Tarskian Geometry and Ruler-and-Compass Constructions

Michael Beeson

Stanford Logic Seminar
April 15, 2014

Alfred Tarski 1918, student in Warsaw, age 17

Famous Axiomatizations of Geometry

- Euclid's *Elements* 300 BCE
- Pasch 1882
- Hilbert 1899
- Pieri 1900 and 1908
- Tarski 1926–27

Euclid and Hilbert had important things in common:

- Systematic treatment building on the work of others
- Emphasis on and development of axiomatic method

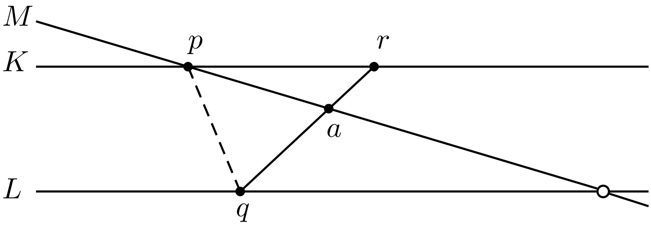
Euclid and Hilbert

Euclid's postulates

1. To draw a straight line from any point to any point.
2. To produce a finite straight line continuously in a straight line.
3. To describe a circle with any centre and distance.
4. All right angles are equal to one another.
5. 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 angles less than two right angles.

Euclid's Postulate 5

Line pq falls on straight lines M and L making angles on the right side less than two right angles. The point indicated by the open circle is asserted to exist.



Line K can always be constructed parallel to L so the important part about Euclid 5 is that any *other* line through p has to meet L.

Doubts about Euclid 5

It was very early felt that Euclid 5 might not quite deserve the status of a "postulate."

- It seems less fundamental
- The meeting point might be very far away, so we can't just "see it in front of our faces".
- It would nicer to have a proof of Euclid 5 and not have to assume it.
- Or at least to replace it with something that seems more fundamental

Theon's manuscript, perhaps 360 CE

Manuscript found in Vatican Library 1808

Then millennia went by . . .

- The library at Alexandria was burned three times under dramatic circumstances (48 BCE, 270 CE, 391 CE)
- Government support dried up, and the Empire itself fell.
- But the Eastern, Greek-speaking part survived another millenium. Its capital was Byzantium, later Constantinople, now Istanbul. The *Elements* survived there.
- The monsoon winds were known, and in the spring a large fleet sailed from the Red Sea area for India, returning in the fall. This commerce may have carried Euclid, too, but I can't find evidence, only rumors.

Adventures of the *Elements*

- From Byzantium to Arabia about 760. Translated under the patronage of the caliph, Harun al-Rashid, about 800.
- Back to Europe with the Muslim conquest of Spain (711-1031).
- New Arabic translation by Nasir al-Din al-Tusi, 1220.
- Sanskrit translation by Jagannātha Samrāt in 1719: the *Rekhāganita*. Was this really the first arrival of the *Elements* in India??
- Translated from Arabic to Latin, Abelard of Bath, 1120.
- Printed in 1482 in Venice, 30 years after the Bible, in Latin translated from Arabic translated from Greek.
- Only one Arabic translation has ever been published, but "thousands" of manuscripts are on microfilm at the Juma Al-Majid Heritage and Cultural Center in Dubai, awaiting examination by historians.
- Theon's Greek edition recovered 1533.

Taken to China by Jesuit missionaries, 1582.

Failed attempts to prove Euclid 5

Failed attempts were published by

- Simplicius (Byzantine, sixth century)
- Al-Abbās ibn Said al-Jawharī (Persian, ninth century)
- Nasir Eddin al Tusi (1201-1274), who did most of his scientific work while imprisoned by the armies of Genghis Khan. Apparently POWs were well-treated!

- Legendre (1752-1833), who continued to try to prove Euclid 5 until the year of his death, when he published a collection of his failed attempts.
- Lambert (whose wrong proof was published posthumously)
- A Ph. D. thesis in 1763 found flaws in 28 different alleged proofs of Euclid 5.

Non-Euclidean geometry

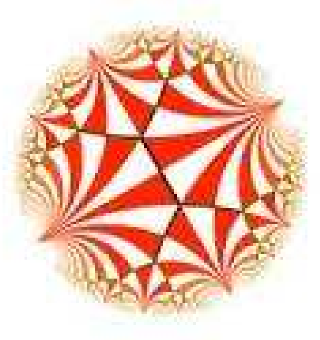
The method of *reductio ad absurdum*, or for short just *reductio*, refers to proof by contradiction. Girolamo Saccheri tried to prove Euclid 5 this way.

- He made long deductions and was in some sense the creator of non-Euclidean geometry.
- But he did not, apparently, understand what he had done, i.e., he continued to believe that more of these deductions would eventually lead to a contradiction.
- For the history of logic, the important point is that Sachceri, and everyone else, still believed geometry was about the one true universe, the space we live in, and logical reasoning just a tool for uncovering truths about that space.

Non-Euclidean Geometry

- Eventually, Gauss, Bolyai, and Lobachevsky realized that it is consistent to assume that there can be several parallels to line L through point P.
- Exactly how this happened is very interesting, both for biographical details and the development of mathematical philosophy. But it is not today's topic.
- Both Euclid 5 and its negation are consistent with Euclid's first four postulates.

A model of non-Euclidean geometry



The picture shows the Poincaré model, in which lines are circular arcs meeting the unit circle at right angles (including diameters of the unit circle), and distance is defined by a certain formula so that the boundary is infinitely far from any interior point.

Credit for the first construction of such a model goes to Beltrami (1868).

Origins of Tarskian Geometry

- 1926-1927 lectures at the University of Warsaw
- Tarski was 26 years old.
- First-order logic was very new.
- The completeness theorem had not even been conjectured yet!
- It was first conjectured in 1928 in Hilbert-Ackermann, p. 68.
- Proved the following year (1929) by Gödel.
- Tarski was on the cutting edge.

Congruence and betweenness

- Congruence of segments $ab \equiv cd$ or just $ab = cd$
- Really a relation between 4 points.
- Betweenness $\mathbf{B}(a, b, c)$ "b is between a and c (on a line)"
- Euclid never mentioned betweenness but assumed its properties implicitly, never proving that lines appearing to meet actually meet.
- But the parallel postulate is an exception to that!
- Pasch 1882 introduced betweenness.
- Use betweenness to define order on a line, and "inside" or "outside" a circle.

Strict or non-strict betweenness?

- Strict means that a is not between a and b.
- Non-strict allows a to be between a and b.
- Hilbert used strict betweenness
- Tarski used non-strict betweenness
- To avoid confusion:

  $\mathbf{B}(a, b, c)$ for strict betweenness,

  $\mathbf{T}(a, b, c)$ for non-strict.
- Of course each can define the other.

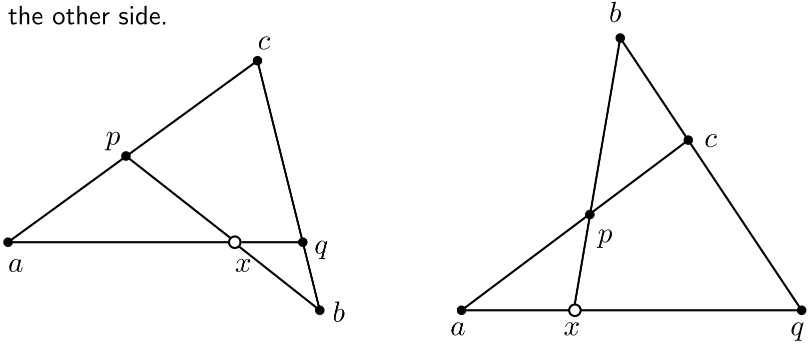
Points, Lines, Angles, and Circles

- Hilbert, following Euclid, took all four as primitive.
- So he needed more relations: “incidence” (a point lies on a line), congruence of angles.
- Tarski, following Pieri, used a language with variables for points only.
- Circles and lines are just ways of constructing more points by their intersection points with other circles and lines.
- Angles can be treated as triples of points, as in “angle ABC”

Pasch's Axiom

This is the axiom that Euclid forgot. Pasch invented it in 1882. It is meant to say that if a line enters a triangle it must exit again. Tarski considered the following versions, which are valid even in 3-space. Tarksi's axiom A7 was first one of these, then later the other one.

Figure: Inner Pasch (left) and outer Pasch (right). Line pb meets triangle acq in one side. The open circles show the points x asserted to exist on the other side.



Moritz Pasch

Extending segments

Euclid's postulate, "To produce a finite straight line continuously in a straight line", corresponds closely to Tarski's axiom A4 that we can extend segment ab by segment cd, to produce a point $p = ext(a, b, c, d)$ such that

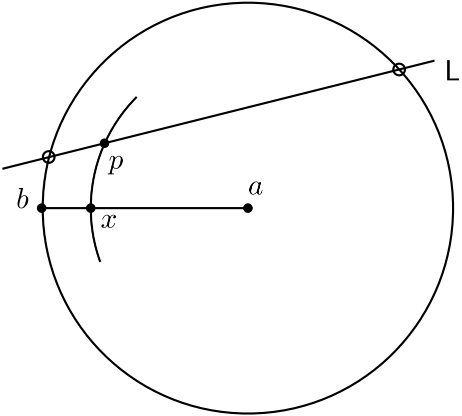
$$\mathbf{T}(a, b, p)$$

$$bp = cd$$

Tarski's axiom says a little more than Euclid's, because it says we can extend ab by a given amount, not just by an unspecified amount.

Line-circle continuity

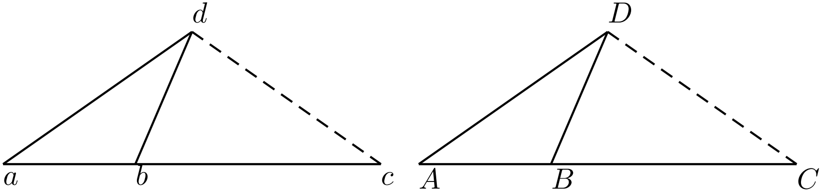
If line L has a point p inside circle C then L meets C.



Segment-circle continuity: if L has a point inside C and a point outside C, then between those two it meets C.

The 5-segment axiom A5

Tarski's 5-segment axiom. cd is determined.



Tarski's 5-segment axiom is a thinly-disguised variant of the SAS criterion for triangle congruence. The triangles we are to prove congruent are dbc and DBC. We are given that $bc = BC$ and $db = DB$. The congruence of angles dbc and DBC is expressed in Tarski's axiom by the congruence of triangles abd and ABD, whose sides are pairwise equal. The conclusion, that $cd = CD$, give the congruence of triangles dbc and DBC.

Euclid and SAS

- Euclid tried to prove SAS in I.4
- Commentators from Geminus on did not accept the proof.
- Euclid should have taken SAS as a postulate.

All right angles equal

- This postulate of Euclid does not follow easily from A5
- But Tarski does not have to assume it: although it is difficult, it can be proved.

First-order geometry

Tarski also added a "continuity axiom" (really an axiom schema) saying essentially that any first-order Dedekind cut is filled.

- This was historic, as the first clear distinction in geometry between first-order and second-order or set-theoretic axioms.
- The corresponding "axiom" in Hilbert's theory wasn't even second-order, but showed that confusion still existed between syntax and semantics.
- The full A11 does not concern us today, as it goes beyond ruler-and-compass constructions.

History of Tarski's geometry

- Prepared for publication in 1940
- But there was a war on, and the paper only appeared in 1967
- A version appeared in a RAND tech report, 1948, revised 1951
- Tarski gave a course at Berkeley, 1956-57 on the foundations of geometry. He and his students Eva Kallin and Scott Taylor proved six of the axioms, leaving only 13.
- Two more were proved in H. N. Gupta's Ph. D. thesis, 1965.
- Gupta also proved inner Pasch implies outer Pasch, and constructed midpoints of segments and perpendiculars based on A1-A9 (no circles).
- Starting with Tarski's notes, and probably a draft of Gupta's work, Wanda Szmielew prepared a manuscript developing geometry from Tarski's axioms.
- This eventually, after her death in 1983, became part I of SST (Szmielew, Schwäbhauser, and Tarski), published in 1985.

Tarski's theory and ruler-and-compass constructions

Not all the points asserted to exist correspond to ruler-and-compass constructions:

- With a ruler and compass, I cannot extend a null segment aa, as I need two distinct points to orient my ruler.
- Tarski uses nonstrict betweenness $\mathbf{T}(a, b, c)$ to formulate inner Pasch and outer Pasch, so there are degenerate cases where one cannot find the point asserted to exist as an intersection of two distinct lines.
- To fix this, we just use $\mathbf{B}$ instead of $\mathbf{T}$ in the two Pasch axioms, and only extend non-null segments.
- Now, these axioms do correspond to ruler-and-compass (in fact, ruler only) constructions.
- But is there a price to be paid? That is, is Pasch with $\mathbf{B}$ equivalent to Pasch with $\mathbf{T}$, or not?
- To answer that question, we have to consider Tarski's other axioms.

Axioms for Betweenness

Tarski's final theory had only *one* betweenness axiom:

$\mathbf{T}(a,b,a) \rightarrow a = b$ identity for betweenness, A6

or in other words

$$\neg \mathbf{B}(a,b,a).$$

But originally there were also these (eliminated in 1956-65):

$\mathbf{T}(a,b,c) \rightarrow \mathbf{T}(c,b,a)$ (A14), symmetry

$\mathbf{T}(a,b,d) \wedge \mathbf{T}(b,c,d) \rightarrow \mathbf{T}(a,b,c)$ (A15), inner transitivity

$\mathbf{T}(a,b,c) \wedge \mathbf{T}(b,c,d) \wedge b \neq c \rightarrow \mathbf{T}(a,b,d)$ (A16), outer transitivity

$\mathbf{T}(a,b,d) \wedge \mathbf{T}(a,c,d) \rightarrow \mathbf{T}(a,b,c) \vee \mathbf{T}(a,c,b)$ (A17), inner connectivity

$\mathbf{T}(a,b,c) \wedge \mathbf{T}(a,b,d) \wedge a \neq b \rightarrow \mathbf{T}(a,c,d) \vee \mathbf{T}(a,d,c)$ (A18), outer connectivity

We have to put back A14 and A15

The degenerate cases of Pasch were used to eliminate A14 and A15. This was not a good bargain: Tarski gave up the connection to ruler-and-compass constructions just to eliminate these two simple and intuitive axioms.

$\mathbf{T}(a,b,c) \rightarrow \mathbf{T}(c,b,a)$ (A14), symmetry
$\mathbf{T}(a,b,d) \wedge \mathbf{T}(b,c,d) \rightarrow \mathbf{T}(a,b,c)$ (A15), inner transitivity

If we put them back, we can derive the degenerate cases of Pasch.

- Now the Skolem function for inner Pasch corresponds to a ruler-and-compass construction!
- So all the terms of this theory correspond to ruler-and-compass constructions, starting from the three fundamental non-collinear points α, β, and γ, or from "given points" that are parameters of the terms.

Congruence axioms

$ab = ba$	(A1) Reflexivity of equidistance
$ab = pq \wedge ab = rs \rightarrow pq = rs$	(A2) Transitivity of equidistance
$ab = cc \rightarrow a = b$	(A3) Identity of equidistance

Dimension axioms

Today we will consider only plane geometry. Then the dimension axioms specialize to

- there are three non-collinear points α, β, and γ. (A8)
- given any two points p and q, and three points a, b, c each equidistant from p and q, then a, b, and c are collinear. (A9)

Note that A9 fails in three dimensions, but holds in two.

Parallel Axiom

Just as Saccheri, Legendre, Bolyai, Gauss, and Lobachevsky discovered, there are *many* propositions that are equivalent to the parallel axiom of Euclid. My favorite is the one Szmielew chose in her original manuscript of 1965: "given three non-collinear points, there exists a point equidistant from those three." In other words, a triangle has a circumscribed circle.

A different version (due to Tarski) appears in SST. For the moment, we stick with Euclid's version, whose only disadvantage is that when formulated in terms of points only, it is a bit longer than the others.

Summary of Tarski's axioms

- Three congruence axioms, A1-A3
- Three betweenness axioms, A6, A14, A15
- Segment extension, A4
- Five-segment axiom, A5
- inner Pasch, A7
- dimension axioms, A8-9
- Parallel axiom, A10
- Line-circle continuity (consequence of A11)

Skolemized Tarski

We can formulate a quantifier-free version of Tarski's (ruler-and-compass) geometry A1-A9 (without the parallel axiom), if we add function symbols as follows:

- $ext(a, b, c, d)$ for the extension of ab by cd, when $a \neq b$.
- $ip(a, b, c, b, q)$ for the point asserted to exist by inner Pasch with $\mathbf{B}$.
- $i\ell c(a, b, c, d)$ for the intersection of the $Line(a, b)$ with the circle of center c passing through d, if that intersection exists.

Defined terms in this language correspond to ruler-and-compass constructions, and the "definedness conditions" for terms can be correctly formulated within the language. We can specify, for example, that these Skolem functions just return their first argument when they are (geometrically) "undefined."

A property of constructions

Suppose we have constructed, or are somehow given, a finite number of points (a "configuration"), in the usual plane $\mathbb{R}^2$. The **diameter** of the configuration is the maximum distance between two of its points.

If we now apply the construction functions ip, $i\ell c$, and ext to some points of the configuration to obtain new points, the diameter of the configuration at most doubles. Indeed, only ext increases the diameter at all, and ext at most doubles the diameter.

This observation is due to Julien Narboux and his colleagues Pierre Boutry and Pascal Schreck.

Herbrand's theorem

This is a general logical theorem about any quantifier-free theory T in first-order logic. It says that if $T \vdash \exists x\, A(x)$ then there exist terms $t_1, \ldots, t_n$ such that

$$T \vdash A(t_1) \vee A(t_2) \ldots A(t_n).$$

The formula A can have other variables not shown explicitly; which of the t_i works probably depends on the values of those variables. Of course, if x is a list of variables, each t_i is a corresponding list of terms.

A fertile three years

- Herbrand proved his theorem in 1928, while Tarski was lecturing in Warsaw, Hilbert-Ackermann was being sent to the printers, and the completeness theorem was in gestation.
- He was a 20-year old Ph. D. student at the time.
- Herbrand was on the right track– he might have proved the completeness theorem, if Gödel hadn't done it at about that same time (publication date 1929).
- And if he hadn't died tragically in a mountain-climbing accident at age 23.

Herbrand during his fatal climbing trip

Herbrand's theorem in geometry

In geometry, where each t_1 represents a ruler-and-compass construction, Herbrand's theorem tells us there will be finitely many constructions, one of which will work in each case.
For example, suppose $A(a, b, p, x, y)$ says that $a \neq b$ implies that $x \neq y$ and x is on $Line(a, b)$ and $y \neq x$ and $xy \perp ab$ and p lies on $Line(x, y)$. Then $\exists x, y A(a, b, p, x, y)$ says that we can find a line through p perpendicular to ab.

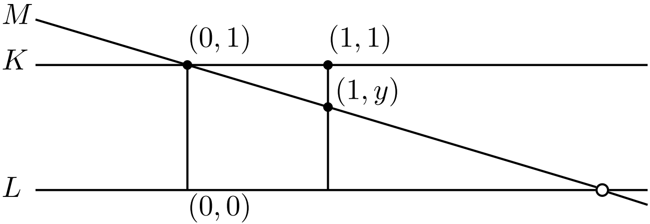
We usually would argue by cases: if p is on ab, we "erect" a perpendicular to ab at p, and if not, we "drop" a perpendicular to p. These two constructions correspond to terms t_{ij} (containing variables a, b, p) such that

$$T \vdash A(a, b, p, t_{11}, t_{12}) \vee A(a, b, p, t_{21}, t_{22})$$

So in this case, Herbrand's theorem only tells us what we already know; but it tells us without doing any geometry at all that any proof of the existence of perpendiculars must provide some number (not necessarily two) of constructions, one of which always works.

Euclid 5 does not follow from A1-A9 and line-circle continuity

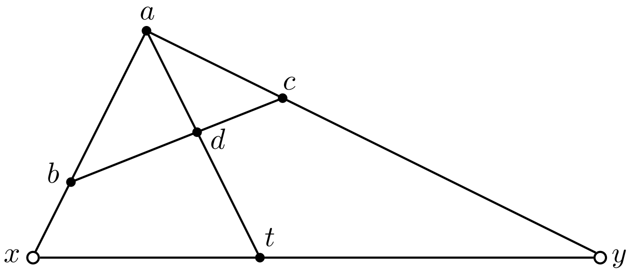
Proof (with Narboux, Boutry, and Schreck.) Suppose it did. Then by Herbrand's theorem, there would be n terms $t_1(y), \ldots, t_n(y)$ such that, provably in A1-A9 plus line-circle, one of the $t_i(y)$ lies on both M and L in this figure:



Now let k be the total number of symbols in all the t_i. The diameter of any configuration constructed from the illustrated points in k construction steps is at most 2^k. But if we choose $y = 2^{-k-2}$, the intersection point of M and L lies more than 2^k from the origin, so it cannot be constructed by one of the t_i. That contradiction completes the proof.

Constructions corresponding to Euclid 5

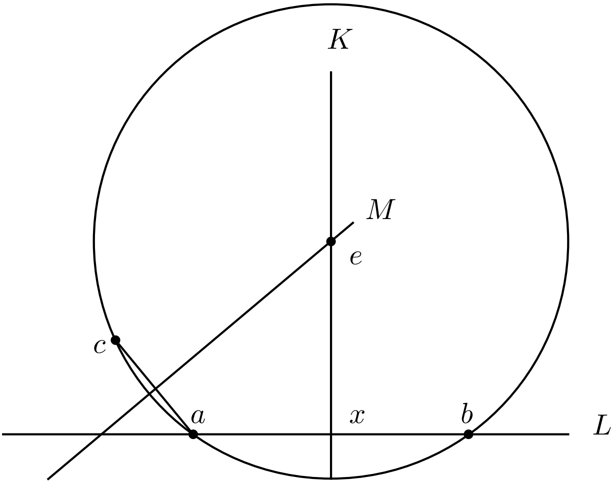
Tarski's version of Euclid 5 does not correspond to a ruler-and-compass construction. It is this:



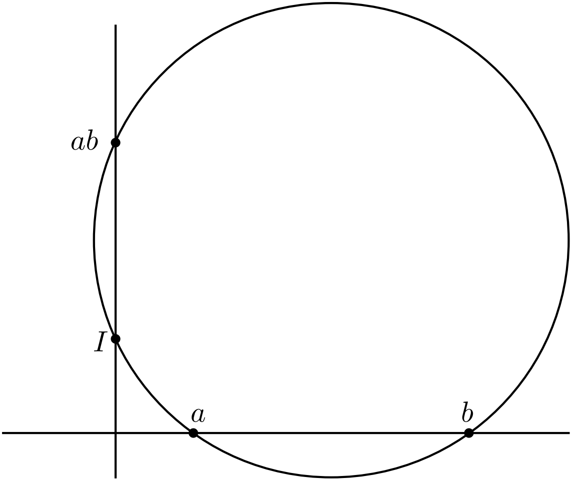
Point t in the interior of angle bac can be connected to the two sides of the angle by some line xy.

Triangle Circumscription (Szmielew's Parallel axiom)

If a, b, and c are not collinear, there exists a point e equidistant from a, b, and c.



Hilbert Multiplication



"Multiplication is defined" is equivalent to Euclid 5

Ruler-and-compass constructions

Herbrand's theorem tells us:

Theorem

Suppose Tarski's A1-A10 plus line-circle continuity proves

$$\forall x\, A(x) \rightarrow \exists y\, C(x, y)$$

with A and C quantifier-free. Then there is an integer n and ruler-and-compass constructions $t_1(x), \ldots, t_n(x)$ such that one of them always constructs y from x. Here both x and y can be lists of variables.

The constructions are of a very specific kind:

- center of a circumscribed circle, $center(a, b, c)$
- segment extensions $ext(a, b, c, d)$
- line intersection points given by inner Pasch: $ip(a, p, c, b, q)$

Reduction of *center* to simpler constructions?

You may think that $center(a, b, c)$ is not a "fundamental" construction, but something you construct in several ruler-and-compass construction steps.

- If we can drop perpendiculars and bisect segments, we can get $center$ from Euclid 5.
- Gupta showed we can do those things from A1-A9, i.e. without circles or any parallel postulate.
- Even with line-circle allowed it is not trivial, but easier.
- But you still need *some* construction for Euclid 5, such as the one Euclid mentions.
- You may need a *very* long ruler.

Two fundamental theorems about Tarskian geometry

Our independence proof can be made to work for the full first-order continuity scheme A11 in place of just line-circle continuity, but unlike the proof you saw, it is not easy. The explanation depends on the following well-known, but difficult to prove, facts.

- Addition and multiplication can be defined geometrically using A1-A10.
- The models of A1-A10 are all isomorphic to planes $\mathbb{F}^2$ over a real-closed field $\mathbb{F}$.

The first fact is due to Descartes, Hilbert, Tarski, Szmielew, and Gupta, and is the main point of Part I of SST.
The second fact is due to Tarski, and was the main theorem he emphasized about his geometry.

There won't be time in this talk to discuss the extension to A11.

Comparison to models of non-Euclidean geometry

While it's amusing and pleasant to show that the independence of Euclid 5 is a consequence of Herbrand's theorem, in the interest of full disclosure, we should point out that

- The Herbrand's theorem proof only works for first-order geometry, while
- the models of non-Euclidean geometry also work for second-order geometry.

A model of Max Dehn

There is another way to prove the independence of Euclid 5, due to Max Dehn:

- Let $\mathbb{F}$ be a non-Archimedian Euclidean field (ordered, and positive elements have square roots, and some element is larger than any integer).
- An element of $\mathbb{F}$ is "finitely bounded" if it is less than some integer.
- Let $\mathbb{M}$ be the finitely bounded elements of $\mathbb{F}$; the model we want is $\mathbb{M}^2$.
- It satisfies A1-A9, and line-circle continuity, but not A10, because there are lines through $(0, 1)$ with nonzero but infinitesimal slope.
- This model doesn't satisfy the whole schema A11 either, but that defect can be fixed using Tarski's work, just as for our proof.

Constructive Geometry

If we use intuitionistic logic, then in Herbrand's theorem, you can take the number of terms to be one.

That is, if you can prove $\exists y\, A(x, y)$, then there is a term (just one term) t such that $A(x, t(x))$ is provable.

- For example, is there a uniform construction of a perpendicular to line L through point p, without a case distinction as to whether p is on L or not?
- Yes, there is. So the theorem that for every line L and point p, there is a perpendicular to L through p is constructively correct.
- The construction is not claimed to be obvious. We are only pointing out what has to be done to prove constructively that there exists a perpendicular to L through any point p.
- Constructive proofs avoid case distinctions and proof by contradiction when proving things exist.

Constructive Tarskian Geometry

An esoteric subject, because you need to know

- Tarski's axiom system
- something about constructive (intuitionistic) logic and mathematics

I have written several papers about constructive geometry; the most recent is about a constructive version of Tarskian geometry. You can find the two most recent by googling their filenames:

ConstructiveGeometryAndTheParallelPostulate.pdf
AxiomatizingConstructiveGeometry.pdf