HILBERT'S AXIOM SYSTEM FOR PLANE GEOMETRY
A SHORT INTRODUCTION

BJØRN JAHREN

Euclid's "Elements" introduced the axiomatic method in geometry, and for more than 2000 years this was the main textbook for students of geometry. But the 19th century brought about a revolution both in the understanding of geometry and of logic and axiomatic method, and it became more and more clear that Euclid's system was incomplete and could not stand up to the modern standards of rigor. The most famous attempt to rectify this was by the great German mathematician *David Hilbert*, who published a new system of axioms in his book "Grundlagen der Geometrie" in 1898. Here we will give a short presentation of Hilbert's axioms with some examples and comments, but with no proofs. For more details, we refer to the rich literature in this field — e. g. the books "Euclidean and non-Euclidean geometries" by M. J. Greenberg and "Geometry: Euclid and beyond" by R. Hartshorne.

Hilbert also treats geometry in 3-space, but we will only consider the 2-dimensional case. We begin by agreeing that the basic ingredients in our study are *points* and *lines* in a *plane*. At the outset the plane is just a set S where the elements P are called points. The lines are, or can at least be naturally identified with certain subsets l of S, and the fundamental relation is the *incidence relation* $P \in l$, which may or may not be satisfied by a point P and a line l. But we also introduce two additional relations: *betweenness*, enabling us to talk about points lying between two given points, and *congruence*, which is needed when we want to compare configurations in different parts of the plane. Hilbert formulated three sets of axioms for these relations: *incidence axioms, betweenness axioms* and *congruence axioms.* In addition to these we also need an *axiom of continuity* to make sure that lines and circles have "enough" points to intersect as they should, and of course the *axiom of parallels.* As we introduce Hilbert's axioms, we will gradually put more and more restrictions on these ingredients, and in the end they will essentially determine Euclidean plane geometry uniquely. The axioms are also *independent*, in the sense that for each axiom $\mathcal{A}$ there is a model satisfying all the rest of the axioms, but not $\mathcal{A}$.

Note that although *circles* also are important objects of study in classical plane geometry, we do not have to postulate them, since, as we shall see, they can be defined in terms of the other notions.

Before we start, maybe a short remark about language is in order: An axiom system is a formal matter, but the following discussion will not be very formalistic. After all, the goal is to give a firm foundation for matters that we all have a clear picture of in our minds, and as soon as we have introduced the various formal notions, we will feel free to discuss them in more common language. For example, although the relation $P \in l$ should, strictly speaking, be read: "P and l are incident", we shall use "l contains P", "P lies on l" or any obviously equivalent such expression.

We are now ready for the first group of axioms, the *incidence axioms:*

I1: *For every pair of distinct points A and B there is a unique line l containing A and B.*

I2: *Every line contains at least two points.*

I3: *There are at least three points that do not lie on the same line.*

We let $\overline{AB}$ denote the unique line containing A and B.

These three axioms already give rise to much interesting geometry, so-called "incidence geometry". Given three points A, B, C, for example, any two of them span a unique line, and it makes sense to talk about the *triangle ABC*. Similarly we can study more complicated configurations. The Cartesian model $\mathbb{R}^2$ of the Euclidean plane, where the lines are the sets of solutions of nontrivial linear equations $ax + by = c$, is an obvious example, as are the subsets obtained if we restrict a, b, c, x, y to be rational numbers ($\mathbb{Q}^2$), the integers ($\mathbb{Z}^2$), or in fact any fixed subring of $\mathbb{R}$. However, *spherical geometry*, where S is a sphere and the *lines* are great circles, is not an example, since any pair of antipodal points lies on infinitely many great circles — hence the uniqueness in I1 does not hold. This can be corrected by identifying every pair of antipodal points on the sphere. Then we obtain an incidence geometry called *the (real) projective plane* $\mathbb{P}^2$. One way to think about the points of $\mathbb{P}^2$ is as lines through the origin in $\mathbb{R}^3$. If the sphere has center at the origin, such a line determines and is determined by the antipodal pair of points of intersection between the line and the sphere. A "line" in $\mathbb{P}^2$ can then be thought of as a plane throuh the origin in $\mathbb{R}^3$, since such a plane intersects the sphere precisely in a great circle. Notice that in this interpretation the incidence relation $P \in l$ corresponds to the relation "the line l is contained in the plane P".

There are also *finite* incidence geometries — the smallest has exactly three points where the lines are the three subsets of two elements.

The next group of axioms deals with the relation "B lies between A and C". In Euclidean geometry this is meaningful for three points A, B and C lying on the same straight line. The finite geometries show that it is not possible to make sense of such a relation on every incidence geometry, so this is a new piece of structure, and we have to declare the properties we need. We will use the notation $A * B * C$ for "B lies between A and C".

Hilbert's *axioms of betweenness* are then:

B1: *If $A * B * C$, then A, B and C are distinct points on a line, and $C * B * A$ also holds.*

B2: *Given two distinct points A and B, there exists a point C such that $A * B * C$.*

B3: *If A, B and C are distinct points on a line, then one and only one of the relations $A * B * C$, $B * C * A$ and $C * A * B$ is satisfied.*

B4: *Let A, B and C be points not on the same line and let l be a line which contains none of them. If $D \in l$ and $A * D * B$, there exists an E on l such that $B * E * C$, or an F on l such that $A * F * C$, but not both.*

If we think of A, B and C as the vertices of a triangle, another formulation of B4 is this: If a line l goes through a side of a triangle but none of its vertices, then it also goes through exactly one of the other sides. This formulation is also called *Pasch's axiom.* Note that this is not true in $\mathbb{R}^n$, $n \geq 3$. Hence I3 and B4 together define the geometry as '2–dimensional'.

In the standard Euclidean plane (and in other examples we shall study later) we can use the concept of *distance* to define betweenness. Namely, we can then define $A * B * C$ to mean that A, B are C are distinct and $d(A, C) = d(A, B) + d(B, C)$, where $d(X, Y)$ is the distance between X and Y. (Check that B1-4 then hold!) This way $\mathbb{Q}^2$ also becomes an example, but not $\mathbb{Z}^2$, since B4 is not satisfied. (Exercise 3.)

Observe also that every open, convex subset K of $\mathbb{R}^2$ (e. g. the interior of a circular disk) satisfies all the axioms so far, if we let the "lines" be the nonempty intersections between lines in $\mathbb{R}^2$ and K, and betweenness is defined as in $\mathbb{R}^2$. (This example will be important later.) The projective plane, however, can not be given such a relation. The reason is that in the spherical model for $\mathbb{P}^2$, the "lines" are great circles where antipodal points have been identified, and these identification spaces can again naturally be identified with circles. But if we have three distinct points on a circle, each of them is equally much "between" the others. Therefore B3 can not be satisfied.

The betweenness relation can be used to define the *segment* AB as the point set consisting of A, B and all the points between A and B:

$$AB = \{A, B\} \cup \{C | A * C * B\}.$$

Similarly we can define the *ray* $\overrightarrow{AB}$ as the set

$$\overrightarrow{AB} = AB \cup \{C | A * B * C\}.$$

If A, B and C are three point not on a line, we can then define the *angle* $\angle BAC$ as the pair consisting of the two rays $\overrightarrow{AB}$ and $\overrightarrow{AC}$.

$$\angle BAC = \{\overrightarrow{AB}, \overrightarrow{AC}\}.$$

Note also that $\overline{AB} = \overrightarrow{AB} \cup \overrightarrow{BA}$.

Betweenness also provides us with a way to distinguish between the two *sides* of a line l. We say that two points A and B are *on the same side of* l if $AB \cap l = \emptyset$. It is not difficult to show, using the axioms, that this is an equivalence relation on the complement of l, and that there are exactly two equivalence classes: the two sides of l. (Exercise 4.) Similarly we say that a point D is *inside* the angle $\angle BAC$ if B and D are on the same side of $\overline{AC}$, and C and D are on the same side of $\overline{AB}$. This way we can distinguish between points inside and outside a triangle. We also say that the angles $\angle BAC$ and $\angle BAD$ are on the same (resp. opposite) side of the ray $\overrightarrow{AB}$ if C and D are on the same (resp. opposite) side of the line $\overline{AB}$.

The same idea can, of course, also be applied to distinguish between the points on a line on either side of a given point. Using this, one can define a linear ordering of all the points on a line. Therefore the axioms of betweenness are sometimes called "axioms of order".

We have now established some of the basic concepts of geometry, but we are missing an important ingredient: we cannot yet compare two different configurations of points and lines. To achieve this, we need to introduce the concept of *congruence*. Intuitively, we may think of two configurations as congruent if there is some kind of "rigid motion" which moves one onto the other. In the Euclidean plane $\mathbb{R}^2$ this can be defined in terms of measurements of angles and distances, such that two configurations are congruent if all their ingredients are "of the same size". These notions have, of course, no meaning on the basis of just the incidence- and betweenness axioms. Hence congruence has to be yet another piece of structure — a relation whose properties must be governed by additional axioms.

There are two basic notions of congruence — congruence of segments and congruence of angles. Congruence of more general configurations can then be defined as a one-one correspondence between the point sets involved such that all corresponding segments and angles are congruent. We use the notation $AB \cong CD$ for "the segment AB is congruent to the segment CD", and similarly for angles or more general configurations. Hilbert's axioms for congruence of segments are:

C1: *Given a segment AB and a ray r from C, there is a uniquely determined point D on r such that $CD \cong AB$.*

C2: *$\cong$ is an equivalence relation on the set of segments.*

C3: *If $A * B * C$ and $A' * B' * C'$ and both $AB \cong A'B'$ and $BC \cong B'C'$, then also $AC \cong A'C'$.*

If betweenness is defined using a distance function (as in the Euclidean plane) we can define $AB \cong CD$ as $d(A, B) = d(C, D)$. C2 and C3 are then automatically satisfied, and C1 becomes a stronger version of B2.

Even without a notion of distance we can use congruence to compare "sizes" of two segments: we say that AB is shorter than CD ($AB < CD$) if there exists a point E such that $C * E * D$ and $AB \cong CE$.

We can now also define what we mean by a *circle:* Given a point O and a segment AB, we define the circle with center O and radius (congruent to) AB as the point set $\{C \in S \,|\, OC \cong AB\}$. Note that this set is nonempty: C1 implies that any line through O intersects the circle in two points.

The axioms for congruence of angles are:

C4: *Given a ray $\overrightarrow{AB}$ and an angle $\angle B'A'C'$, there are angles $\angle BAE$ and $\angle BAF$ on opposite sides of $\overline{AB}$ such that $\angle BAE \cong \angle BAF \cong \angle B'A'C'$.*

C5: *$\cong$ is an equivalence relation on the set of angles.*

C6: *Given triangles ABC and $A'B'C'$. If $AB \cong A'B'$, $AC \cong A'C'$ and $\angle BAC \cong \angle B'A'C'$, then the two triangles are congruent — i. e. $BC \cong B'C'$, $\angle ABC \cong \angle A'B'C'$ and $\angle BCA \cong \angle B'C'A'$.*

C4 and C5 are the obvious analogues of C1 and C2, but note that C4 says that we can construct an arbitrary angle on *both* sides of a given ray. C6 says that a triangle is determined up to congruence by any angle and its adjacent sides. This statement is often referred to as the "SAS" (side–angle–side) congruence criterion.

In the Euclidean plane $\mathbb{R}^2$ we define congruence as equivalence under actions of the *Euclidean group* of transformations of $\mathbb{R}^2$. This is generated by rotations and translations, and can also be characterized as the set of transformation of $\mathbb{R}^2$ which preserve all distances. It is quite instructive to prove that the congruence axioms hold with this definition.

These three groups contain the most basic axioms, and they are sufficient to prove a large number of propositions in book I of the "Elements". However, when we begin to study circles and "constructions with ruler and compass", we need criteria saying that circles intersect (have common points) with other circles or lines when our intuition tells us that they should. The next axiom provides such a criterion.

First a couple of definitions:

Definition: Let Γ be a circle with center O and radius OA. We say that a point B is *inside* Γ if $OB < OA$ and *outside* if $OA < OB$.

We say that a line or another circle is *tangent to* Γ if they have exactly one point in common with Γ.

We can now formulate Hilbert's axiom E:

E: *Given two circles* Γ *and* Δ *such that* Δ *contains points both inside and outside* Γ*. Then* Γ *and* Δ *have common points. (They "intersect".)*

(It follows from the other axioms that they will then intersect in exactly *two* points.) This is an example of what we call a *continuity axiom.* The following variation is actually a consequence of axiom E:

E': *If a line* l *contains points both inside and outside the circle* Γ*, then* l *and* Γ *will intersect. (Again in exactly two points.)*

Hilbert gives the *Axiom of parallels* the following formulation — often called "Playfair's axiom" (after John Playfair i 1795, although it goes back to Proclus in the fifth century):

P: *(Playfair's axiom) Given a line* l *and a point* P *not on the line. Then there is at most one line* m *through* P *which does not intersect* l*.*

If the lines m and l do not intersect, we say that they are *parallel*, and we write $m \parallel l$. The *existence* of a line m through P parallel to l can be shown to follow from the other axioms, so the real content of the axiom is the uniqueness.

With these axioms we are able to prove all the results in Euclid's "Elements" I–IV, but they do not yet determine the Euclidean plane uniquely. The standard plane ($\mathbb{R}^2$ with the structure defined so far) is an example, and it is an instructive exercise to prove this in detail, but we obtain other examples by replacing the real numbers by another ordered field where every element has a square root! For uniqueness we need a stronger continuity axiom, as for instance *Dedekind's axiom:*

D: *If a line l is a disjoint union of two subset T_1 and T_2 such that all the points of T_1 are on the same side of T_2 and vice versa, then there is a unique point $A \in l$ such that if $B_1 \in T_1$ and $B_2 \in T_2$, then either $A = B_1$, $A = B_2$ or $B_1 * A * B_2$.*

This is a completeness axiom with roots in Dedekind's definition of the real numbers, and an important consequence is that the geometry on any line can be identified with the geometry on $\mathbb{R}$. One can show that it implies axiom E, and together with the groups of axioms I*, B*, C* and P it does determine Euclidean geometry completely.

Finally we mention that axiom D also implies another famous continuity axiom, the *Axiom of Archimedes:*

A: *Given two segments AB and CD, we can find points $C = C_0, \dots, C_n$ on $\overrightarrow{CD}$, such that $C_iC_{i+1} \cong AB$ for every $i < n$ and $CD < CC_n$.*

("Given a segment AB, then every other segment can be covered by a finite number of congruent copies of AB".)

Using this axiom we can introduce notions of distance and length such that AB has length one, say, and a geometry with the axioms I*, B*, C* P, E and A can be identified with a subset of the standard Euclidean plane.

Exercises.

1. Find all incidence geometries with four or five points.
2. Let V be a vector space of dimension at least 2 over a field F. Show that V satisfies I1-3, if we define *lines* to be sets of the form $\{A + tB | t \in F\}$, where $A, B \in V$, $B \neq 0$.
3. Show that $\mathbb{Q}^2$ satisfies axioms B1–4, but $\mathbb{Z}^2$ does not.
4. Prove that 'being on the same side of the line ℓ' is an equivalence relation on the complement of ℓ, with exactly two equivalence classes.
5. Show that $\mathbb{Q}^2$ does not satisfy C1. Try to determine conditions on an algebraic extension F of $\mathbb{Q}$ such that F^2 will satisfy C1.
6. Show that the center of a circle is uniquely determined.
7. Discuss which axioms are needed in order to bisect a given segment.
8. Which axioms are satisfied by $\overline{\mathbb{Q}}^2$, where $\overline{\mathbb{Q}}$ is the algebraic closure of $\mathbb{Q}$?
9. Show that the axiom of Archimedes can be used to define a length function on segments.
10. Suppose given a geometry with incidence, betweenness and congruence, and let r be a ray with vertex $\mathcal{O}$. r determines a unique line ℓ containing it.

    Show that we can give ℓ the structure of an ordered abelian group, with $\mathcal{O}$ as neutral element and such that $\mathcal{P} \geqslant \mathcal{O}$ if and only if $\mathcal{P} \in r$.

    Show that two different rays give rise to isomorphic ordered groups.