F(y)\|.$$

As F is an isometry,

$$\|F(x) - F(z)\| = \|x - z\|, \quad \text{and} \quad \|F(x) - F(y)\| = \|x - y\|.$$

Hence $\lambda = \mu$, and

$$F((1-\lambda)x + \lambda y) = (1-\lambda)F(x) + \lambda F(y)$$

If $\lambda > 1$, we instead write

$$y = \frac{1}{\lambda}z + \left(1 - \frac{1}{\lambda}\right)x$$

to consider y as on the line segment xz and repeat the argument. Similarly, if $\lambda < 0$, we consider x on the line segment yz with

$$x = \left(1 - \frac{-\lambda}{1-\lambda}\right)z + \left(\frac{-\lambda}{1-\lambda}\right)y.$$

Hence F is affine - $F(x) = Ax + v$. As F is an isometry, and translation by v is an isometry, the map

$$x \mapsto Ax$$

is an isometry. Hence

$$\begin{aligned}\langle Ax, Ay\rangle &= \frac{1}{4}\left(\|A(x+y)\|^2 - \|A(x-y)\|^2\right)\\ &= \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right) = \langle x, y\rangle\end{aligned}$$

for all $x, y \in \mathbb{R}^n$. We therefore have that

$$\langle A^T Ax, y\rangle = \langle Ax, Ay\rangle = \langle x, y\rangle$$

for all $x, y \in \mathbb{R}^n$. This implies $A^T A = I$, and so $A \in \mathrm{O}_n(\mathbb{R})$. □

Exercise 2.41. *Using Theorem 2.40, show that affine lines (or more generally affine spaces) can be defined in n-dimensional Euclidean space independent of a choice of affine co-ordinate system. It suffices to show that the image of an affine line/space under an isometry of* $\mathrm{Isom}(\mathbb{R}^n)$ *is an affine line/space.*

In two and three dimensions $\mathrm{O}_n(\mathbb{R})$ describes exactly the set of rotations, and the set of reflections followed by a rotation. Hence, at least in a visualisable number of dimensions, rigid motions are exactly those obtained by reflecting, rotating, and translating.

Definition 2.42. *We call an isometry of $\mathbb{E}^n$ orientation preserving if, as an affine transformation $Ax + b$,* $\det(A) > 0$. *Otherwise, it is called orientation reversing.*

Remark 2.43. *Recall that* $\det(A) = \pm 1$ *for all orthogonal matrices, so really we say a transformation is orientation preserving if its determinant is* 1*, and orientation reversing if its determinant is* -1.

Example 2.44. *Some examples of planar isometries and their effects on orientation include:*

- *Translations are orientation preserving.*
- *Rotations around a point in the plane are orientation preserving.*
- *If we define reflection in a line ℓ as follows, we obtain an orientation reversing isometry:*
  - *If p is on ℓ, a reflection maps p to p,*
  - *If p is not on ℓ, construct a line $\tilde{\ell}$ through p perpendicular to ℓ, intersecting ℓ at q. Then p is mapped to the unique other point $\tilde{p}$ on $\tilde{\ell}$ such that $d(p, q) = d(\tilde{p}, q)$.*

Exercise 2.45. *It is not immediately obvious that reflection, as defined, is an isometry. By constructing some congruent triangles, show that it is.*

Reflections, and their generalisations to higher dimensions, form an important class of isometries. They are examples of what are called *involutions*: function $f : X \to X$ such that $f(f(x)) = x$ for all $x \in X$. We will also see that every isometry can be expressed in terms of reflections.

Lemma 2.46. *Let ABC be a non-degenerate triangle in $\mathbb{R}^2$, $\Phi \in \mathrm{Isom}(\mathbb{R}^2)$, and let $DEF = \Phi(ABC)$ be the image of the triangle. Then there is a set of three isometries $r_1, r_2, r_3 \in \mathrm{Isom}(\mathbb{R}^2)$, each of which is a reflection or the identity, such that*

$$r_3 r_2 r_1(ABC) = DEF$$

Theorem 2.47. *The group $\mathrm{Isom}(\mathbb{E}^2)$ is generated by reflections.*

Proof. It is enough to show it for $\mathrm{Isom}(\mathbb{R}^2)$. Let $f \in \mathrm{Isom}(\mathbb{R}^2)$, and pick any non-degenerate ABC. Then, by Lemma 2.46, we can find reflection or identity maps $r_1, r_2, r_3 \in \mathrm{Isom}(\mathbb{R}^2)$ such that

$$f(A) = r_3 r_2 r_1(A),\ f(B) = r_3 r_2 r_1(B),\ f(C) = r_3 r_2 r_1(C).$$

Hence, by Theorem 2.34, $f = r_3 r_2 r_1$. Thus every isometry can be expressed in terms of at most three reflections. □

Remark 2.48. *This implies that to show that some geometry concept, such as an ellipse, can be defined in any 2-dimensional Euclidean space, it is enough to show that reflections act appropriately on it e.g. the reflection of an ellipse in any line is still an ellipse, so ellipses are a purely "geometric" idea.*

Remark 2.49. *Theorem 2.47 extends to $\mathrm{Isom}(\mathbb{E}^n)$, with every isometry expressible in terms of at most $(n + 1)$ reflections in a codimension 1 hyperplane. Note also that none of the geometry involved used any notion of parallel lines, so this generation property holds in any absolute geometry on $\mathbb{R}^n$.*

Example 2.50. *How could we write the translation $T_v(x) = x + v$ in terms of reflections? Consider the triangle*

$$A = 0,\ B = \frac{1}{2}v,$$

and C any other point on the line ℓ_1 perpendicular to AB through B. We can write $C = \frac{1}{2}v + w$ with $\langle v, w\rangle = 0$. As $T_v(A) = v$, we take r_1 to be reflection in the line ℓ_1:

$$r_1(A) = v,\ r_1(B) = B,\ r_1(C) = C.$$

The translations of B is $B' = T_v(B) = \frac{3}{2}v$. If we take r_2 to be the reflection in the line ℓ_2 perpendicular to BB', through the midpoint v, we find

$$r_2r_1(A) = v, ; r_2r_1(B) = \frac{3}{2}v,\ r_2r_1(C) = ???$$

As ℓ_2 is parallel to ℓ_1, $r_2r_1(C)) = r_2(C)$ is a point on the line perpendicular to ℓ_1 through C. This line is parallel to AB, and so by a simple rectangle argument, we must have $r_2(C) = \frac{3}{2}v + w = T_v(C)$. Hence $T_v = r_2r_1$.

We round out our discussion of analytic Euclidean geometry with a neat result, proved using both analytic and synthetic geometry for comparison.

Definition 2.51. *Given a line with an orientation, we define the signed length of the segment between two points A, B on this line by $|AB|$ is given by the distance if B is on the positive side of A with respect to the orientation, or by the negative of the distance if B is on the negative side of A*

Theorem 2.52 (Ceva's Theorem)**.** *Given a triangle ABC and a point O not on the sides, let D, E, F be the intersections of (extensions of) AO, BO, CO with BC, AC, AB respectively. Then (even as signed lengths)*

$$\frac{|BD|}{|DC|}\frac{|CE|}{|EA|}\frac{|AF|}{|FB|} = 1.$$

Proof. We can assume we are working in $\mathbb{R}^2$, and that O is the origin by translation. Hence A, B, C, D, E, F are given by vectors a, b, c, d, e, f. By construction, d lies on the line $\{ta \mid t \in \mathbb{R}\}$ and the line

$$\{(1-\lambda)b + \lambda c \mid \lambda \in \mathbb{R}\}$$

Hence

$$d = t_d a = b + \lambda_d(c - b)$$

for some $t_d, \lambda_d \in \mathbb{R}$ satisfying this equality. We then have that

$$|BD| = \|b - d\| = |\lambda_d|\|c - b\|,$$
$$|DC| = \|c - d\| = |(1 - \lambda_d)\|c - b\|,$$

and hence

$$\frac{|BD|}{|DC|} = \left|\frac{\lambda_d}{1 - \lambda_d}\right|.$$

Similarly,

$$e = t_e b = c + \lambda_e(a - c), \quad \frac{|CE|}{|EA|} = \left|\frac{\lambda_e}{1 - \lambda_e}\right|$$
$$f = t_f c = a + \lambda_e(b - a), \quad \frac{|AF|}{|FB|} = \left|\frac{\lambda_f}{1 - \lambda_f}\right|$$

Let us now try to eliminate b from our simultaneous vector equations. We find

$$\frac{t_f}{\lambda_f}c - \frac{1 - \lambda_f}{\lambda_f}a = b = \frac{-\lambda_d}{1 - \lambda_d}c + \frac{t_d}{1 - \lambda_d}a.$$

As our triangle is non-degenerate, a and c are linearly independent, and so we can equate coefficients:

$$\frac{t_f}{\lambda_f} = -\frac{\lambda_d}{1 - \lambda_d}, \quad -\frac{1 - \lambda_f}{\lambda_f} = \frac{t_d}{1 - \lambda_d},$$

the second of which can be rearranged to give

$$\frac{\lambda_d}{t_d} = \frac{-\lambda_d \lambda_f}{(1 - \lambda_d)(1 - \lambda_f)}.$$

Similarly eliminating c leads to

$$\frac{t_d}{\lambda_d} = -\frac{\lambda_e}{1 - \lambda_e}.$$

Thus

$$\frac{\lambda_d \lambda_e \lambda_f}{(1 - \lambda_d)(1 - \lambda_e)(1 - \lambda_f)} = \frac{-t_d}{\lambda_d} \times \frac{-\lambda_d}{t + d} = 1.$$

The absolute value of the left hand side is precisely the product

$$\frac{|BD|}{|DC|} \frac{|CE|}{|EA|} \frac{|AF|}{|FB|}$$

and so the result follows. In fact, it holds for signed lengths, as the equality holds without absolute values. □

In the analytic proof, the fact that Ceva's theorem holds for signed lengths is a simple consequence of the proof method For the synthetic proof, it is also built into the proof, arising from the two the cases of O inside the triangle and O outside the triangle simultaneously

Proof. We first note that, if O is inside the triangle, all the signed lengths are positive, which if O is outside the triangle, there will be an even number of negative lengths, so the signs will cancel. Thus, the claim will hold for signed lengths if it holds for unsigned lengths.

Let $\triangle ABC|$ denote the area of the triangle ABC. Note that COD and BOD have a common perpendicular height h to the bases CD and BD, and so

$$\frac{|\triangle BOD|}{\triangle COD|} = \frac{\frac{1}{2}|BD|h}{\frac{1}{2}|CD|h} = \frac{|BD|}{|DC|}.$$

Similarly

$$\frac{|\triangle BAD|}{|\triangle CAD|} = \frac{||BD|}{|DC|}.$$

We have that

$$\begin{aligned} |\triangle AOB| = |\triangle BAD| - |\triangle BOD| &= \frac{|BD|}{|DC|}\left(|\triangle CAD| - |\triangle COD|\right) \\ &= \frac{|BD|}{|DC|}|\triangle COA|. \end{aligned}$$

Hence

$$\frac{|\triangle AOB|}{|\triangle COA|} = |BD||DC.$$

Similarly

$$\frac{|\triangle COA|}{|\triangle BOC|} = |AF||FB, \quad \text{and} \quad \frac{|\triangle BOC|}{|\triangle AOB|} = |CE||EA.$$

Hence

$$\frac{|BD|}{|DC|}\frac{|CE|}{|EA|}\frac{|AF|}{|FB|} = [\frac{|\triangle AOB||\triangle BOC||\triangle COA|}{|\triangle COA||\triangle AOB||\triangle BOC|} = 1$$

□

3 Polygons and triangulations

Definition 3.1. *A polygon P is a finite collection of (at least three) points $V(P)$ in the plane, called vertices, and straight lines $E(P)$, called edges such that*

- *The endpoints of each edge are vertices. We call say the endpoints of an edge are adjacent,*
- *Each vertex is contained in exactly two edges,*
- *No two edges intersect outside of a vertex,*
- *Given edges e_1, e_2 with endpoints $\{v_1, v_2\}$ and $\{v_2, v_3\}$ respectively, v_1, v_2, v_3 are not collinear,*
- *All vertices are connected: given $v, w \in V(P)$, there exists a sequence of edges $e_1, \ldots, e_n \in E(P)$ such that e_i, e_{i+1} share a common vertex, v is an endpoint of e_1 and w is an endpoint of e_n*

We call a polygon with n vertices an n-gon.

Remark 3.2. *It is convenient to identify an edge e with its endpoint $v - w$. With this notation, we can given the connected condition as there exists a sequence of vertices $v, v_1, \ldots, v_n, w$ such that $v - v_1, v_1 - v_2, \ldots, v_n - w$ are edges of P. Also note that a polygon has an equal number of edges as vertices.*

Fact 3.3. *A polygon P bounds a set $I(P)$ in the plane, called the interior. We will write*

$$\overline{P} = I(P) \cup E(P) \cup V(P)$$

when we want to refer to the closed figure bounded by the polygon.

Example 3.4. *Here are some examples of polygons and non-polygons*
Use your imagination for now

The simplest example of a polygon is a triangle, and by breaking a polygon up into triangles, we can reduce many mathematical questions to questions about triangles.

Definition 3.5. *A triangulation $\mathcal{T}$ of a polygon P is a collection of triangles such that*

- $\bigcup_{T \in \mathcal{T}} V(T) \supseteq V(P)$,
- $\bigcup_{T \in \mathcal{T}} E(T) \supseteq E(P)$,
- $\bigcup_{T \in \mathcal{T}} \overline{T} = \overline{P}$,
- *For any distinct $T_1, T_2 \in \mathcal{T}$, $I(T_1) \cap I(T_2) = \emptyset$.*

We call a triangulation strong *if*

$$\bigcup_{T \in \mathcal{T}} V(T) = V(P)$$

Example 3.6. *Here are some examples of triangulations. All except the second are strong.*

Use your imagination for now

Remark 3.7. *To cut down on notation, we will occasionally use $V(\mathcal{T})$ to refer to the union of all vertices in the triangulation, $E(\mathcal{T})$ for the union of all edges in the triangulation, etc.*

To reduce to studying triangles, we need to be certain that a triangulation always exists, and ideally how to construct one. We give two methods to do so: one is mathematically convenient, while the other is more practical.

Definition 3.8. *A diagonal of a polygon is a straight line d connecting two non-adjacent vertices such that d is contained entirely within $I(P)$ (other than its endpoints on the boundary of P).*

By taking a diagonal of a polygon, we dissect in into two polygons with fewer sides. Assuming we can find a diagonal, we can therefore take iterated diagonals to construct a triangulation. Unfortunately, we cannot just take two non-adjacent vertices at random, and need to put a bit of though into finding a diagonal.

Lemma 3.9. *Every polygon with at least 4 vertices has a diagonal.*

Proof. We first want to find a vertex with an interior angle less than π. Fix an orientation of the plane, and pick a leftmost vertex v. This has neighbours u and v. At least one of u or w is strictly to the right of v, as v is leftmost and u, v, w are not collinear. Hence $\angle uvw < \pi$. If $u - w$ forms a diagonal, we are done. Otherwise the boundary of P must cross to the left of $u - w$, and so there is a vertex of P in the triangle uvw. Let z be the leftmost such vertex. We claim $v - z$ is a diagonal. If not, the boundary of P must cross to the left of $v - z$, and so there is a vertex in the triangle $vzw \subset uvw$, contradicting the leftmost-ness of z □

Theorem 3.10. *Every polygon P with $n \geq 3$ vertices has a strong triangulation into $n - 2$ triangles.*

Proof. We induct on the number of vertices. For $n = 3$, P is a triangle, and so we are done. Now suppose a triangulation exists for all polygons with k vertices for $3 \leq k < n$, and let P be a polygon with n vertices. Then P has a diagonal, with separates P into an i-gon and a j-gon, with $i + j = n + 2$. By induction, these admit strong triangulations into $i - 2$ and $j - 2$ triangles respectively. This gives a strong triangulation of P into $i + j - 4 = n - 2$ triangles. □

This guarantees the existences of a triangulation, but is an inefficient way to construct one. A more constructive approach arises via ear clipping.

Definition 3.11. *An ear of a polygon P is a set of 3 consecutive vertices u, v, w such that $u - w$ is a diagonal of P. We say two ears overlap if the triangles they define intersect in their interiors.*

Lemma 3.12. *Every polygon with $n \geq 4$ vertices has two non-overlapping ears.*

Proof. P has a diagonal and the diagonal splits P into two polygons P_1, P_2¿ If $n = 4$, both are triangles and we are done. If $n > 4$, then one of the polygons has at least 4 vertices, and we can apply induction to conclude such ears exist. □

We can build a strong triangulation by finding an ear of P, "clipping" it off, and repeating the process with our new, smaller polygon. This is much faster than just taking random diagonals. Algorithms for finding triangulations based around diagonals have $O(n^4)$ runtimes, which ear-clipping algorithms have $O(n^2)$ runtimes. Neither are mathematically optimal, with most commonly used algorithms having a runtime of $O(n \log n)$. There is even an algorithm due to Chazelle that has a linear runtime, though it is too complicated to be implemented in a practical way.

Exercise 3.13. *In general, triangulations are not unique. Start with different ears/diagonals, and you will end up with different triangles. Show there exists a polygon P with a unique strong triangulation into $n - 2$ triangles for every $n \geq 3$*

A neat application of triangulations is the Art Gallery Problem: in a gallery, what is the minimum number of stationary rotating guards needed to watch the whole gallery? While we cannot answer this in general, we can give a good upper bound, at least for galleries with no central pillars.

To do so, we reformulate the problem geometrically. We say a set of points S in a polygon $\overline{P}$ is a guard set for S if, for every $p \in I(P)$, there is an $s_p \in S$ such that the line $s_p - p$ is contained in $\overline{P}$. The Art Gallery Problem asks for the minimum size of S. We will give an upper bound of $\lfloor \frac{n}{3} \rfloor$ vertices.

To prove this, we need a short combinatorial lemma about colours and colourings of triangulations.

Lemma 3.14. *Given a polygon P with strong triangulation $\mathcal{T}$, we can always colour the vertices of the triangles with 3 colours such that no two vertices of the same colour are endpoints of an edge.*

Proof. We induct on n. If $n = 3$, finding such a colouring is trivial. If $n > 3$, then we can find a diagonal $x - y$ of the strong triangulation cutting P into two polygons P_1, P_2 with strictly fewer vertices and induced strong triangulations. By induction, each of these admits such a colouring. In particular, x and y will have different colours in P_1, and x and y will have different colours in P_2. By renaming colours in P_2, we can assume that x has the same colour in both P_1 and P_2, and y has the same colour in P_1 and P_2. Thus we can merge our colourings to obtain a colouring of $\mathcal{T}$ as needed. □

Remark 3.15. *This only holds for strong triangulations. In general, you need 4 colours - though it is hard to show that you* only *need 4.*

Theorem 3.16. *For a polygon P with n vertices, there exists a guard set of $\lfloor \frac{n}{3} \rfloor$ vertices.*

Proof. Let $\mathcal{T}$ be a strong triangulation of P and colour the vertices with red, blue, and green as in Lemma 3.14. Of these three colours, one can appear on at most $\lfloor \frac{n}{3} \rfloor$ vertices. Suppose it is red. Then, any point of P is in a triangle with a red vertex, and is hence guarded by this red vertex. □

Example 3.17. *The accompanying example was not in a well guarded gallery, and will need to be recovered from those holding it for ransom.*

3.1 Lattice polygons

A particularly nice class of polygons are those whose vertices have integer coordinates (for some affine coordinate system). In particular, computations of perimeters reduce to dealing with right angled triangles, and computations of area reduce to counting.

Definition 3.18. *A plane lattice is a discrete subset $\Lambda \subset \mathbb{R}^2$ defined by*

$$\Lambda = \{mv_1 + nv_2 \mid m, n \in \mathbb{Z}^2\}$$

for $v_1, v_2 \in \mathbb{R}^2$ a pair of linearly independent vectors.

The standard lattice in $\mathbb{R}^2$ is $\mathbb{Z}^2 \subset \mathbb{R}^2$. For all that follows, we will fix $\Lambda = \mathbb{Z}^2$, though all essentially all the arguments will work for an arbitrary lattice. The resulting formulae may be a bit messier though.

Definition 3.19. *We say a polygon P is a lattice polygon if $V(P) \subset \Lambda$. Given a lattice polygon P, we divde the elements of Λ into interior, boundary, and exterior points*

Interior: *Elements of Λ contained within $I(P)$,*

Boundary: *Elements of Λ contained within $V(P) \cup E(P)$,*

Exterior: *All other elements of Λ.*

We denote by B_P and I_P the number of boundary and interior points, respectively.

Our next goal will be to show that the area bounded by P is a function of B_P and I_P, reducing area computation to counting point. We start by cutting P into nice triangles.

Definition 3.20. *A lattice triangle is called elementary if it has 3 boundary vertices and no interior vertices. A triangulation of a lattice polygon P is called elementary if it is made up of elementary triangles.*

Example 3.21. *Here are some examples of elementary and non-elementary triangles.*

Use your imagination for now

Lemma 3.22. *Every elementary triangle has equal area, given by $\frac{1}{2}$ for $\Lambda = \mathbb{Z}^2$.*

Proof. Let T be an elementary triangle. We can assume one of the vertices is $(0,0)$. Let the other two be

$$p = av_1 + bv_2, \quad q = cv_1 + dv_2.$$

The area of this triangle is given by half the norm of the cross product:

$$A = \frac{1}{2}\|p \times q\| = \frac{1}{2}|ad - bc|\|v_1 \times v_2\|$$

Hence the claim would follow from $ad - bc = \pm 1$. Equivalently, it would be enough to show that the linear map

$$v_1 \mapsto p = av_1 + bv_2, \quad v_2 \mapsto q = cv_1 + dv_2$$

is invertible, with inverse given by a matrix with integer entries.

To see this, complete the triangle to a parallelogram P with vertices $0, p, q, p+q$. As T is elementary, then P contains no interior vertices, and exactly 4 boundary vertices. To see this, note that if P had a "problem" vertex $mv_1 + nv_2$, then $p + q - mv_1 - nv_2$ would be a "problem" vertex for T. The parallelogram P clearly tiles the plane, and none of these translates can contain a problem vertex. Hence, every element of Λ must be a vertex of a copy of P in the tiling, and vice-versa:

$$\Lambda = \{Ap + Bq \mid A, B \in \mathbb{Z}\}$$

In particular, there is a linear map whose matrix has integer entries such that

$$p \mapsto v_1, \quad q \mapsto v_2.$$

Thus, the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is invertible with inverse with integer entries, and so has determinant $ad - bc = \pm 1$ as needed. □

Remark 3.23. *To prove the statement about "problem" vertices more rigorously, we note that $x \in P$ if and only if*

$$x = \lambda p + \mu q, \quad \textit{where} \quad 0 \leq \lambda, \mu \leq 1$$

while $x \in T$ if and only if

$$x = \lambda p + \mu q, \quad \textit{where} \quad 0 \leq \lambda, \mu \leq 1, \ \lambda + \mu \leq 1.$$

The claim then follows from some algebra.

Lemma 3.24. *Every lattice polygon has an elementary triangulation.*

Proof. We know there exists a (strong) triangulation $\mathcal{T}$ of every polygon P. If every triangle of $\mathcal{T}$ is already elementary, we are done. Otherwise, there is a triangle with an interior point or an excess boundary point. If it is an interior point, we use this interior point to divide the triangle into three smaller triangles. If it is an excess boundary point, we use it to divide the triangle into two smaller triangles. This gives a new triangulation with fewer "problem" vertices. We can repeat the process, splitting triangles and reducing the number of problem vertices each time. As there were only finitely many problem vertices to begin with, we must eventually obtain a triangulation with no problem vertices, i.e. an elementary triangulation. ☐

Thus, if we can figure out how many elementary triangles are in an elementary triangulation of a polygon P, we can work out its easy very easily!

Lemma 3.25. *Let P be a lattice polygon, and $\mathcal{T}$ be an elementary triangulation containing T triangles. Then*

1. $|V(\mathcal{T})| = B_P + I_P$,
2. $2|E(\mathcal{T})| = 3T + B_P$
3. *There are $T = 2I_P + B_P - 2$ triangles in $\mathcal{T}$.*

Proof.

1. Every vertex of the triangulation is either interior to P or on the boundary.
2. Every triangle contains three edges, and every pair of adjacent boundary vertices defines a sign edge of a triangle. Thus, the total number of edges, counted with multiplicity, is $3T+B_P)$. Every edge of a triangle is contained in two triangles, or is contained in one triangle and joins two boundary vertices of P, and so appear twice in this count. Thus
$$3T + B_P = 2|E(\mathcal{T})|.$$
3. We will induct on T. If $T = 1$, then P is an elementary triangle and the claim follows. Now suppose the relationship is true for triangulations of polygons into $T - 1$ triangles, and suppose P has a triangulation into T triangles. Pick a triangle with an edge on the boundary of P.
   - If this triangle only has one edge on the boundary of P, we delete it to obtain a triangulation of a polygon with $B_P + 1$ boundary vertices and $I_P - 1$ interior vertices into $T - 1$ elementary triangles. By induction
   $$T - 1 = 2(I_P - 1) + B_P + 1 - 2, \quad \Rightarrow \quad T = 2I_P + B_P - 2.$$
   - If this triangle has exactly two edges on the boundary, then we delete both and their shared vertex to obtain a triangulation of a polygon with $B_P - 1$ boundary vertices and I_P interior vertices into $T - 1$ triangles. By induction
   $$T - 1 = 2I_P + (B_P - 1) - 2 \quad \Rightarrow \quad T = 2I_P + B_P - 2.$$
   - If this triangle has three edges on the boundary, then P was an elementary triangle and we have already considered this.

□

This leads use to Pick's theorem for the area of a lattice polygon

Theorem 3.26. *Given a lattice polygon P on $\Lambda = \mathbb{Z}^2$, we have*

$$\text{Area}(P) = I_P + \frac{B_P}{2} - 1$$

Proof. We have an elementary triangulation into $2I_P + B_P - 2$ elementary triangles, each of which as an area of $\frac{1}{2}$, and so the claim follows. □

3.2 Polygons with holes and the Euler characteristic

Obviously, we can handle polygons with holes in the middle by comparing the (full) exterior polygon with the interior polygon, but we can give a bit of a slicker approach via an invariant called the Euler characteristic. But first, what is a polygon with a hole in the middle.

Definition 3.27. *A polygon with holes P consists of a polygon P_{ext}, called the exterior boundary, and a number of non-intersecting interior polygons $P_1, \ldots, P_n$, called holes, such that no interior polygon is contained within the interior of another. We write $h(P) = n$ for the number of holes of P.*

Given a polygon with holes, we refer to the boundary of P_{ext} as the exterior boundary and the boundaries of $P_1, \ldots, P_n$ are the interior boundaries. The interior of P is the set

$$I(P) = I(P_{ext}) \setminus (I(P_1) \cup \cdots \cup I(P_n)) .$$

We can naturally extend many of our results about triangulations and to polygons with holes, but we have to slightly modify Pick's formula to count the number of holes, via the Euler characteristic.

Definition 3.28. *Given a polygon with holes P and a triangulation $\mathcal{T}$, we define the Euler characteristic with respect to $\mathcal{T}$ of P by*

$$\chi_{\mathcal{T}}(P) = |V(\mathcal{T})| - |E(\mathcal{T})| + |F(\mathcal{T}|$$

where we now write $F(\mathcal{T})$ for the set of triangles of $\mathcal{T}$. It will now be convenient to refer to these triangles as faces.

Fact 3.29. *For any polygon with holes, the Euler characteristic depends only on P, not on the triangulation. Specifically*

$$\chi_{\mathcal{T}}(P) = |V(P)| - |E(P)| + (1 - h(P))$$

for all triangulations $\mathcal{T}$. As such, we often just write $\chi(P)$ for this common value.

Remark 3.30. *Proving this can be a worthwhile exercise. An argument could go as follows:*

1. *Extend the Euler characteristic to subdivisions of P into arbitrary polygons and consider what happens when you merge two faces of a subdivision.*

2. *Show that, for any two triangulations of P, there is a subdivision $\mathcal{S}$ of P such that every face of $\mathcal{S}$ is contained in a triangle of each triangulation. Hence conclude that $\chi_{\mathcal{T}}(P)$ is independent of $\mathcal{T}$.*

3. *Compute $\chi_{\mathcal{T}}(P)$ for a strong triangulation, but considering how the number of edges and faces change as you delete diagonals.*

With this, we can give a generalisation of Pick's theorem. This could also be shown just by subtracting the appropriate areas, but also follows immediately from taking an elementary triangulation.

Theorem 3.31. *Let P be a lattice polygon with holes. Then*

$$\text{Area}(P) = I_P + \frac{B_P}{2} - \chi(P).$$

Proof. Take an elementary triangulation $\mathcal{T}$ of P, with $|F(\mathcal{T})|$ triangles. Then

$$\begin{aligned} |F(\mathcal{T})| &= \chi_{\mathcal{T}}(P) - |V(\mathcal{T})| + |E(\mathcal{T})| \\ &= \chi(P) - B_P - I_P + \frac{3|F(\mathcal{T}) + B_P}{2} \end{aligned}$$

Solving this for $|F(\mathcal{T})|$, we find

$$\text{Area}(P) = \frac{1}{2}|F(\mathcal{T})| = I_P + \frac{B_P}{2} - \chi(P).$$

□

4 Elliptic Geometry

5 Hyperbolic Geometry

6 Projective geometry

7 Summary of main results

8 Appendix A: A review of linear algebra over the reals

Definition 8.1. *A (real) vector space is a set V with two operations:*

Addition *a map* $+ : V \times V \to V$,

Scalar Multiplication *a map* $\cdot \mathbb{R} \times V \to V$

and a distinguished element $0 \in V$, *satisfying the following*

1. $(x + y) + z = x + (y + z)$,
2. $x + y = y + x$,
3. $x + 0 = x$,
4. $\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y$,
5. $(\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x$,
6. $\lambda \cdot (\mu \cdot x) = (\lambda\mu) \cdot x$
7. $0 \cdot x = 0$.

We will usually omit the dot in scalar multiplication.

Example 8.2. *The main example is* $\mathbb{R}^n$, *the set of* n*-tuples of real numbers*

$$x = (x_1, \dots, x_n)$$

with componentwise addition and scalar multiplication:

$$\begin{aligned}(x_1, \dots, x_n) + (y_1, \dots, y_n) &= (x_1 + y_1, \dots, x_n + y_n),\\ \lambda(x_1, \dots, x_n) &= (\lambda x_1, \dots, \lambda x_n),\\ 0 &= (0, \dots, 0).\end{aligned}$$

Definition 8.3. *A basis for* V *is a minimal set of vectors* $\{v_1, \dots, v_N\} \subset V$ *such that every* $x \in V$ *can be written in the form*

$$x = \lambda_1 v_1 + \cdots \lambda_N v_N.$$

We call N *the dimension of* V

Example 8.4. *The vector space* $\mathbb{R}^n$ *has a basis, called the standard basis, consisting of* $e_1, \dots, e_n$, *where* e_i *has a* 1 *in the* i^{th} *entry and zeroes elsewhere. As such,* $\mathbb{R}^n$ *is an* n*-dimensional vector space.*

Definition 8.5. *A map* $T : V \to W$ *between vector spaces is called a linear map (or a homomorphism) if*

$$T(x + y) = T(x) + T(y) T(\lambda x) = \lambda T(x)$$

A linear map T *with an inverse map* $T^{-1} : W \to V$ *is called an isomorphism. If* $T : V \to W$ *is an isomorphism, we say* V *and* W *are isomorphic.*

Exercise 8.6. *Show that every vector space of dimension n is isomorphic to $\mathbb{R}^n$.*

Fact 8.7. *Every linear map $T : \mathbb{R}^m \to \mathbb{R}^n$ can be represented by an $(n \times m)$-matrix A, such that*

$$T(x) = Ax$$

is given by matrix multiplication. As such, we will often omit the brackets when discussing linear maps.

Definition 8.8. *An inner product on a vector space V is a map $I : V \times V \to \mathbb{R}$ such that*

1. $I(\lambda x + \mu y, z) = \lambda I(x, z) + \mu I(y, z)$ *for all* $x, y, z \in V$, $\lambda, \mu \in \mathbb{R}$,
2. $I(z, \lambda x + \mu y) = \lambda I(z, x) + \mu I(z, y)$ *for all* $x, y, z \in V$, $\lambda, \mu \in \mathbb{R}$,
3. $I(x, y) = I(y, x)$ *for all* $x, y \in V$,
4. $I(x, x) > 0$ *for all* $x \in V$, $x \neq 0$.

These properties are referred to as bilinearity, symmetry, and positive-definiteness respectively.

Remark 8.9. *Note that properties (1) and (3) imply (2), leading to a bit of redundancy for sake of clarity.*

Example 8.10. *The standard inner product on $\mathbb{R}^n$ (sometimes called the dot product) is given by*

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle = x_1 y_1 + \cdots + x_n y_n.$$

Fact 8.11. *Given a matrix A, matrix multiplication interacts with the standard inner product as follows*

$$\langle x, Ay \rangle = \langle A^T x, y \rangle$$

where A^T is the transpose.

Definition 8.12. *Given a vector space V with an inner product I, we say $v, w \in V$ are orthogonal if $I(v, w) = 0$. A set $\{v_1, \ldots, v_n\}$ is orthonormal if*

$$I(v_i, v_j) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

Example 8.13. *The standard basis of $\mathbb{R}^n$ is orthonormal with respect to the standard inner product.*

Definition 8.14. *A norm on V is a function $\| \cdot \| : V \to \mathbb{R}_{\geq 0}$ such that*

1. $\|\lambda v\| = |\lambda| \|v\|$ *for all* $\lambda \in \mathbb{R}$, $v \in V$,
2. $\|v\| = 0$ *if and only if* $v = 0$,

3. $\|v+w\| \leq \|v\| + \|w\|$

Example 8.15. *The Euclidean norm on* $\mathbb{R}^n$ *is given by*

$$\|(x_1, \ldots, x_n)\| = \sqrt{x_1^2 + \cdots + x_n^2}$$

which give the (Euclidean) length of the vector x*.*

Remark 8.16. *This inequality is usually called the triangle inequality. By taking* $v = a - b$ *and* $w = b - c$*, we can rewrite it as*

$$\|a - c\| \leq \|a - b\| + \|b - c\|.$$

If we think of a, b, c *as describing the vertices of a triangle* ABC*, this says*

$$|AC| \leq |AB| + |BC|$$

which is precisely the condition needed to 3 line segments to be able to form a triangle!

Exercise 8.17. *Let* V *be a vector space with inner product* I*. Show that*

$$\|x\| := \sqrt{I(x, x)}$$

defines a norm on V*. Show that for such norms, we have that*

$$\|v + w\| = \|v\| + \|w\|$$

if and only if $v = 0$ *or* $w = \lambda v$ *for some real* λ*.*

Example 8.18. *The Euclidean norm on* $\mathbb{R}^n$ *comes from the standard inner product.*

Fact 8.19. *A norm* $\|\cdot\|$ *coming from an inner product satisfies the parallelogram law*

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

Theorem 8.20. *Let* V *be a vector space with norm* $\|\cdot\|$*. If the norm satisfies the parallelogram law, then there exists a unique inner product* I *on* V *inducing* $\|\cdot\|$*.*

Proof. We claim that

$$I(x, y) := \frac{1}{4}\left(\|x + y\|^2 - \|x - y\|^2\right)$$

defines an inner product on V. Symmetry and positive-definiteness are quick to check, while bilinearity can be derived from the parallelogram law. I induces the norm

$$I(x, x) = \frac{1}{4}\|2x\|^2 = \|x\|^2.$$

To see that I is unique, suppose

$$\|x\| = \sqrt{J(x,x)}$$

for some inner product J. Then, using bilinearity of J, we find

$$I(x,y) = \frac{1}{4}\left(J(x+y,x+y) - J(x-y,x-y)\right) = J(x,y).$$

Thus, the inner product is unique. □

9 Appendix B: A review of groups

Definition 9.1. *A group is a set G, with a multiplication map $G \times G \to G$ and a distinguished identity element $e \in G$ such that*

- *$a(bc) = (ab)c$ for all $a, b, c \in G$,*
- *$ae = ea = a$ for all $a \in G$,*
- *For all $a \in G$, there exists $a^{-1} \in G$ such that*

  $$aa^{-1} = a^{-1}a = e$$

  called the inverse.

Example 9.2. *The set of integers $(\mathbb{Z}, +)$ with addition is a group. The identity element is 0, and the inverse is the negative.*

Example 9.3. *The set $\mathrm{GL}_n(\mathbb{R})$ of invertible $(n \times n)$-matrices with matrix multiplication is a group. The identity is the identity matrix*

$$I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

and inverse given my the matrix inverse. For $n = 2$, the inverse is given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

This is called the general linear group.

Example 9.4. *The special linear group is a subgroup of $\mathrm{GL}_n(\mathbb{R})$ given by*

$$\mathrm{SL}_n(\mathbb{R}) = \{M \in \mathrm{GL}_n(\mathbb{R}) \mid \det M = 1\}$$

Example 9.5. *The orthogonal group is a subgroup of* $\mathrm{GL}_n(\mathbb{R})$ *given by*

$$\mathrm{O}_n(\mathbb{R}) = \{M \in \mathrm{GL}_n(\mathbb{R}) \mid MM^T = M^T M = I\},$$

the set of matrices whose transpose is the inverse. A simple exercise is to show that the determinant of every matrix in this group is ± 1*. Orthogonal matrices represent a combination of reflections and rotations of* $\mathbb{R}^n$

Example 9.6. *The special orthogonal group is the intersection of* $\mathrm{SL}_n(\mathbb{R})$ *with* $\mathrm{O}_n(\mathbb{R})$*, given by*

$$\begin{aligned} \mathrm{SO}_n(\mathbb{R}) &= \{M \in \mathrm{SL}_n(\mathbb{R}) \mid MM^T = M^T M = I\} \\ &= \{M \in \mathrm{O}_n(\mathbb{R}) \mid \det M = 1\}. \end{aligned}$$

Elements of this group represent rotations of $\mathbb{R}^n$*.*

Exercise 9.7. *Show that*

$$\mathrm{SO}_2(\mathbb{R}) = \left\{ \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \mid \theta \in [0, 2\pi) \right\}.$$

Definition 9.8. *A map between groups* $f : G \to H$ *is called a homomorphism if*

$$f(ab) = f(a)f(b)$$

for all $a, b \in G$*. An invertible is called an isomorphism. If* $f : G \to H$ *is an isomorphism, we say* G *and* H *are isomorphic.*