This is the first lecture of *Introduction to mathematics* which I teach at Department of Logic, Charles University. However, it can also be used as a stand-alone text. Note that this is still a draft and as such may contain errors.

Radek Honzik, December 2011

Truth vs. deducibility: Euclid's axiomatization

We start by a paraphrase of Euler's citation: in lecturing, one can either achieve correctness or clarity, hardly ever both. This section is supposed to provide a motivation and possibly simple formulation of the issues we will deal with later on in more details (such as "proof", "consistency", "theory" etc.). It should show that in mathematics one needs to be careful of one's assumptions before he or she draws conclusions.

Euclid's Elements

Euclid of Alexandria (ca. 300 B.C.E.), his work *Elements* (at least attributed to him; more probably he has collected and summarized earlier knowledge).

First treatise which tries to be highly deductive in nature: from a list of obvious propositions (*postulates*) it derives (by means of intuitively acceptable rules) *theorems* or *propositions*.

Description of Elements:

- Four books on the study of plane figures: triangles, quadrilaterals (čtyřúhelník), and circles. The famous theorem of Pythagoras is the 47th proposition of the first book.
- Two books on the theory of ratio and proportion and of the theory of similar figures.
- Next three books are about whole numbers (for instance he proves that there are infinitely many primes).[1]
- The 10th book deals with lengths of some irrational value (which can be described by a square root).
- Final three books are about three-dimension geometry (for instance the proof that there are exactly five regular solids).

Most arguments in Elements (with some omissions in the number-theoretic part) use the *axiomatic method*. From basic propositions, and definitions, Euclid derives by purely logical means other theorems. For instance, the circularity (vicious circle) argument must be avoided.

Book I contains five "postulates" which are more narrowly geometrical. Note that "line" was for Greeks more like a "line segment" (úsečka) then "line" (přímka). This distinction is however quite subtle and we will not be very careful about this.[2]

(0.1) **Euclid's postulates**

E1. For all two points there is a line segment passing through the two points.
E2. Every line segment can be always extended.
E3. For every point and every length r, there exists a circle with its center at the point and with radius r.
E4. All right angles are equal.
E5. If a straight line segment falling on two straight lines makes the interior angles on the same side less than two right angles, the two straight lines, if extended indefinitely, meet on that side on which are the angles less than two right angles.

Remark 0.1 We will refer as *core* to the postulates E1–E4 above plus all other definitions and rules introduced by Euclid in his book before E5 is introduced. This is patently incorrect

[1] **Just argue that if p were the greatest prime, then consider the number $p!+1$. The number $p!$ is clearly divided by every number $1 < n \leq p$. It follows that $p!+1$ cannot be divided by any such n. By the Fundamental theorem of arithmetics (also included in Elements) each number, and also $p!+1$, has a (unique) prime factorization, and hence the factorization of $p!+1$ must contain a prime number $> p$.

[2] The cause for the preference of line segments over lines is the proverbial fear of Greeks of the actual infinity.

in that we do not say what these are, but we again refer to Euler's citation above: the exact formulation of all assumption would be very long. We will need to leave some things to the intuition.

It follows that we can define two lines to be *parallel* if they do not meet. The "parallel postulate" E5 is much more complicated than the previous four, and from the beginning commentators suggested that it should be deducible from the previous postulates. Maybe even Euclid felt this because its introduction is delayed until proposition 29 of Book I.

A Scottish editor, Robert Simson, provided in the 19th century the following equivalent, now more frequent, reformulation of E5:

E5*. Given any line p in plane and any point P in that plane that does not lie on p, there is exactly one line q that passes through the point P and does not meet the line p.

However, as stated E5* is *not* equivalent to E5 over the core.[3] The point is that E5 is stated as in implication, not equivalence. Assume we strengthen E5 to E5+ as follows:

E5+. 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 extended indefinitely, meet on that side on which are the angles less than two right angles. If on the other hand, the interior angles are two right angles on one side (and so also on the other side), then the lines will never meet.

We can now show the following to get some experience with geometrical arguments before we go on:

Claim 0.2 *(**)E5* and E5+ are equivalent over the core.*

Proof. (**)First assume that E5+ holds and we want to prove E5*. The second part of E5+ implies that there is at least one parallel to the line p through a given point Q: use two perpendiculars (kolmice), the first one passing through the point Q crossing the line p at the right angle, and the second parallel passing through Q, perpendicular to the first perpendicular. The first part of E5+ implies that there is at most one parallel: if there were two, one of them would need to have on one side an angle less than two right angles, thus by E5+ the lines would meet on this side, thus contradicting the initial assumption that the lines are parallel.

Now assume that E5* holds and we want to prove E5+. From the assumption that there is at least one parallel we show the second part of E5+: if the lines with the interior angle equal to two right angles meet on one side, then by equality of right angles they would meet on the other side as well. It follows that every line which goes through the point meets one of the lines (because on one side it would form an angle of size less the right angle – using shodné úhly, thus meeting the line). This is a contradiction, and so the second part of E5+ is proved. Assume now that there is at most one parallel. From this we want to conclude the first part of E5+: if there were lines which have the sum of interior angles less than two right angles, and which do not meet, than we can draw a line above the line which has the angle less than the right angle and which would still form an angle less than the right angle, and which would not meet the second line. Contradiction. □

How is it that Euclid states E5 without the converse as in E5+? The reason is that Euclid thought that he has proved the converse, in fact it follows from the proposition 17 in Book I, which states that "in any triangle the sum of any two angles is less than two right angles."

However, we will later see that Euclid forgot one assumption which as it turns out is necessary to show this converse proposition. He used the following assumption, but he thought that is is so evident that it needn't be mentioned:

[3] This is a matter of some confusion in literature. The reason is that it matters which exact formulation of Euclid's postulates we use.

(0.2) **The forgotten Euclid's postulate**

E0. Every two distinct lines meet in at most one point.

We now show that E0 plus core suffices to show the following weakening of E5*:

Claim 0.3 *(**) Given any line p in the plane and any point Q in the plane that does not lie on p, there is at least one line q that passes through the point Q and does not meet the line p.*

Proof. (**) Construct two perpendiculars, obtaining a line q passing through Q such that the interior angles formed by p, q and the first perpendicular are equal to two right angles. If p and q meet on one side, then they would meet (by equality of right angles) on the other side as well, thus contradicting E0. □

We later show that E0 cannot be proved from the core.

Remark 0.4 Perhaps this is a place to stop and clear a possible misunderstanding. How can two lines meet in more than one point? is it not obviously false? Two views are mixed up here: (1) A statement is true (in our space as we see it); (2) A statement is deducible from the postulates. With the benefit of new and astonishing mathematical results obtained in early 30's in the 20th century, we have a framework in which we clearly see that these two views are principally different. The question is, does from the core follow that a line cannot be in some sense curved so that two lines *can* meet in two points?

Remark 0.5 A note on terminology. We say that a set of axioms $\mathscr{A}$ is *consistent* if we cannot deduce a contradiction (a statement of the form $\varphi \wedge \neg\varphi$) from $\mathscr{A}$. Note that if we can deduce a single contradiction from $\mathscr{A}$, then we can deduce *every statement* from $\mathscr{A}$. This means that $\mathscr{A}$ is completely useless. (See Exercise at the end of this text).

Attempts to prove or defend the fifth postulate

Returning back to history, we state one of the many attempted "proofs" that E5 follows the core. Proclus, an early commentator of Elements, suggested to following argument as a proof that E5 is in fact derivable from the core: He tried to prove that there is at most one parallel line to a given p which passes through Q. Let q be a parallel line passing through Q. He argued that any other line r which passes through Q must enter between the lines p and q. The distance from r to q would be increasing and because the distance between parallel lines p, q must be constant, the line r must eventually meet p, thus showing that there is at most one parallel.

What is the flaw with Proclus' argument? The flaw is that the assumption that the distance between two parallel lines is constant is in fact equivalent to E5. See Exercises at the end of this section for more equivalences to E5.

After the decline of the Hellenistic (and Roman) civilisation in Middle Ages, the study of Elements was taken over by Arab and Islamic commentators (they also gave a number of false proofs).

The revival of Western interest started at the 16th century. Curiously, it was the Newton's use of Euclid's geometry in his *Principia Mathematica* to formulate the theory of gravitation that may have stalled the mathematical study of the postulates. They were taken to be "the description of the reality", and hence the postulates became absolute: true, or false. It was blasphemous, to some extent, to claim that E5 may be false. One might say that the mathematical sophistication at the time of Newton was certainly high enough to allow arguments for the possible falsity of E5, but "psychology" prevented them from viewing the issue from the strictly logical viewpoint. This "psychological" approach to space was later supported by Immanuel Kant's *a priori* intuition of space and time (these categories make possible all perception, and hence cannot be verified or refuted by experiments – which of course employ perception).

Slowly, some mathematicians at least derived some consequences in geometries where E5 was taken to be false. For instance Gerolamo Saccheri, an Italian Jesuit, published in 1733 a treatise *Euclid Freed of Every Flaw.* Saccheri was able to formulate three geometries which are supposedly consistent with the core: one where every triangle has the sum of its angles two right angles (case E), one where the sum is always less than two right angles (case L) and one where the sum is greater than two angles (G). However, still embedded in the psychological tradition of E5 being true, he tried to derive a contradiction in cases L and G, thus proving E5. He managed to show that, using the tacit assumption E0, the case G can never happen. But he was unable to refute the case L. He even drew some nontrivial consequences for geometries satisfying L, but than (in apparent desperation from the lack of mathematical arguments) supposedly refuted L by some theologically based arguments about infinity.

Following Saccheri, a Swiss mathematician J.H.Lambert derived some new consequences from the case L, for instance that there would be what philosophers called "the absolute measure of length": a triangle would be uniquely determined by its angles. This was universally (most notably by one Wolff, a follower of Leibniz) considered as counterintuitive, and hence false.

Discovery of non-Euclidean geometries

By discovery we mean in this section the study of the consequences of the negation of E5, disregarding its "truth" or "falsity" status. This approach witnesses a gradual change in paradigm of mathematics: the study of some principles solely due to their mathematical appeal, with the hidden supposition that "a nice mathematics cannot be a false mathematics".

Note however that this does not prevent a legitimate discussion which geometry is true in our universe. Indeed, it just tells us that if we think that the core describes a unique geometry, than we are sadly mistaken.

Hyperbolic geometry

Two names are connected with the discovery of the non-Euclidean "hyperbolic"[4] geometry: a Hungarian mathematician J.Bolyai (1802–1860) and N.Lobachevskii (1792–1856) in Russia.[5]

In a world where the exchange of ideas was notoriously slow or close to non-existent, the two men did practically the same work without knowing about each other. Here is their point of departure:[6]

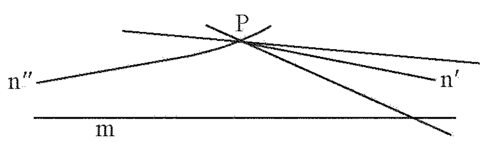

Fig.1

In their geometry, the postulate E0 holds true, and hence there exists at least one parallel to a given line m in every point P (see Claim 0.3). As Fig 1 suggests, however, they considered a situation in which there exists infinitely many parallels meeting P, thus failing E5. One group of parallels will be above the line n' and the other group above the line n''.

[4]Taken from the *hyperbolic paraboloid*, a saddle-shaped place, which can be used to model this geometry.

[5]The role of another famous German mathematician C.F.Gauss (1777–1855) in discovery of non-Euclidean geometries is disputed. He did not publish comprehensive results in the area, although he later claimed that he knew about them even before he read the works of J.Bolyai or N.Lobachevskii. This claim cannot be verified, however.

[6]This Fig.1, and also Fig's 3 and 4, are taken from the book The Road to Reality, by Roger Penrose. Their copyright is hereby acknowledged.

Elliptic, or Riemann geometry

A German mathematician B.Riemann (1826–1866) developed in the framework of his study of geometrical *manifolds* in complex analysis theory of generalized spaces, among them a space which is called elliptic, or Riemann, in his honour. In this space, the postulate E0 fails, and moreover, there are no parallels to p through a point P. We will describe a canonical space for this geometry in Fig. 4 (a two-dimensional surface of a ball).

Consistency of non-Euclidean geometries

Implicit in the work of J.Bolyai, N.Lobachevskii, and G.Riemann concerning abstractly defined "formal spaces" was the argument that the postulate E5 is independent of the core (plus E0). That is both E5 and its negation is consistent with the core (plus E0). The Riemann geometry shows that the core does not prove E0, and moreover if E0 fails, it is consistent that there are no parallels.

It is a shift in paradigm, to allow for "formal constructions" and "formal spaces" and claim that they are relevant to questions such as whether or not is E5 deducible from the core. We will learn more details in subsequent lessons, but in intuitive words this change in paradigm represents the move from the question such as "is φ true in our world, or in the realm of number, etc..." to "can φ be deduced from the axioms which we take to describe our world, or the realm of numbers, etc...". In principle, it may be that these two concepts are equivalent, but pending the exact formulation of "deduction" means, it transpired in 1930's by work of the Austrian mathematician K.Gödel that these concepts are essentially different.

The following theorem is an informal formulation of one of the Gödel's result, where we use (in keeping with the common practice) the word *model* to refer to the universes discussed above:

Theorem 0.6 (K.Gödel) *(**) A proposition φ is deducible from a set of axioms $\mathscr{A}$ if and only if φ is true in every model where all axioms in $\mathscr{A}$ are true.*

In fact for our purposes, we need just one (the easier) implication from Theorem 0.6: if φ is deducible from a set of axioms $\mathscr{A}$, then it is true in every model of $\mathscr{A}$, or equivalently, if there is a model where all axioms in $\mathscr{A}$ and $\neg\varphi$ hold, then φ is not deducible from $\mathscr{A}$.

It now follows that we can show that E5 is not deducible from the core by constructing a model where the axioms or postulates in the core are true, but E5 fails.

The first mathematician who explicitly stated that it is consistent that the core holds and simultaneously the negation of E5 holds was an Italian Eugenio Beltrami.

Model 1: Our space

First we state the obvious: the axioms E1–E5 are true in the intuitive world as we see it. That is why we call our space an *Euclidean space*. In general, any space which satisfies E1–E5 (and has the usual notion of distance) is Euclidean. We say that it is *consistent* (relative to E1–E4) that E5 is true.

Model 2: Hyperbolic space

Consider the following picture:

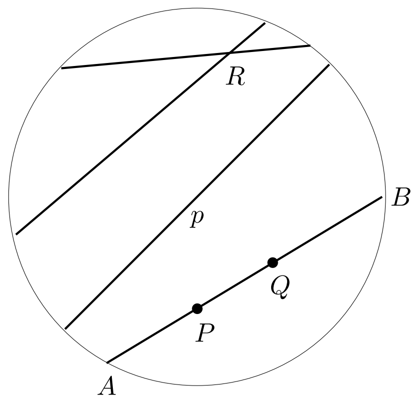

Fig.2: Hyperbolic space (the projective, or Felix Klein's,[7] representation)

We will show that under proper interpretation, the space visualised by the picture satisfies the axioms E0,E1–E4, ¬E5, thus showing by Theorem 0.6 that E5 is not deducible from E0,E1–E4.[8] Geometry connected with spaces like this is called the *hyperbolic geometry.* In our interpretation, the universe is just the interior of the circle (without its boundary), a point will be the usual point, a line will be a line segment going from one edge of the circle to the other edge.[9] We also need to interpret the notion of a distance: given two points P, Q on the line connecting (in our Euclidean space) the points A, B (which are however *not* in the universe of the hyperbolic geometry), we measure their distance $d(P, Q)$ by the following formula

$$d(P, Q) = -\frac{1}{2} \log \mathrm{CR}(A, P, B, Q), \tag{0.3}$$

where $\mathrm{CR}(A, P, B, Q)$ is called the cross-ratio of the the points A, P, B, Q is defined by

$$\mathrm{CR}(A, P, B, Q) = \frac{|AP|}{|AB|} \frac{|QB|}{|PQ|}. \tag{0.4}$$

The import of the formula is that if we for instance move the point Q to B, then the distance $|QB|$ tends to zero, and so the cross-ratio tends to zero, and the distance $d(P, Q)$ tends to infinity.

We will now deal with the axioms:

- E0. This postulate is clearly true.
- E1. This postulate is clearly true.
- E2. This is true, if we interpret the extension using the formula (0.3): given a line segment PQ, we can extend it so that the length of the segment is greater than any given $r \in \mathbb{R}$.
- E3. A circle is just an object the points of which have the same distance from the origin: we do have a distance as in (0.3), so we have a circle. Note that the representation of a circle in our Euclidean picture Fig 2 may not look like a circle, but it is a circle inside of the space.
- E4. E4 is very vaguely formulated. Instead, we introduce another model which has correct angles (and thus implicitly verifies E4; note that angles in Felix Klein's representation are measured differently). The model with correct angles is called a *conformal representation* of the hyperbolic geometry, sometimes referred to as a *Poincaré disc*:[10]

[7] A German mathematician, 1849-1925. The word "projective" refers refers to a projective transformation of the usual plane to that of the interior of the circle. Such transformation preserve the "cross-ratio" defined below.

[8] * In fact we show that if E0,E1–E4, E5 are consistent, so are the postulates E0,E1–E4,¬E5. In other words the Euclidean geometry is consistent if and only if the non-Euclidean geometry is consistent.

[9] The notion of a circle and of an angle is introduced below.

[10] J.H.Poincaré was a famous French mathematician and physicist, 1854–1912.

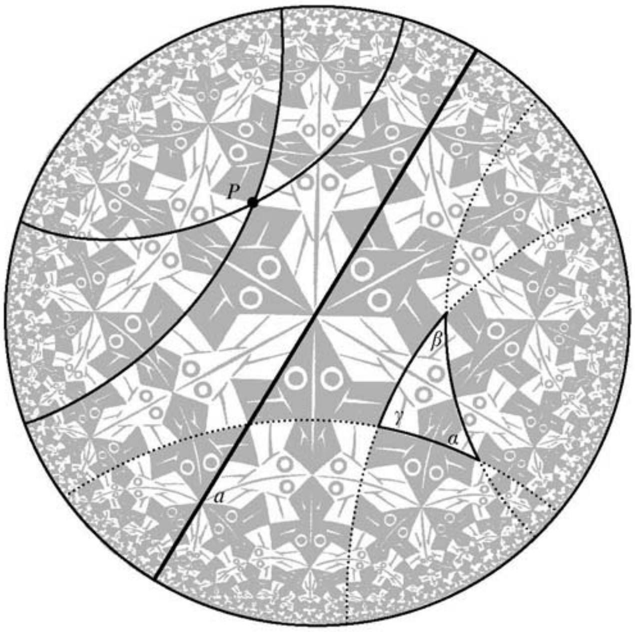

Fig.3: Poincaré disc (with one of the Escher's picture in the background to illustrate the nature of shapes in this representation). Notice two parallels in the point P.

Instead of lines, use parts of circles intersecting our given circle in the (Euclidean) right angle (see Fig 3). In this representation, angles in the hyperbolic geometry are measured exactly the same as in the Euclidean geometry. In the projective representation, angles are measured differently.
Note that in the conformal representation, a "straight line" is modeled as a curve, a part of circle. It shows that although we can imagine the core postulates to describe "straight lines", they fall short of this task.

¬E5. As illustrated in Fig 2, there is more than one line which goes through R and does not meet p: in fact there are infinitely many of these.

Model 3: Riemann spaces

Great circles (hlavní kružnice) on the sphere as lines.

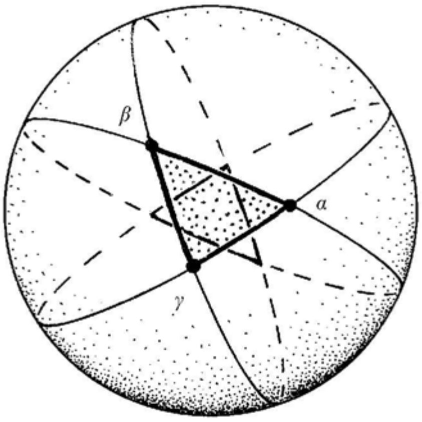

Fig.4: Spherical two-dimensional space.

***

Exercises.

1. Show that proposition 17 "in any triangle the sum of any two angles is less than two right angles" implies the converse in E5+: "if on the other hand, the interior angles are equal to two right angles, then the lines will never meet."

2. Show that E0,E1–E4 is equivalent to E1*,E2–E4, where E1* states that "For all two distinct points there is exactly one line passing through the two points."

3. *Show that a set of axioms $\mathscr{A}$ is inconsistent (i.e. we can deduce a contradiction from $\mathscr{A}$) if and only if we can deduce every statement from $\mathscr{A}$.

4. *Playfair's axioms* states that "Given any line p in plane and any point P in that plane that does not lie on p, there is at most one line q that passes through the point P and does not meet the line p." (note that this is a weakened version of E5*). Show that over E1–E4, Playfair's axiom is equivalent to E5.

5. ** The following statements are equivalent to E5 over the core.
   (i) The sum of the angles in every triangle is 180°.
   (ii) There exists a triangle whose angles add up to 180°.
   (iii) The sum of the angles is the same for every triangle.
   (iv) There exists a pair of similar, but not congruent, triangles.
   (v) Every triangle can be circumscribed.
   (vi) If three angles of a quadrilateral are right angles, then the fourth angle is also a right angle.
   (vii) There exists a quadrilateral of which all angles are right angles.
   (viii) There exists a pair of straight lines that are at constant distance from each other.
   (ix) Two lines that are parallel to the same line are also parallel to each other.
   (x) Given two parallel lines, any line that intersects one of them also intersects the other.
   (xi) In a right-angled triangle, the square of the hypotenuse equals the sum of the squares of the other two sides (Pythagoras' Theorem).
   (xii) There is no upper limit to the area of a triangle.

Remark 0.7 Consider the following proposition which is know as Golbach's conjecture "Every even number > 2 is a sum of two primes" (where 1 is not considered as a prime). It is not known today whether or not is Golbach's conjecture true in the natural model of numbers. However, a "natural model" is a very abstract thing (being an infinite set, no one can really analyze it), and so the question is approached from the angle whether or not is Golbach's conjecture deducible from the axioms which describe natural numbers.[11] In analogy with E5 we now see that perhaps there is a third option regarding the Golbach's hypothesis: it may be independent, that is neither true of false. This result is widely considered as not very attractive: we got used to different, but still completely sensible, geometries, but the existence of different, and still sensible, natural numbers is another kettle of fish (one version of natural numbers would be the one where Golbach's conjecture holds true, the other one where it is false).[12]

[11] By Gödel's results we know that this angle is strictly weaker, but in practice this is the best we have.

[12] We sweep some things under the carpet here: one can show that if Golbach's conjecture is independent, then in fact it must be *true* in the real natural numbers $\mathbb{N}$.