Brouwer and Euclid

Michael Beeson

July 18, 2017

Abstract

We explore the relationship between Brouwer's intuitionistic mathematics and Euclidean geometry. Brouwer wrote a paper in 1949 called *The contradictority of elementary geometry*. In that paper, he showed that a certain classical consequence of the parallel postulate implies Markov's principle, which he found intuitionistically unacceptable. But Euclid's geometry, having served as a beacon of clear and correct reasoning for two millennia, is not so easily discarded.

Brouwer started from a "theorem" that is not in Euclid, and requires Markov's principle for its proof. That means that Brouwer's paper did not address the question whether Euclid's *Elements* really requires Markov's principle. In this paper we show that there is a coherent theory of "non-Markovian Euclidean geometry." We show in some detail that our theory is an adequate formal rendering of (at least) Euclid's Book I, and suffices to define geometric arithmetic, thus refining the author's previous investigations (which include Markov's principle as an axiom).

Philosophically, Brouwer's proof that his version of the parallel postulate implies Markov's principle could be read just as well as geometric evidence for the truth of Markov's principle, if one thinks the geometrical "intersection theorem" with which Brouwer started is geometrically evident.

Contents

1

1   Introduction

Brouwer, in founding the philosophy of mathematics known as "intuition-ism", rejected many of the mathematical results that were obtained in the

nineteenth century or before. The rejected body of mathematics has become known as "classical mathematics". The word "classical" also brings to mind the world of ancient Greece, where Euclid and his predecessors laid the foundations of modern mathematical reasoning in the third century BCE. Euclidean geometry was for two millennia the *sine qua non* of careful reasoning; every educated person in Europe studied it; the American Declaration of Independence was modeled on Euclidean reasoning. Brouwer did not directly challenge Euclid by name in any publication, but in 1949 he published a paper [7] with the title, *Contradictority of Elementary Geometry.*

Brouwer was never much of a diplomat. If his personality had been more diplomatic, he might have pointed out that, on a certain reading, certain theorems of elementary geometry related to the parallel postulate may seem (intuitionistically) contradictory; but that every classical theorem permits various refinements, once constructive distinctions are taken into account, and Euclid's parallel postulate and its consequences are not exceptions.

What would Euclid have written, if he had come after Brouwer, instead of before? Of course, he would not have thrown up his hands, thinking geometry is contradictory, and gone into investment banking instead. We will show in this paper that, if one is careful about the formulation of the axioms, Euclidean geometry is perfectly consistent with Brouwer's intuitionism.

What Brouwer calls a "contradiction" has two parts: (i) Brouwer rejects a certain property of the ordering of points on a line known as Markov's principle, and (ii) Brouwer shows that a certain "intersection theorem" (a classical consequence of Euclid's parallel postulate) implies Markov's principle, even though it appears not to imply the law of the excluded middle. This result would be better summarized by the statement

> Intuitionistic "elementary geometry" must distinguish between different propositions classically equivalent to Euclid's parallel postulate.

In previous work [4, 3], such distinctions were made, and a coherent theory of constructive geometry developed. But that theory would not have met with Brouwer's approval, because Markov's principle is assumed as an axiom. The reason for that was simple pragmatism: it enables one to argue by contradiction and cases, as long as one is trying to prove betweenness and congruence assertions about specific points, rather than assertions that more points exist with certain properties.[1]

Thus, the door is open (and has been open for 68 years) for a development of Euclidean geometry that explicitly avoids not only the law of the excluded middle, but also Markov's principle. We call this theory "non-Markovian Euclidean geometry", or for short just "non-Markovian geometry." Perhaps it should be called "intuitionistic geometry", if one feels that the rejection of Markov's principle is fundamental to intuitionism.

---

[1] For logicians: allowing Markov's principle enabled the double negation interpretation to work, allowing the "importation" of a certain class of geometrical results, whose classical proofs are long and complicated.

The key concepts of non-Markovian geometry are the concepts of "distinct points" and "positive angles." The concept that a and b are distinct is written $a\#b$, and is stronger than simple inequality $a \neq b$. Intuitively, $a\#b$ means that we have a positive lower bound on how far apart a and b are, although in axiomatic geometry there is of course no notion of "distance." The concept that angle abc is "positive", written $0 < abc$, means intuitively that we have a lower bound on how different the directions of the rays bc and ba are.

Both these concepts will be *defined* in terms of betweenness and congruence, rather than be introduced as primitive. In particular the correct definition of "positive angle" is not obvious *a priori*. In order to ensure that the axioms do not imply Markov's principle immediately, Pasch's axiom must be restricted to assume that certain angles are positive. Then the definition of "positive angle" must be broad enough to permit the applications of Pasch that are needed to "bootstrap" geometry. But finally, we should be able to prove that a positive angle is simply the apex angle of an isosceles triangle (whose three points are distinct). Then Euclid Book I can be proved, when angles are assumed to be positive (and have positive supplements) and triangles, by definition of "triangle", have distinct vertices. In the last section we give a metamathematical theorem to the effect that this claim can be extended to (at least) Books II and III as well.[2]

There have been some papers on related subjects in the seven decades since Brouwer's rejection of "elementary geometry", and the obligation arises to explain in what relation the present work stands to those papers. Almost all of those papers were about projective geometry, or affine geometry, or Desarguesian geometry, rather than Euclidean geometry, and also were based on (or included) apartness, which is rejected in this work. Here is a brief, possibly incomplete, list of such papers. The first was Heyting's 1925 thesis (published two years later as [12] and again 34 years later as [13].) Heyting's student van Dalen continued work on intuitionistic projective spaces in [30, 31, 32]; see also [34] and [20, 21, 23, 22]. Aside from the previous work of the present author [3, 5], the only previous paper on constructive Euclidean geometry was by Lombard and Vesley [19], who followed Heyting in taking apartness as primitive.

2   Versions of the parallel postulate

The "parallel postulate" of Euclid has been reformulated many times in the history of geometry, as efforts to prove it led instead to many different equivalent propositions. But not all these versions are equivalent using intuitionistic logic. Euclid's version requires two lines to meet, if two specified angles "make less than two right angles." A different version due to the Englishman Playfair (1729), was popularized by Hilbert; it asserts the impossibility of two different lines parallel to a given line through the same point. In so doing it makes no existential assertion, unlike Euclid's version, which asserts the existence of an intersection point.

---

[2]This claim cannot be taken too literally, as of course Euclid does contain gaps and errors.

In this section, we will review several versions of the parallel postulate, and then discuss Brouwer's 1949 proof. All these versions of the parallel postulate, including Euclid's own, are consistent with intuitionistic logic. Brouwer's paper shows that his version of the parallel postulate implies a certain ordering principle, known as Markov's principle, which Brouwer believed to be contrary to the nature of the intuitionistic continuum, for reasons far removed from Euclidean geometry.

2.1 Euclid's parallel axiom

Euclid's postulate 5 is

> *If a straight line falling on two straight lines make 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 the two right angles.*

We consider the formal expression of Euclid's parallel axiom. Let L and M be two straight lines, and let pq be the "straight line falling on" L and M, with p on M and q on L. We think that what Euclid meant by "makes the interior angles on the same side less than two right angles" was that, if K is another line through p, making the interior angles with pq equal to two right angles, then M would lie in the interior of one of those interior angles (see Fig. 1).

Euclid did not define "angle", and did not define "lies in the interior of an angle", but these issues of precision have little to do with intuitionism. Assuming for the moment that we understand the notions of "angle" and "alternate interior angle", then we can state Euclid's parallel axiom, using three more points to "witness" that one ray of line M emanating from p lies in the interior of one of the interior angles made by K. Fig. 1 illustrates the postulate. The point asserted to exist is shown by a small open circle (a convention we will follow throughout).
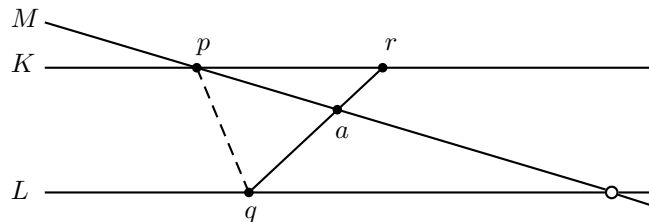
Figure 1: Euclid 5: M and L must meet on the right side, provided $\mathbf{B}(q, a, r)$ and pq makes alternate interior angles equal with K and L.

In this formulation, the point is asserted to exist "on the right side"; more precisely, "on the same side of pq as a." Euclid did not define "on the same side of", although his postulate uses that phrase. That notion was, apparently, first defined by M. Pasch in 1882 (on page 27 of [25]).

2.2 Playfair's axiom

"Playfair's axiom" is the version of the parallel axiom adopted by Hilbert in [15]. That version, unlike all the other versions, makes no existence assertion at all, but only asserts that there cannot exist two different lines parallel to a given line through a given point.
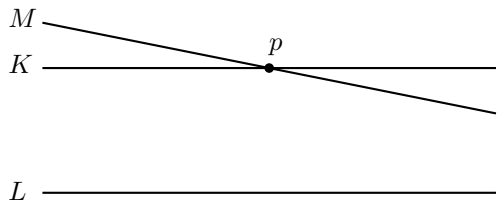
Figure 2: Playfair: if K and L are parallel, M and L are not parallel.

The conclusion of Playfair's axiom is that M and L are not parallel. By definition, parallel lines are lines that do not meet, so the conclusion is that M and L cannot fail to meet. That is, not not there exists an intersection point. Since $\neg\neg\exists$ is equivalent to $\neg\forall\neg$, no existential quantifier is needed to express Playfair's axiom.

2.3 Brouwer's intersection theorem

Brouwer considers a version of the parallel postulate similar to Euclid's, but without the "witness" a in the hypothesis testifying to the side on which the angles are less than a right angle. It is similar to a variant of Playfair's axiom in which "M and L cannot fail to meet" is replaced by "M and L meet." Rather than calling it a "postulate", Brouwer refers to the "intersection theorem of Euclidean plane geometry", which he states as "a common point can be found for any two lines a and ℓ in the Euclidean plane which can neither coincide nor be parallel." See Fig. 3.
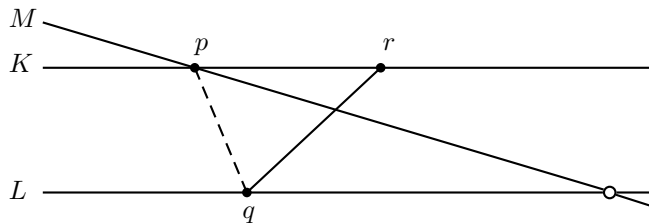
Figure 3: Brouwer's intersection theorem: M and L must meet, provided pq makes alternate interior angles equal with K and L.

This version of the parallel postulate was rediscovered (or re-invented?) in [2, 4], where it was called the "strong parallel postulate." There it was

introduced in an axiomatic setting that included the stability of between-ness, or Markov's principle. As we shall discuss below, this version of the parallel postulate, in the absence of Markov's principle, needs a stronger hypothesis to make constructive sense: he should have required that M and K make a positive angle, i.e., are "positively non-collinear". Without that hypothesis, it is hardly surprising that Brouwer's intersection theorem implies Markov's principle, for the hypothesis that M and K are unequal lines is a negative one, but the conclusion that M meets L is a positive one.

2.4   Euclid 5 formulated in Tarski's language

Euclid's version of the parallel postulate mentions angles, and the concept of "corresponding interior angles" made by a transversal. Here we give a formulation of Euclid's parallel postulate, expressed without mentioning angles. Tarski's well-known axiomatization of geometry only talks about points; angles are discussed indirectly, as triples of points; hence it is of interest to formulate Euclid 5 in that language. See Fig. 4. In order to eliminate the hypothesis about alternate interior angles, we replace it by the hypothesis that the shaded triangles in the figure are congruent. Congruence of triangles just means that the corresponding sides are congruent. This formulation of the axiom then mentions only points and the relations of betweenness and congruence, yet it is conceptually faithful to Euclid's formulation. There are dozens of propositions that are classically equivalent to Euclid 5, and Tarski chose a different one of those to serve as his parallel axiom; but we wish to follow Euclid closely.
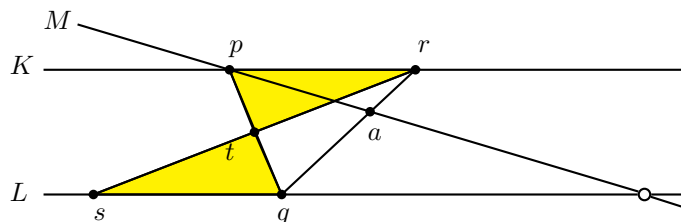
Figure 4: Euclid 5. Transversal pq of lines M and L makes corresponding interior angles less than two right angles, as witnessed by a. The shaded triangles are assumed congruent. Then M meets L as indicated by the open circle.

3   Constructive Euclidean geometry

3.1   Is Euclid constructive?

In constructive mathematics, if one proves something exists, one has to show how to construct it. In Euclid's geometry, the means of construction

are not arbitrary computer programs, but ruler and compass. Therefore it is natural to look for quantifier-free axioms, with function symbols for the basic ruler-and-compass constructions. The terms of such a theory correspond to ruler-and-compass constructions. These constructions should depend continuously on parameters, since we want to allow an interpretation of geometry in which the points are given by successive approximations. We can see the continuity of ruler-and-compass constructions dramatically in computer animations, in which one can select some of the original points and drag them, and the entire construction "follows along." We expect that if one constructively proves that points forming a certain configuration exist, then the construction can be done "uniformly", i.e., by a single construction depending continuously on parameters.

To illustrate what we mean by a uniform construction, we consider an important example. There are two well-known classical constructions for constructing a perpendicular to line L through point p: one of them is called "dropping a perpendicular", and works when p is not on L. The other is called "erecting a perpendicular", and works when p is on L. Classically we may argue by cases and conclude that for every p and L, there exists a perpendicular to L through p. But constructively, we are not allowed to argue by cases. If we want to prove that for every p and L, there exists a perpendicular to L through p, then we must give a single, "uniform" ruler-and-compass construction that works for any p, whether or not p is on L.

The other type of argument (besides argument by cases) that is famously not allowed in constructive mathematics is proof by contradiction. There are some common points of confusion about this restriction. The main thing one is not allowed to do is to prove an existential statement by contradiction. For example, we are not allowed to prove that there exists a perpendicular to L through x by assuming there is none, and reaching a contradiction. From the constructive point of view, that proof of course proves *something*, but that something is weaker than existence. We write it $\neg\neg\,\exists$, and constructively, one cannot delete the two negation signs.

However, when proof by contradiction is used only to prove that two points are equal, or two segments are congruent, or that one point is between two others, the situation seems qualitatively different. There is no "existential information" missing from such a proof. Nothing is being asserted to exist, let alone being asserted to exist without being constructed. The question then arises, whether such proofs are constructively acceptable. That is, whether one is or one is not allowed to prove equalities, congruences, and inequality or order relations between points by contradiction. Consider for the moment two points x and y on a line. If we derive a contradiction from the assumption $x \neq y$, then $x = y$. Formally this becomes the implication

$$\neg\, x \neq y \rightarrow x = y.$$

In the style of Euclid: *Things that are not unequal are equal.* It has its intuitive grounding in the idea that $x = y$ does not make any existential statement. This principle is named "the stability of equality." (Generally any relation is "stable" if it is implied by its double negation.)

8

The principle that $x < y$ can be proved by contradiction can be expressed by

$$\neg\neg\, x < y \rightarrow x < y.$$

or equivalently

$$\neg\, y \leq x \rightarrow x < y.$$

Since order on a line is not a primitive relation in geometry, we consider instead the corresponding axiom for the betweenness relation, namely

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

This is known as "the stability of betweenness." It is also known as "Markov's principle", since it was adopted by Markov as a basic principle of Russian constructive mathematics.

Here is a geometric way of thinking about Markov's principle, rather than as a principle about ordering the continuum. Markov's principle reduces to the assertion that, given two points s and t, we can find two circles with centers s and t that separate the two points. The obvious candidate for the radius is half the distance between s and t, so the question boils down to Markov's principle for numbers: if that radius is not not positive, must it be positive? The stability of betweenness thus has the same philosophical status as Markov's principle: it is self-justifying, not provable from other constructive axioms, and leads to no trouble in constructive mathematics, while simplifying many proofs.

Brouwer was not a person to accept a principle because it was useful, convenient, and apparently harmless, in the sense that it does not interfere with the constructibility of solutions proved to exist with its aid. Before Brouwer would accept a principle, he wanted it to be true, and he saw no reason why Markov's principle has to be true.

3.2 The form of Euclid's theorems and proofs

It is helpful to remember that Euclid did not work in first-order logic. His theorems, and their proofs, have a fairly simple structure: Given some points, lines, and circles bearing certain relations, then there exist some further points bearing certain relations to each other and the original points. This logical simplicity implies that (although this may not be obvious at first consideration) if we allow the stability axioms for equality and betweenness, then essentially the only differences between classical and constructive geometry are the two requirements:

- You may not prove existence statements by contradiction; you must provide a construction.

- The construction you provide must be uniform; that is, it must be proved to work without an argument by cases.

Sometimes, when doing constructive mathematics, one may use a mental picture in which one imagines a point p as having a not-quite-yet-determined location. For example, think of a point p which is very close to line L. We may turn up our microscope and we still can't see whether p is or is not on L. We think "we do not know whether p is on L or not." Our construction of a perpendicular must be visualized to work on such

9

points p. Of course, this is just a mental picture and is not used in actual proofs. It can be thought of as a way of conceptualizing "we do not have an algorithm for determining whether p is on L or not."

We illustrate these principles with a second example. Consider the problem of finding the reflection of point p in line L. Once we know how to construct a perpendicular to L through p, it is still not trivial to find the reflection of p in L. Of course, if p is on L, then it is its own reflection, and if p is not on L, then we can just drop a perpendicular to L, meeting L at the foot f, and extend the segment pf an equal length on the other side of f to get the reflection. But what about the case when we do not know whether p is or is not on L? Of course, that sentence technically makes no sense; but it illustrates the point that we are not allowed to argue by cases. The solution to this problem may not be immediately obvious; see [4].

The description given above of the form of Euclid's theorems is supported by Avigad *et. al.* in [1]:

> Euclidean proofs do little more than introduce objects satisfy-
> ing lists of atomic (or negation atomic) assertions, and then
> draw further atomic (or negation atomic) conclusions from
> these, in a simple linear fashion. There are two minor depar-
> tures from this pattern. Sometimes a Euclidean proof involves
> a case split; for example, if ab and cd are unequal segments,
> then one is longer than the other, and one can argue that a
> desired conclusion follows in either case. The other exception
> is that Euclid sometimes uses a *reductio*; for example, if the
> supposition that ab and cd are unequal yields a contradiction
> then one can conclude that ab and cd are equal.

These arguments are constructively acceptable, if we have the stability of congruence and betweenness. In [4], several examples such arguments in Euclid are examined, including I.6 and I.24. If the conclusion is a congruence or equality statement, even an argument based on "of two unequal segments, one is longer than the other" does not require Markov's principle, but only the stability of congruence.

3.3   Euclidean geometry with Markov's principle

In [4], I advocated adopting the stability of equality, congruence, and betweenness as axioms of constructive geometry. In particular, I argued that Euclid Books I-IV can be formalized using intuitionistic logic plus those two stability principles, using Euclid's version of the parallel postulate instead of Hilbert's, and otherwise making no changes in Hilbert's axioms. In [3], a version of constructive Euclidean geometry is given that is based on Tarski's language and axioms, slightly modified. In addition to using Euclid 5 instead of Tarski's parallel axiom, we found it necessary to use strict betweenness. Hilbert used strict betweenness, but Tarski used non-strict betweenness. To avoid confusion, in this paper we use $\mathbf{B}(a, b, c)$ for strict betweenness and $\mathbf{T}(a, b, c)$ for non-strict betweenness ($\mathbf{T}$ being the first initial of "Tarski"). For details of the axioms see [3]; the essential idea is that after replacing $\mathbf{T}$ by $\mathbf{B}$, we have to "put back" some axioms

that Tarski originally used, but later found clever derivations of from the remaining axioms, using in an essential way the "degenerate cases" of those axioms where non-strict betweenness holds.

In these two long papers, I showed that Euclid 5 together with the stability of congruence and betweenness allows the formalization, not only of Euclid Books I–IV, but also of the construction of coordinates, so that one can construct within geometry the field operations defined on a fixed line L (taken as the x-axis).

The consistency of Euclidean geometry with stability axioms, at least relative to classical Euclidean geometry, is obvious, since all the axioms are classically valid. It may be more instructive to note that Euclidean geometry also has a model in the recursive reals, i.e. those real numbers given by recursive Cauchy sequences with a specified rate of convergence. This model can be formalized in arithmetic, and the resulting interpretation is, using standard recursive realizability, consistent with Markov's principle and Church's thesis for arithmetic. Thus, also in the sense of recursive mathematics, there is nothing inconsistent about the stability of equality and betweenness.

Julien Narboux pointed out that the stability of equality can be derived from the stability of congruence. The proof is given, in the context of Tarski's theories, in Lemma 7.1 of [4]. The converse is also easy to prove. So there are really only two stability axioms to consider: stability of congruence and stability of betweenness.

3.4 Angles and angle ordering

As is well-known, betweenness is never explicitly mentioned in Euclid. There are three ways that betweenness occurs implicitly in the statements of Euclid's propositions: collinearity, ordering of segments, and ordering of angles. Euclid takes these concepts as undefined, and assumes (in the common notions) the basic properties of ordering. Hilbert also took angles and congruence of angles as primitive notions, but unlike Euclid, he defined angle ordering. Tarski defined all three notions. The basic properties of angle ordering then become theorems. These developments are spelled out in [26], Chapter 11, with attention to constructivity in [3], §8.11. We here review that treatment to see whether and where Markov's principle was used.

The concept "x lies on $Ray(b, a)$" is needed to define angles. In [3] §8.11, we defined "x lies on $Ray(b, a)$" by $\neg(\neg\mathbf{T}(b, x, a) \wedge \neg\mathbf{T}(b, a, x))$. Here $\mathbf{T}$ is non-strict betweenness. That definition will not do if we do not have Markov's principle. Instead, we should use the definition

$$\exists e\,(\mathbf{B}(e, b, x) \wedge \mathbf{B}(e, b, a) \wedge eb = ea).$$

That is, x is on the opposite side of b from the reflection e of a in b. In particular, the reflection e must exist for x to be on the ray. Of course, if we assume Markov's principle, then the two definitions are equivalent.

Then "abc and ABC are the same angle" means that the same points lie on $Ray(b, a)$ as on $Ray(B, A)$ and the same points lie on $Ray(b, c)$ as on $Ray(B, C)$. Then two angles abc and ABC are congruent if by changing

a, c, A, and C to other points on those same rays, we can make $ab = AB$ and $bc = BC$ and $ac = AC$.

We say "f lies in the interior of angle abc if there is a "crossbar" uv, with u on $Ray(b,a)$ and v on $Ray(b,c)$, with u, b, and v distinct, and for some point e we have $\mathbf{B}(u,e,v)$ and $\mathbf{B}(b,e,f)$. Then if f lies in the interior of abc, it also lies in the interior of any $a'bc'$ that is the "same angle" as abc.

4 Brouwer's 1949 paper

4.1 What Brouwer actually proved

What did Brouwer mean by asserting the inconsistency of Euclidean geometry? In this section we answer that question. Brouwer worked extensively with "the continuum", which we denote here by $\mathbb{R}$. We refer to members of the continuum as "real numbers", rather than using Brouwer's terminologies for numbers given by certain kinds of sequences. Geometry has a model in which the points are pairs of real numbers (x,y). We refer to this model as "the model $\mathbb{R}^2$."

Real numbers are given by sequences of rationals, and order in $\mathbb{R}$ is defined in terms of order in the rationals and the concept of sequence, so it implicitly depends on the natural numbers. Order in geometry, on the other hand, is defined in terms of betweenness on a line, which in turn is given by some axioms about betweenness. In Brouwer's 1949 paper, he is concerned only with the model $\mathbb{R}^2$ of geometry, and not with axiomatic geometry.

Here is what Brouwer proved in [7]:

Theorem 1 (Brouwer) *In the model $\mathbb{R}^2$, the statement "the strong parallel postulate implies Markov's principle" holds.*

Proof. Let L be the x-axis and P be the point $(0,1)$. Assume $\neg\neg \epsilon > 0$. Let $a = (1, 1 - \epsilon)$. According to the strong parallel principle, the line through P and a meets L. The point of intersection is $(x,0)$, where by similar triangles x is to 1 as 1 is to ϵ. (The existence of x is guaranteed by the strong parallel postulate, not by division by ϵ, which would not be legal under the weak assumption $\epsilon \neq 0$.) Now Brouwer appeals to the fact that $\mathbb{R}$ satisfies the Archimedean axiom: there is a positive integer N, which Brouwer chooses in the form 2^{n_ℓ}, greater than $|x|$. Now we have $x < N$, and both sides of the inequality have multiplicative inverses. Brouwer says (line 12 of his paper, where his ρ_ℓ is our $-\epsilon$)

$$\rho_\ell < -2^{-n_\ell}$$

or in our notation,

$$\epsilon > \frac{1}{N}.$$

Brouwer offers no further justification, but we think some is necessary. Now $\epsilon = 1/x$, so to justify Brouwer's conclusion we would need to infer $1/x > 1/N$ from $x < N$ and the existence of $1/x$ and $1/N$. We know

$N > 0$, but we only know $\neg\neg \; x > 0$; hence we do not know $xN > 0$ and the last step of the following argument cannot be carried out

$$\frac{1}{x} - \frac{1}{N} = \frac{N - x}{xN} > 0$$

since the last step would require $xN > 0$, which we do not have. Instead, we get only

$$\neg\neg \; \epsilon > \frac{1}{N}.$$

We can, however, justify Brouwer's argument using apartness. The principle of apartness, which holds in $\mathbb{R}^2$, says that if $\alpha < \beta$, then

$$z < \beta \vee \alpha < z.$$

It follows that if $\alpha > 0$ and $\neg\neg \; \beta > 0$, then $\alpha + \beta > 0$, since by apartness, $\alpha + \beta > 0 \vee \alpha + \beta < \alpha$, and the second case, $\alpha + \beta < \alpha$, implies $\beta < 0$, which contradicts $\neg\neg \; \beta > 0$. Applying this principle with $\beta = 1/x - 1/N$ and $\alpha = 1/2N$, we have

$$\frac{1}{x} - \frac{1}{N} + \frac{1}{2N} = \frac{1}{x} - \frac{1}{2N} > 0$$

and hence $\epsilon > 1/2N > 0$ as desired. That completes the proof.

Brouwer used the Archimedean axiom to get a positive upper bound on x. That is not necessary, and neither is the use of apartness. We now show how to remove these two principles from Brouwer's proof. First, we observe that from the hypothesis $\neg\neg \; \epsilon > 0$ and the equation $x = 1/\epsilon$, we have $\neg\neg \; x > 0$. From $\neg\neg \; x > 0$ we have $\neg\neg \; x = |x|$. By the stability of equality, we have $x = |x|$.

Now, instead of using Archimedes's axiom to get a positive upper bound on x, we could just as well use $|x| + 1$. Then we have

$$0 < \frac{1}{2(|x| + 1)} < \frac{1}{|x|} = \frac{1}{x} = \epsilon$$

and hence $\epsilon > 0$. The use of Archimedes's axiom is a red herring (for this proof—not necessarily in general for intuitionism). That Brouwer used it anyway shows how far from Brouwer's mind was any consideration of whether his argument was first-order or not.

4.2 Implications of Brouwer's theorem for axiomatic geometry

Brouwer's theorem is formulated as a theorem about the Euclidean plane $\mathbb{R}^2$. But for half a century already at the time of Brouwer's publication, ever since Hilbert's 1899 book [15], geometry had moved on from discussing the one true plane to axiomatic formulations. In Brouwer's paper, he did not consider the question whether Markov's principle is contradictory in some axiomatic system for Euclidean geometry. That may have been because of his aversion to axiomatics in general, or it may have been the lack of any development at all of constructive Euclidean geometry at

that time, or it may have been for some other reason entirely. But now, the question seems natural.[3] Can we use Brouwer's proof to show that the axioms of Euclidean geometry (as formulated for example in [3], but without Markov's principle) allow one to deduce Markov's principle from Brouwer's version of the parallel axiom?

We showed above that Brouwer's use of the Archimedean axiom and his (implicit) use of apartness are easily eliminated. The one remaining issue is his use of coordinate geometry. What Brouwer's proof shows is that, if some theory of Euclidean geometry suffices to define coordinates and arithmetic, then in that theory, Brouwer's version of the parallel postulate implies Markov's principle. That is, after all, not too surprising, given that the hypothesis of Brouwer's "intersection theorem" is the negative statement that the two lines M and K are not identical, but the conclusion is a positive existence statement. Markov's principle is "built-in."

Later in this paper, we will give a formal theory EG for Euclidean geometry without Markov's principle. The theory EG plus Markov's principle has been extensively studied in [3, 4]. The fact that an attractive theory of geometry can be formulated without using Markov's principle as an axiom, and without any obvious way to prove Markov's principle, leads to the following questions and conclusions:

(1) Does EG (which does not have MP) plus Euclid 5 suffice to define coordinates and geometric arithmetic? For short we say "EG can define arithmetic" to describe this property. We claim in this paper that EG can define arithmetic.

(2) Then Brouwer's proof shows that EG plus Brouwer's intersection theorem proves MP.

(3) We know [4] that EG + MP proves Brouwer's intersection theorem, which is there called the "strong parallel postulate", or SPP for short. Hence SPP is actually equivalent to MP in EG.

HA is "Heyting's arithmetic", the standard formal theory of arithmetic with intuitionistic logic. In the context of HA, "Markov's principle" is the name usually given to

$$\neg\neg \exists x\, P \to \exists x\, P \qquad (P \text{ primitive recursive}).$$

It is well-known (see e.g. [29]) that this principle is not provable in HA. Consider the interpretation of EG in HA determined by representing points as pairs of (indices of) recursive real numbers. To verify that the interpretations of the axioms of EG are provable in HA, we do not need Markov's principle, since the stability of betweenness is not an axiom of EG. We also do not need Markov's principle to verify any of the other axioms of EG, including Euclid 5, since the coordinates of the point asserted to exist by Euclid 5 involve only positive denominators. To show

---

[3]The referee pointed out that Brouwer was the thesis advisor, two years after the publication of the article on the contradictority of elementary geometry, of Johanna Adriana Geldof's thesis [9]. This thesis has nothing specifically intuitionistic in it, except one sentence near the beginning saying that it assumes any two elements are either equal or not equal, and also nothing specifically Euclidean, but it does show that Brouwer's aversion to axiomatics was not absolute.

that EG does not prove the geometric Markov's principle (stability of betweenness) it therefore suffices to prove that the recursive interpretation of the geometric Markov's principle is equivalent to the arithmetic Markov's principle. But that is an easy exercise. Hence

(4) EG does not prove Markov's principle.

Since Brouwer's "intersection theorem" is equivalent to Markov's principle, EG does not prove that "theorem." In particular Brouwer's "intersection theorem" is not verifiable in the recursive interpretation (unless we assume the arithmetic Markov principle).

Brouwer never mentioned (in the paper under consideration, or anywhere else as far as I know) Euclid, or Euclid's axioms; nor did he mention Hilbert, or Hilbert's axioms, or any axiomatic system whatever. He worked simply with the plane $\mathbb{R}^2$ using coordinate geometry. He showed that if $\mathbb{R}^2$ satisfies the "intersection theorem", then $\mathbb{R}$ satisfies Markov's principle. The title of his paper, however, claims that "elementary geometry is contradictory." To reach that conclusion from his result, one would need to believe that the intersection theorem is part of elementary geometry, and that Markov's principle is contradictory. Of course the intersection theorem is a part of *classical* elementary geometry, but its proof requires Markov's principle, so it is not a part of *intuitionistic* elementary geometry, unless one assumes Markov's principle or an axiom that implies Markov's principle. In particular, the intersection theorem in question does not follow from Euclid's formulation of the parallel postulate.

4.3   Why did Brouwer reject Markov's principle?

Brouwer had claimed in that same year (1949) in [8] that Markov's principle is contradictory. His proof used real numbers that are limits of sequences generated by the "creative subject", who is allowed to examine at each stage of mathematical construction, all proofs developed at earlier stages. Brouwer regarded this as an improvement over a paper published the previous year [6], in which he showed that Markov's principle was "unlikely to be provable." In view of these results, Brouwer viewed his version of the parallel axiom as "contradictory."

Here is a sketch of Brouwer's refutation of MP, in more modern terms. He used Kripke's schema (KS), according to which any proposition ϕ is equivalent to a proposition of the form $\alpha > 0$, for some real number α. Taking ϕ to be $A \vee \neg A$, and noting that $\neg\neg (A \vee \neg A)$ is intuitionistically valid, we have $\neg\neg \alpha > 0$, so by Markov's principle $\alpha > 0$; that is, $A \vee \neg A$. Hence KS implies the law of the excluded middle. But the fan theorem (uniform continuity of functions on 2^N) refutes the law of the excluded middle. Hence KS plus the fan theorem refutes MP. Brouwer believed, at the time of writing the papers we are discussing, that his theory of the creative subject justified KS, and hence, that MP had been refuted.

Krike's schema has found few (if any) adherents in the 70 years since Brouwer advocated it. On the other hand, Markov's principle is consistent with most commonly-studied intuitionistic theories; in particular with the theories of intuitionistic analysis including Brouwer's fan theorem and bar

theorem as axioms. See [28] and [18] for proofs of this consistency using variations of recursive realizability.

4.4 Two sides

In Brouwer's paper [7], he also considered an ordering principle that we call "two-sides":

$$x \neq 0 \rightarrow x < 0 \vee x > 0.$$

The name is chosen because the principle can be thought of as saying that there are "two sides" of the y-axis: every point not on the y-axis lies on the left half of the plane or on the right half.

We digress to show that two-sides can be expressed in the language of geometry. Fix two points 0 and 1, and define -1 to be the endpoint of the extension of the line segment from 1 to 0 by itself. Then $x < 0$ can be defined as $\mathbf{B}(x, 0, 1)$. Two-sides can be expressed as

$$x \neq 0 \rightarrow \mathbf{B}(x, 0, 1) \vee \mathbf{B}(-1, 0, x).$$

Brouwer rejected not only Markov's principle, but also two-sides. Two-sides is not a theorem of EG, even with the help of Markov's principle, as shown in [3]. Proof sketch: the axioms of EG can be expressed, after introducing some function symbols, without using $\exists$ or $\vee$. Then cut-elimination can be used to show that no disjunctive theorems can be proved (unless one of the disjuncts can be proved.) Two-sides does imply Markov's principle, since if we assume $\neg\neg\ x > 0$, then $x \neq 0$, so by two-sides, $x > 0 \vee x < 0$; but $x < 0$ contradicts $\neg\neg\ x > 0$, so that case is ruled out, and we conclude $x > 0$. Thus two-sides is stronger than Markov's principle.

5 Axioms of non-Markovian geometry

Brouwer's objection to Markov's principle led us to consider whether Markov's principle is really necessary for Euclidean geometry. In this paper, we introduce a theory EG of Euclidean geometry, with the stability of congruence (and hence the stability of equality) but without the stability of betweenness (which is also called Markov's principle). EG has Euclid 5 for its parallel postulate, so it corresponds closely to Euclid. The axioms of EG are listed for reference in §5.11, but we shall introduce them gradually, with explanations.

It probably does not matter whether we take a Hilbert-type formulation or a Tarski-type formulation as in [3], but for the sake of precision we must pick a specific list of axioms, and it is far simpler to work with the simple language and short list of axioms of Tarski's theory. As a starting point, we consider the constructive version of Tarski's axioms given in [3], with Markov's principle deleted from the list of axioms. We also need to modify two other axioms (segment extension and Pasch) to ensure that they do not immediately imply Markov's principle. We postpone the details of those modifications of the axioms to the next two sections.

EG has line–circle continuity: a line through a point (strictly) inside a circle meets the circle in two distinct points. We also include circle–circle continuity as an axiom.[4]

What we seek, then, is an axiomatization of EG such that each axiom is equivalent (using MP) to an axiom of intuitionistic Tarski geometry (as defined in [3]), such that EG does not imply Markov's principle, and yet suffices for the development of Euclidean geometry. More specifically, we want EG to satisfy these criteria:

- EG suffices to prove the correctness of the uniform constructions given in [3], namely, uniform perpendicular and reflection in a point. This permits the assignment of coordinates (x, y) to each point, given two fixed perpendicular lines to serve as the x-axis and y-axis.

- EG suffices to prove that every pair (x, y) occurs as the coordinates of some (unique) point.

- EG suffices to define the (uniform) addition and multiplication of (signed) points on the x-axis, and the construction of square roots of non-negative points.

- EG suffices to formalize the arguments of Euclid Book I, and probably Books II–IV as well.

The papers [3] and [4] established these facts for a theory including Markov's principle, so the task here is to find a modified version of this theory that does not imply Markov's principle but still satisfies the criteria listed above.

In particular, we give EG the parallel axiom Euclid 5 (just as in [4, 3] rather than the "strong parallel axiom" used by Brouwer. It follows that, in some sense, Brouwer was criticizing a "straw man", in that the parallel postulate that he found unsatisfactory is not actually Euclid's parallel postulate, and Euclid 5 does not suffer from the flaw (if it is a flaw) that Brouwer pointed out. The strong parallel postulate and Euclid 5 are equivalent in Euclidean geometry with Markov's principle (as shown in [4]), but they are *not* equivalent if Markov's principle is dropped, since (at least with the aid of the apartness axioms) the strong parallel postulate implies Markov's principle, while Euclid 5 does not.

There is no philosophical advantage (for the present purposes) in trying to choose a minimal set of axioms for EG, and indeed there are reasons (discussed below) to be generous in taking more axioms than probably are necessary. We assume two versions of Pasch, and both line–circle and circle–circle continuity. We modify Pasch's axioms to avoid degeneracies that imply discontinuous dependence (as we did in [3]) and also to avoid near-degeneracies that imply Markov's principle. The remarkable conclusion is that we can then derive Euclid Book I (and probably II-IV), coordinates, and arithmetic, without Markov's principle, though we do need to add hypotheses that angles are positive and vertices are distinct.

---

[4]For our present purposes, there is little to be gained by trying to eliminate circle–circle continuity as an axiom, or in general, by trying to minimize the number of axioms, since our aim is simply to demonstrate the viability of intuitionistic geometry without Markov's principle.

5.1 Markov's principle and betweenness

Tarski's geometry, and its variant EG, include a minimal set of axioms about betweenness. Remember that we use strict betweenness $\mathbf{B}$ rather than non-strict betweenness $\mathbf{T}$ as in Tarski. If we have Markov's principle, then $\mathbf{B}$ and $\mathbf{T}$ are interdefinable. But without Markov's principle, there is no apparent way to define $\mathbf{B}$ from $\mathbf{T}$, so it is good that we took $\mathbf{B}$ as fundamental.

Tarski's final theory [27] had only one betweenness axiom, known as (A6) or "the identity axiom for betweenness":

$$\mathbf{T}(a,b,a) \to a = b.$$

In terms of strict betweenness, that becomes $\neg\mathbf{B}(a,x,a)$, or otherwise expressed, $\mathbf{B}(a,b,c) \to a \neq c$. We also refer to this axiom as (A6). The original version of Tarski's theory had more betweenness axioms (see [27], p. 188). These were all shown eventually to be superfluous in classical Tarski geometry, through the work of Eva Kallin, Scott Taylor, Tarski himself, and especially Tarski's student H. N. Gupta [11]. These proofs appear in [26]. Here we give the axiom numbers from [27], names by which they are known, and also the theorem numbers of their proofs in [26]:

$\mathbf{T}(a,b,c) \to \mathbf{T}(c,b,a)$	(A14), symmetry, Satz 3.2
$\mathbf{T}(a,b,d) \wedge \mathbf{T}(b,c,d) \to \mathbf{T}(a,b,c)$	(A15), inner transitivity, Satz 3.5a
$\mathbf{T}(a,b,c) \wedge \mathbf{T}(b,c,d) \wedge b \neq c \to$	
$\qquad \mathbf{T}(a,b,d)$	(A16), outer transitivity, Satz 3.7b
$\mathbf{T}(a,b,d) \wedge \mathbf{T}(a,c,d) \to$	
$\qquad \mathbf{T}(a,b,c) \vee \mathbf{T}(a,c,b)$	(A17), inner connectivity, Satz 5.3
$\mathbf{T}(a,b,c) \wedge \mathbf{T}(a,b,d) \wedge a \neq b \to$	
$\qquad \mathbf{T}(a,c,d) \vee \mathbf{T}(a,d,c)$	(A18), outer connectivity, Satz 5.1

Our theory of constructive geometry in [3]) has three betweenness axioms (numbered as in [27], with an "i" added for "intuitionistic"):

$\neg\mathbf{B}(a,b,a)$	(A6-i)
$\mathbf{B}(a,b,c) \to \mathbf{B}(c,b,a)$	(A14-i), symmetry of betweenness
$\mathbf{B}(a,b,d) \wedge \mathbf{B}(b,c,d) \to \mathbf{B}(a,b,c)$	(A15-i), inner transitivity

In [3], we appealed to Gödel's double-negation interpretation to "import" the long, complicated proofs of the classically superfluous axioms A17 and A18. (Of course the conclusion has to be double negated to eliminate the disjunction.) Since $\mathbf{T}$ is defined negatively, that still works even without Markov's principle, for the versions of these axioms using $\mathbf{T}$. Versions of A17 and A18 in which the disjunction is not double-negated of course are not constructively valid, and if we do double-negate the disjunction, then the double negation interpretation still works with $\mathbf{T}$ replaced by $\mathbf{B}$.

The following "connectivity principle" is equivalent to A17, and it is what is actually needed for several very basic lemmas, such as "two lines intersect in at most one point."

If B and C are both between A and D, and neither is between A and the other, then they are equal.

By applying the double negation interpretation as described above to the complicated proof of A17 in [26]. Therefore, theoretically, there is no need to add this formula as an axiom. However, few readers are going to check the proof in [26] and verify the double negation interpretation, so we might as well just ask them to accept the connectivity principle as an axiom. We therefore include it as an axiom of EG.

The question remains, did we really add enough betweenness axioms? The answer is, yes we did, because we can successfully define coordinates and arithmetic, and with that geometrically defined arithmetic, the points on a line form a Euclidean field. The axioms for Euclidean fields (without assuming Markov's principle) are discussed in §7.2.

5.2 Distinct points and segment extension

Tarski's segment extension axiom provides for an extension of segment ab by segment cd; the result is a point, sometimes written $ext(a,b,c,d)$, about which the axiom asserts, if $p = ext(a,b,c,d)$, that $bp = cd$ and $\mathbf{B}(a,b,p)$. Tarski's theory has no condition on a and b, but in constructive geometry (with Markov's principle), as discussed in [3], we have to require $a \neq b$; that is, only non-null segments can be extended.

We will show next that the (unmodified) extension axiom implies Markov's principle. In preparation we review two facts, the uniqueness of extension and the stability of equality. The extension of segment ab by segment pq is unique: if there are two points c such that $\mathbf{B}(a,b,c)$ and $bc = pq$ then those two points are equal, as can be proved using other axioms of geometry. The principle of "stability of equality" was introduced in § 3.1: $\neg\neg\, x = y \rightarrow x = y$.

Now we prove that the extension axiom implies Markov's principle. Suppose $\neg\neg\, B(a,b,c)$; then $a \neq b$, so $p = ext(a,b,b,c)$ is defined; but $bc = bp$ and $\mathbf{B}(a,b,p)$ by the extension axiom. Using the uniqueness of extension (doubly negated), from $\neg\neg\, B(a,b,c)$ we can prove $\neg\neg\, p = c$. Then, by the stability of equality, $p = c$. Then from $\mathbf{B}(a,b,p)$ we have $\mathbf{B}(a,b,c)$, the conclusion of Markov's principle.

Therefore, if one wishes to do geometry without Markov's principle, one must modify the extension axiom. We do so by restricting the applicability of segment extension to positive segments. That concept is defined as follows: We say segment ab is *positive*, or equivalently, that a and b are *distinct*, if

$$\exists e\, \mathbf{B}(e,a,b) \vee \exists e\, \mathbf{B}(a,b,e) \vee \exists e\, \mathbf{B}(a,e,b).$$

We write this as $a\#b$. Since the word "length" carries the idea of measurement by numbers, we avoid the phrase "has positive length", but it may be helpful to mention it once.

Once $a\#b$ is defined, the segment extension axiom has additional subtleties. We said above that we want the axiom to say that positive segments can be extended. But we also want the notion "positive segment" to be closed under congruence; that is,

$$ab = cd \wedge a\#b \rightarrow c\#d.$$

We wish this principle to be a theorem. We therefore need to have our segment extension axiom say, "any segment congruent to a positive segment can be extended," rather than just "any positive segment can be extended."

A second subtlety arises from the distinction between **T** and **B**. In constructive geometry with Markov's principle we could just say

$$a \neq b \rightarrow \mathbf{T}(a, b, ext(a, b, c, d))$$

and then it can be proved that

$$a \neq b \wedge c \neq d \rightarrow \mathbf{B}(a, b, ext(a, b, c, d)).$$

But in non-Markovian geometry, we need to change $\neq$ to $\#$ and take *both* axioms. The following show what we mean, but they still deal only with the second subtlety and not the first.

$$a\#b \rightarrow \exists e \, (\mathbf{T}(a, b, e) \wedge be = cd).$$
$$a\#b \wedge c\#d \rightarrow \exists e \, (\mathbf{B}(a, b, e) \wedge be = cd).$$

Of course, in the presence of Markov's principle, these are equivalent to the form (A4-i) used in [3].

Combining the two solutions, we now give the final form of the extension axiom of EG:

$$A\#B \wedge ab = AB \rightarrow \exists e \, (\mathbf{T}(a, b, e) \wedge be = cd).$$
$$A\#B \wedge ab = AB \wedge C \neq D \wedge cd = CD \rightarrow \exists e \, (\mathbf{B}(a, b, e) \wedge be = cd).$$

Lemma 1 *If $a\#b$ and $ab = cd$ then $c\#d$.*

Proof. Suppose $a\#b$ and $ab = cd$. There are three cases. Case 1, for some e we have $\mathbf{B}(a, b, e)$. Then $b\#e$, so we can extend cd by be to a point f with $\mathbf{B}(c, d, f)$. Hence $c\#d$. Case 2, for some e we have $\mathbf{B}(e, a, b)$, is treated similarly. Case 3, for some e we have $\mathbf{B}(a, e, b)$. Then first extend ab by ab to e; so $\mathbf{B}(a, b, e)$, reducing to case 1. That completes the proof.

The symbol $a\#b$ is traditionally used for apartness, a concept introduced by Heyting. We use the word "distinct" instead of "apart", because we do not include the traditional apartness axiom $a\#b \rightarrow a\#c \vee b\#c$. The reasons why not are discussed in [4]. We use "unequal" for the negative relation $a \neq b$, and "distinct" or its synonym "different" for $a\#b$.

Remark. One might think that one could adopt one of the three betweenness conditions as the definition of $a\#b$ and prove the other two to be equivalent. That can be done using some lemmas about betweenness, for example, to prove the third property, we would need outer transitivity. But the proof of outer transitivity begins by extending a segment whose endpoints cannot be seen to be distinct except by the third property itself; so that approach is circular and does not work. Therefore, we adopt the definition given instead of a shorter one.

It is often said that "two points determine a line." In intuitionistic geometry, the correct statement is "two *distinct* points determine a line."

A line can be thought of as all the possible extensions of a segment. Intuitively, if we do not know that two points are distinct, the direction of the segment connecting them is not clear and a line is not (yet) determined.

The following fundamental principle of order on a line would have to be taken as an axiom, if it were not provable. However, in non-Markovian geometry, we note that the hypothesis requires b and c to be distinct, not just unequal.

Lemma 2 (Outer transitivity)

$$\mathbf{B}(a,b,c) \wedge \mathbf{B}(b,c,d) \wedge b\#c \to \mathbf{B}(a,b,d).$$

and also

$$\mathbf{B}(a,b,c) \wedge \mathbf{B}(b,c,d) \wedge b\#c \to \mathbf{B}(a,c,d).$$

Proof. The first form follows from the second, using the symmetry of betweenness. We prove the second form. Suppose the hypothesis. Since $\mathbf{B}(a,b,c)$, we have $a\#c$. Hence segment ac can be extended by cd to point e. Then by the 5-segment axiom, applied with d at the top in Fig. 5, and acd and ace along the bases of the two figures, we conclude $ed = dd$. But $ed = dd$ implies $e = d$, and by definition of e and the extension axiom A1-i, we have $\mathbf{B}(a,c,e)$, since $c\#d$ by hypothesis. Hence $\mathbf{B}(a,c,d)$ as desired. That completes the proof of the lemma.

5.3 The five-segment axiom and SAS

Tarski replaced Hilbert's fourth and fifth congruence axioms (angle transport and SAS) with an elegant axiom, known as the five-segment axiom. This axiom is best explained not through its formal statement, but through Fig. 5. The 5-segment axiom says that in Fig. 5, the length of the dashed segment cd is determined by the lengths of the other four segments in the left-hand triangle. Formally, if the four solid segments in the first triangle are pairwise congruent to the corresponding segments in the second triangle, then the dashed segments are also congruent.
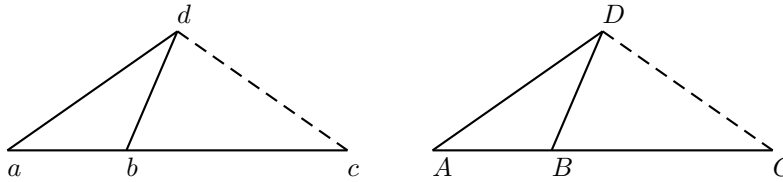
Figure 5: The 5-segment axiom. If $a\#b$ and the corresponding solid segments are congruent, then $cd = CD$.

The 5-segment axiom is a thinly-disguised variant of the SAS criterion for triangle congruence. To see this, refer to the figure. The triangles we are to prove congruent are dbc and DBC. We are given that bc is congruent to BC and db is congruent to DB. The congruence of angles dbc and DBC is expressed in the 5-segment axiom by the congruence of

triangles abd and ABD, whose sides are pairwise equal. The conclusion, that cd is congruent to CD, gives the congruence of triangles dbc and DBC. In Chapter 11 of [26], one can find a formal proof of the SAS criterion from the 5-segment axiom. It is easily adapted to prove the SAS criterion in non-Markovian geometry (which in non-Markovian geometry requires that the angle in question be positive).

The formal statement of the axiom uses non-strict betweenness and does not specify that d is not collinear with ab, though it does require $a \neq b$. Allowing d to be collinear with ab permits the use of this axiom to derive properties of betweenness, and it is not only constructively valid, but very useful.

The question arises whether we ought to strengthen the hypothesis $a \neq b$ to $a\#b$, or even perhaps require $\mathbf{B}(a, b, c)$ (which would imply $a\#b$, but also $b\#c$). The latter would weaken the axiom.[5] The former seems philosophically prudent, since without it, perhaps the direction of the line ab is uncertain. We therefore require $a\#b$ in this axiom.[6]

This axiom enables one to replace reasoning about angles with reasoning about congruence of segments.[7] We would like to emphasize that the 5-segment axiom is often just as easy to use as SAS. Here is an illustrative example:

Lemma 3 *Vertical angles are congruent.*

Proof. Let angles abc and dbe be vertical angles; so $\mathbf{B}(a, b, e)$ and $\mathbf{B}(c, b, d)$. We may suppose without loss of generality that $bd = bc = ab = be$. We must show $ac = de$. Consider the two configurations $abec$ and $dbce$. The hypotheses of the 5-segment axiom are fulfilled, because $ab = db$, $be = bc$, $ec = ce$, and $bc = be$. Then by the 5-segment axiom, $ac = de$ as desired. That completes the proof. The reader is urged to compare this proof to Euclid's proof.

5.4   Lines, rays, triangles, and right angles

The segment extension axiom allows us to create a new point on the line containing a and b only when a and b are distinct points. Without Markov's principle, we have to distinguish between pairs of unequal points, and pairs of distinct points. In some sense only distinct points determine a segment, because only then does the segment extension axiom apply. Although Tarski's language speaks only about points, the import of the axiom is that we are allowed to construct the line containing a and b only

---

[5] The theorem "all null segments are equal" would not be provable; but adding it back as a new axiom would fix that. However, there seems no point in doing that.

[6] The unmodified 5-segment axiom holds in our Kripke model where Markov's principle fails, so we are not mathematically compelled to modify it.

[7] The history of this axiom is as follows. The key idea (replacing reasoning about angles by reasoning about congruence of segments) was introduced (in 1904) by J. Mollerup [24]. His system has an axiom closely related to the 5-segment axiom, and easily proved equivalent. Tarski's version, however, is slightly simpler in formulation. Mollerup (without comment) gives a reference to Veronese [33]. Veronese does have a theorem (on page 241) with the same diagram as the 5-segment axiom, and closely related, but he does not suggest an axiom related to this diagram.

when $a\#b$. Instead of "two unequal points determine a line", we have "two distinct points determine a line".

The relation of collinearity is defined by

$$L(a,b,c) \;\leftrightarrow\; \neg\neg\,(\mathbf{B}(c,a,b) \vee c = a \vee \mathbf{B}(a,c,b) \vee c = b \vee \mathbf{B}(a,b,c)).$$

Pushing one negation sign inwards, $L(a,b,c)$ can be expressed in a negative form. Similarly, the notion that x is on the ray from b through c can be expressed by $L(b,c,x) \wedge \neg\mathbf{B}(x,b,c)$. We write this as $Ray(b,c,x)$. The ray itself is not treated as a mathematical object; only the relation between three points is formalized. Nevertheless we speak of x lying on $Ray(b,c)$, which is the same as saying $Ray(b,c,x)$.

Lemma 4 *Let a and b be distinct points. Let c and d be any points. Then we can lay segment cd off on $Ray(a,b)$, in the sense that there exists a point x with $Ray(a,b,x)$ and $ax = cd$. Moreover, point x is unique.*

Proof. Since $a\#b$, segment ba can be extended to point e with $ae = ab$ and $\mathbf{B}(e,a,b)$. (We say e is the reflection of b in a.) Then segment ea can be extended by cd to point x. Then $ax = cd$ and $\mathbf{T}(e,a,x)$, by the segment extension axiom. But that is the meaning of $Ray(a,b,x)$, so the first claim of the lemma is proved. It remains to prove the uniqueness of x. Suppose y is another point on $Ray(a,b)$ with $ay = cd$. By the definition of $Ray(a,b,x)$, that means for some point e (the reflection of b in a) we have $\mathbf{T}(e,a,x) \wedge \mathbf{T}(e,a,y) \wedge ax = ay$. Now apply the 5-segment axiom to the following two configurations in the form of Fig. 5: the first has eax on the base and y on the top, the second has eay on the base and also has y on the top. The hypotheses are fulfilled because $ax = ay$. The conclusion is $yx = yy$. By Axiom A3, $x = y$. That completes the proof.

Definition 1 *Angle abc is a right angle if a, b, and c are distinct points, and if $d = ext(a,b,a,b)$ (so d is the reflection of a in b), then triangles abc and dbc are congruent triangles.*[8]

It can be proved that this notion respects the congruence relation on angles. Euclid's Postulate 4, that all right angles are congruent, is a theorem in Tarski's geometry.[9] It can be proved (and without Markov's principle). We do not have space in this paper to give the proof, but we will outline it. The details can be found in [26], Part I, Chapter 10.

Reflection in a point is an isometry. That is, if B is the midpoint of AC and also of PQ, then $AP = CQ$. The relation "ABC is a right angle" is preserved under reflection in a point. Reflection in a line is an isometry. And the key result (Satz 10.12 in [26]): Given right angles ABC and ABF with BC congruent to BF, then AC is congruent to AF. (Remember there is no dimension axiom, so picture the two angles in different planes.) Here is a sketch of the proof:

---

[8]We note that in [26], the predicate $R(a,b,c)$ allows the case $b = c$, but our definition of "right angle" requires three distinct points.

[9]It is a much *simpler* theorem in Hilbert's geometry, because Hilbert takes as an axiom that an angle can be copied on a given side of a given line, and the copy is unique. These facts require non-trivial proofs from the axioms of EG or Tarski.

Let M be the midpoint of CF (which exists without needing circles since triangle BCF is isosceles). Let D be the reflection of F in B, so ABD is congruent to ABF. Here is the key: triangle ABD is the reflection of ABC in the line BM. Since reflection in a line is a isometry, those triangles are congruent, proving Satz 10.12. Now given two right angles, using Euclid's Prop. 23 we can copy one of them into the position of the second angle in Satz 10.12, and applying Satz 10.12, the angles are congruent.

Lemma 5 *If abc is a right angle, then cba is a right angle.*

Proof. The two angles obtained by reflecting abc first in ba and then in bc are vertical angles, hence congruent. In fact the proof of Lemma 3 works directly, without needing to first prove that an angle congruent to a right angle is a right angle.

5.5   Positive angles

Corresponding to the notion of "distinct points", there is, intuitively, a notion of "positive angle", which we write $abc > 0$ (it is not necessary to write $\angle abc > 0$). We also write this as $0 < abc$; both our abbreviations for a statement involving betweenness (given below). Saying $abc > 0$ is stronger than simply requiring that a, b, and c are not collinear (with a and c on the same side of b) in the same way that apartness is stronger than inequality. It is our immediate aim to define this notion. We wish this notion to be defined in such a way that, after we prove that coordinates can be geometrically defined, the angle between $p = (x, y)$ and $q = (b, 0)$ with vertex at $(0, 0)$ is a positive angle if and only if $y > 0$.[10] But this cannot be the definition, as many theorems must be proved before we can construct coordinates. For example, midpoints and perpendiculars are needed.

We must define "positive angle" at the outset, because (as we shall see in the next section) we need to restrict the hypotheses of Pasch's axiom by requiring an angle to be positive. Our first use of Pasch's axiom will be to repair Euclid's constructions in Propositions I.10 to I.10, culminating in the theorem that every segment ab with $a\#b$ has a midpoint. In that proof, we need to apply Pasch's axiom in situations where the vertex angle is $30°$, $60°$, and $120°$. Therefore, the definition of "positive angle" must immediately imply that those angles are positive. More generally, we need to know that the angles of a right triangle are positive. (Remember that "triangle" implies the vertices are distinct.)

The following conversations, following the pattern of "light bulb" jokes, illustrate the basic concepts of non-Markovian geometry:

> How many points does it take to determine a line? *Three, two to lay a straightedge on and one to check that the other two are distinct.*

---

[10] Here $y > 0$ refers to order on a line, defined axiomatically by betweenness, namely $\mathbf{B}(-1, 0, y)$, where -1 and 0 are two specified points on a specified line.

> How many points does it take to determine a positive angle?
> *Six, three to determine the vertex and sides, one for each side to check the other two are distinct, and the sixth to check that the sides do not coincide.*

Definition 2 *(i) abc is an* apex angle *if there are distinct points u, v on $Ray(b, a)$ and $Ray(b, c)$ respectively such that $bu = bv$ and $b \# u$.*

(ii) abc is an angle of a right triangle if bac or bca is a right angle, or more generally, if angle abc is congruent to an angle ABC such that BAC or BCA is a right angle. "Right angle" is defined in Definition 1).

(iii) $abc > 0$ ("abc is a positive angle") if abc is an apex angle, or a right angle, or an angle of a right triangle.

We will later prove that every positive angle is an apex angle, but that can be done only after developing some geometry, using the definition above. In other words, if we were to take "apex angle" as the definition of "positive angle", we could not prove that an angle of a right triangle is positive, because to do so we need some intermediate results that cannot be justified until we know that an angle of a right triangle is positive. Specifically, we will see below that a correct formulation of Pasch's axiom for non-Markovian geometry requires certain angles to be positive, and to prove that angles of a right triangle are apex angles, we need such angles to be positive.

The formula $abc > 0$, written out in primitive notation, is existential, since not only is there an explicit $\exists$ and an explicit disjunction, but also distinctness involves an existential quantifier. This is good, since we intend it to express the existence of a positive lower bound on the angle abc.

We also need to express $abc < \pi$, i.e., angle abc is positively different from a "straight angle." We just need to say that the supplement of abc is a positive angle.

Definition 3 *Angle abc has a positive supplement, or $abc < \pi$, is defined by*

$$abc < \pi \leftrightarrow \exists d \left(\mathbf{B}(a, b, d) \wedge dbc > 0 \right).$$

We emphasize that the notation $abc > 0$ does not imply the assignment of a measure of any kind to angles. It is just a statement that a point can be found to witness, using the betweenness relation, that the angle is not zero. To express that angle abc is both positive and has a positive supplement, we abbreviate the conjunction of the two statements $0 < abc$ and $abc < \pi$ as $0 < abc < \pi$.

We can define certain specific angles as follows:

$$abc = 60^\circ \leftrightarrow ab = bc \wedge ab = ac \wedge \#(a, b, c)$$

$$abc = 120^\circ \leftrightarrow \exists d \left(\mathbf{B}(a, b, d) \wedge ab = bd \wedge bc = bd \wedge cd = bc \wedge \#(a, b, c) \right)$$

$$abc = 30^\circ \leftrightarrow \exists d \left(\mathbf{B}(a, d, b) \wedge ad = cd \wedge ad = db \wedge \#(a, b, c) \right)$$

$$abc = 150^\circ \leftrightarrow \exists d \left(\mathbf{B}(a, b, d) \wedge cad = 30^\circ \right)$$

We would like to prove that each of these four angles is a positive angle. That is neither obvious nor easy. In fact, to do so it is necessary to have a strong lower dimension axiom. That issue will be discussed below; but the lower dimension axiom that we use guarantees the existence of an equilateral triangle whose sides have midpoints, whose altitudes have distinct endpoints, and whose medians meet in a central point. That configuration directly implies that $60°$ is an apex angle (hence positive), and $30°$ is an angle of a right triangle.

Lemma 6 *If $abc = 60°$ or $abc = 120°$ or $abc = 30°$ or $abc = 150°$ then $abc > 0$.*
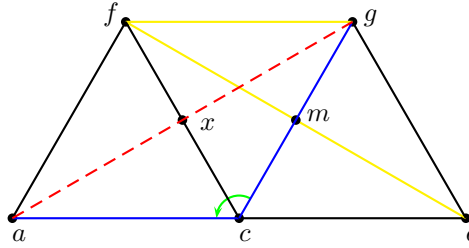

Figure 6: The $120°$ angle acg is positive, since $\mathbf{B}(a, x, g)$.

Proof. The cases of $30°$ and $60°$ follow directly from the lower dimension axiom. We take up the case $abc = 120°$. We will show that the "tiling" shown in Fig. 6 can be constructed. Then $acg = 120°$ and acg is an apex angle because $ac = cg$ and $\mathbf{B}(a, x, g)$.

Remark. It is not trivial to construct the three equilateral triangles shown in the figure and show that $\mathbf{B}(a, c, e)$. This has to be done directly from the axioms, and Pasch cannot be used since we as yet have no positive angles. Euclid 5 will have to be used, as the construction does not work in non-Euclidean geometry.

We start with the equilateral triangle fac and the midpoint x of side fc, whose existence has to be assumed in the lower dimension axiom discussed in § 5.10. Triangle fxa is congruent to triangle cxa, so angle axc is a right angle. Then cxa is also a right angle, by Lemma 5. Hence $gc = ac$, by definition of right angle. Then $ac = ca = fa = gc$. Now cxg is a right angle, so gxc is also a right angle, so $fg = gc$. Let m be the midpoint of gc.

Consider the two lines ac and fg. The transversal fc makes alternate interior angle equal, as witnessed by f, g, x, a, and c. Technically, our formulation of Euclid 5 does not mention angles at all; only the congruence of triangles fgx and cax is needed. The condition $\mathbf{B}(g, m, c)$ fulfills the hypotheses of Euclid 5 (which would traditionally be expressed as "the corresponding interior angles mfc and ecf make less than two right angles"). Hence fm meets the line through ac in a point e. That is, $\mathbf{B}(a, c, e) \wedge \mathbf{B}(f, m, e)$.

Because m is the midpoint of gc, angles fmg and fmc are right angles. Therefore cme and gme are also right angles. Since cmf is a right angle,

$ce = fc = gc$. Since emc is a right angle, $ge = ce$ and triangle gce is equilateral. Then angle acg is, by definition, a 120° angle. But $a\#q$ because $\mathbf{B}(a, x, q)$. That completes the case of a 120° angle. fmc is a right angle, and $mf = me$, also $fc = ce$.

Now consider a 150° angle. Fig. 7 illustrates the construction.
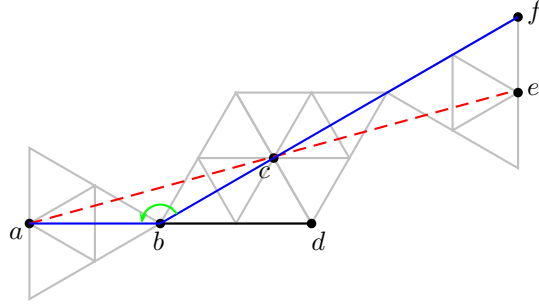

Figure 7: $abc = 150°$ because $cbd = 30°$, and $a\#c$ because $\mathbf{B}(a, c, e)$.

The illustrated network of equilateral triangles can be proved to exist using the 5-segment axiom and Euclid 5, starting from point b, as in the 120° proof above. Then abc is a positive angle, since $\mathbf{B}(a, c, e)$ shows that $a\#c$. But abc is a 150° angle, since the witnesses called for in the definition of 150° angle are provided by d and e. points. That completes the proof of the lemma.

Remark. Fig. 7 illustrates the statement that it takes six points to determine a (positive) angle in non-Markovian geometry. Point b is the vertex; point d witnesses that ab determines a ray; point f witnesses that bc determines a ray; point e witnesses that those two rays are distinct.

Definition 4 *A* triangle *is an ordered triple of distinct points abc such that all three angles abc, bca, and cab are positive angles with positive supplements.*

5.6   Non-Markovian inner Pasch

Tarski used (at various times) "outer Pasch" and "inner Pasch", both formulated using non-strict betweenness $\mathbf{T}$ and allowing various degenerate cases that are not constructively valid. These axioms are illustrated in Fig. 8. As axioms they have two important advantages over other forms of Pasch: they are purely existential, and they do not depend on the dimension. That is, they hold in $\mathbb{R}^n$ for any n, in contrast to the version of Pasch that says, if a line enters a triangle it must exit, which fails in $\mathbb{R}^3$.

In [4], we formulated what seemed the most general constructively valid version of inner Pasch, replacing $\mathbf{T}$ by $\mathbf{B}$ in the hypothesis $\mathbf{T}(a, p, c)$ and in the two conclusions, and requiring points a, b, c in Fig. 8 to be not collinear. If Markov's principle is dropped, and inner Pasch is not modified, Pasch's axiom will imply Markov's principle.
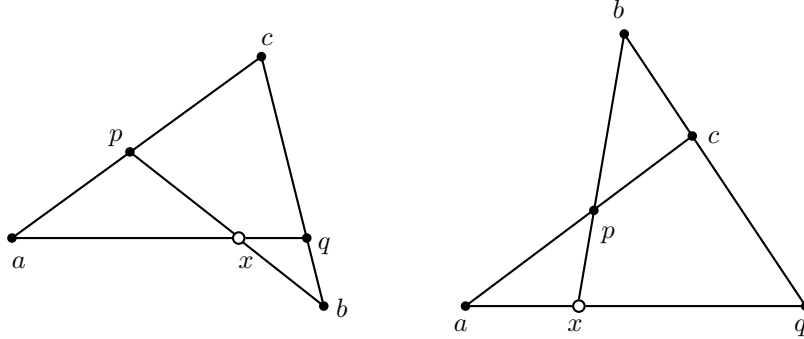
Figure 8: Inner Pasch (left) and outer Pasch (right). Line pb meets triangle acq in one side ac, and meets an extension of side cq. Then it also meets the third side aq. The open circles show the points asserted to exist.

Theorem 2 *Inner Pasch without modifications implies Markov's principle.*

Remark. The theorem refers to the version of inner Pasch described above, and used in [4].

Proof. Let p and v be any two distinct points, and assume unrestricted inner Pasch. (In Fig. 8, v would lie on pb extended; think of v fixed, while b can move along the line vp towards p). We will prove $\neg\neg\,\mathbf{B}(p,b,v)$ implies $\mathbf{B}(p,b,v)$. To that end, assume $\neg\neg\,\mathbf{B}(p,b,v)$; we have to prove $\mathbf{B}(p,b,v)$. Let a and c be points with $\mathbf{B}(a,p,c)$ such that $\neg L(a,v,c)$ (for example, a and c can be taken to lie on the perpendicular to pv at p.) We have $\neg L(a,b,c)$, since $L(a,b,c)$ implies $b = p$ (by the definition of L). Hence we can apply unrestricted inner Pasch, obtaining the existence of x with $\mathbf{B}(p,x,b)$. Now we have $\mathbf{B}(p,x,b)$ and $\mathbf{B}(p,b,v)$. By the inner transitivity of betweenness, we have $\mathbf{B}(p,b,v)$. That completes the proof.

Remark. In that proof, we used the inner transitivity of betweenness and the existence of points a and c not collinear with v, but these things can be proved from unrestricted Pasch, since as we show below, they can be proved even from restricted inner Pasch. Here we only intend to demonstrate the need to restrict inner Pasch in some way.

Our formulation of inner Pasch (suitable for non-Markovian geometry) uses strict betweenness $\mathbf{B}$ in all four places. So the hypotheses include $\mathbf{B}(a,p,c) \wedge \mathbf{B}(b,q,c)$, and the conclusion is $\exists x\,(\mathbf{B}(a,x,q) \wedge \mathbf{B}(p,x,b))$. In a quantifier-free form, x is replaced by a Skolem term. We also need to make sure the whole figure does not degenerate into a line, by requiring the hypothesis $0 < acb < \pi$.

For comparison: Tarski used **T** instead of **B**, and did not care if the figure degenerated. For constructive geometry, that is wrong, since it leads to discontinuous dependence of x on the parameters. In [4], the hypothesis **T**(b, q, c) was retained but **B** was used for the other hypothesis and the two conclusions. That is, the degenerate cases $q = c$ and $q = b$ are still allowed. This is constructively sensible as the lines aq and pb still are transverse, so have a unique intersection point, constructible with a ruler. To avoid discontinuity, we also required $\neg L(a, b, c)$.

In non-Markovian geometry, the negatively-phrased non-collinearity is not enough, as the lemma above shows. Instead we need $0 < acb < \pi$. The hypothesis $0 < acb < \pi$ is expressed in the way defined in § 5.5. That is an existential statement, but it occurs only in the hypotheses of inner Pasch, so inner Pasch is still equivalent to a quantifier-free axiom (when a Skolem term is used for x).

There are, however, other possible hypotheses that we might also use in connection with inner Pasch; that is, other possible sufficient conditions than $0 < acb < \pi$ that should support the conclusion of inner Pasch without implying Markov's principle. One such hypothesis is $0 < qpa < \pi$. (Or equivalently, $0 < pqb < \pi$.) This may seem strange (after all line pq is not even drawn in the diagram for Pasch's axiom), but it is the natural condition that we need when we use inner Pasch to prove Euclid's exterior angle theorem. We were not able to prove the exterior angle theorem unless we take the form of inner Pasch with this hypothesis as an axiom, as well as the more natural $0 < acb < \pi$. Of course, after developing the theory of perpendiculars, we can prove these conditions equivalent, but to get off the ground we need both versions of inner Pasch.

5.7   Non-Markovian outer Pasch

Please refer to Fig. 8. The hypotheses of unrestricted outer Pasch in [4] are **B**(a, p, c) and **T**(b, c, q), as well as $\neg L(a, b, c)$. The conclusion is that there exists an x such that **B**(a, x, q) and **B**(b, p, x).

Outer Pasch, like inner Pasch, in its unrestricted form implies Markov's principle. The argument is similar to the one given for inner Pasch; in the second part of Fig. 8, we allow a to move towards q, as for inner Pasch we allowed b to move towards p. Therefore, outer Pasch needs to have its hypotheses strengthened. As for inner Pasch, we replace **T** by **B**, and replace the non-collinearity hypotheses by requiring an angle to be between 0 and π. But which angle?

The hypothesis we choose to formulate non-Markovian outer Pasch is this:
$$0 < baq < \pi \lor 0 < abq < \pi.$$

This choice is a pragmatic one: it enables the uses of outer Pasch that we need (in particular for the crossbar theorem). Eventually, we will prove that if a, b, and c are three distinct points, and one of three angles formed is positive, so are the other two. After that, the various possible versions of non-Markovian outer Pasch will be equivalent.

Our theory EG includes both non-Markovian inner Pasch and non-Markovian outer Pasch as axioms. Gupta proved ([11]; see also [26], Satz

9.3), that inner Pasch implies outer Pasch; but this is a long development, and although we did not find a use of Markov's principle, we are not willing to certify that none is necessary. As mentioned above, we are not aiming in this paper to find a minimal set of axioms for EG, but instead a sensible set of axioms that provides a smooth non-Markovian constructive development of Euclid. Therefore we include both forms of Pasch.

5.8   Line–circle and circle–circle continuity

We say that point p is *inside the circle with center a passing through c* provided p lies on a diameter of the circle.[11] The line–circle continuity axiom says that if p is inside the circle with center a passing through c, then there exist two points x and y on the circle with $\mathbf{B}(x,p,y)$. This axiom is true in $\mathbb{R}^n$ for any n, i.e., not just in plane geometry. Without any dimension axiom, "circles" become spheres (in R^3) or hyperspheres.

We also need a non-strict version of line–circle continuity in which $\mathbf{B}$ is replaced by non-strict betweenness $\mathbf{T}$. Classically, the two are easily proved equivalent, but the case distinction whether p is strictly inside or on the circle is not legitimate intuitionistically. Applied to Descartes's geometric construction of square roots, the non-strict version implies that non-negative segments have square roots, and the strict version implies that positive segments have positive square roots. We therefore need both versions of line–circle continuity to establish the existence of coordinates and define arithmetic.

The circle–circle continuity axiom says that if C and K are circles with distinct centers, and C has a point non-strictly inside K, and a point non-strictly outside circle K, then there is a point lying on both C and K. This immediately implies the corresponding strict version, but not conversely, so we take the non-strict version as our axiom.[12]

A stronger version of circle–circle continuity has the additional hypothesis that p is any point not lying on the line L containing the centers of C and K, and the additional conclusion that C and K have an intersection point on the opposite side of L from p; that is, there are points e and x with e on both C and K and x on L and $\mathbf{B}(e,x,p)$. We do *not* assume this stronger version of circle–circle continuity as an axiom.[13]

5.9   Segment-circle continuity

The segment-circle continuity principle says that if p is a point inside a circle, and q is a point outside, then there is a point on the circle be-

---

[11]It will not do to say p is equal to the center or lies on a radius, since we may not know which alternative holds.

[12]We avoided "degenerate cases" of Pasch's axiom, in which a two-dimensional picture degenerates to one dimension. Using $\mathbf{T}$ in the circle–circle axiom is not a degenerate case in this sense; when the circles are tangent, the picture is still two-dimensional.

[13]This means that Euclid's proof of I.4 (angle bisection) is not directly formalizable in EG, and his proof of the existence of a midpoint is consequently also not directly formalizable. Instead, we construct midpoints using Gupta's theorem about the existence of a midpoint of an isosceles triangle, combined with Euclid's construction of an equilateral triangle on a given segment. Once midpoints are in hand, perpendiculars can be constructed, and the stronger version of circle–circle continuity can be proved.

tween p and q. This principle has been suggested as an axiom by many authors, including Tarski (see [27]). But a detailed study shows that it is inadequate; an irremovable circularity arises in formalizing Euclid without a dimension axiom. If we try to construct dropped perpendiculars (Euclid I.12) using segment-circle continuity, to check the hypotheses we need the triangle inequality (I.20). But I.19 is needed for I.20, and I.7 for I.19. In Prop. I.7, the two triangles that are supposed to coincide may lie in different planes; that possibility has to be removed by an additional hypothesis. Even with its statement thus corrected, I.7 is more difficult to prove than Euclid thought, since he took for granted the fact that an angle cannot be less than itself, but that principle is actually the essential content of I.7. Ever since Hilbert [15], angle inequality has been regarded as a defined concept, and proving I.7 then requires dropped perpendiculars (I.12). But this is circular. The conclusion is that segment-circle continuity is not a suitable axiom to use in formalizing Euclid.[14]

5.10   Dimension axioms

When considering plane geometry, one needs a *lower dimension axiom* providing for the existence of three non-collinear points. Non-collinearity will not be enough in non-Markovian geometry; we need to ensure that there exist three points forming a positive angle with a positive supplement. In other words, we need three distinct points α, β, and γ such that $0 < \alpha\beta\gamma < \pi$.[15]

We take as an axiom that there are three points, given by constants α_1, α_2, α_3, that form an equilateral triangle. Then we add a further axiom, introducing three more constants for the midpoints of the three sides, and one final constant for the common intersection point of two of the medians. Thus the three vertices are all distinct, and two of the medians have distinct endpoints. See Fig. 9. The formal expression of this axiom is given in the next section, where all the axioms are listed for reference.

The upper dimension axiom for plane geometry says that if a and b are distinct, and three points are each equidistant from a and b, then those three points are collinear. If we do not use any upper dimension axiom, or we replace 3 by some larger integer, then as we have already noted, inner and outer Pasch still make sense, and circle–circle continuity becomes

---

[14]line–circle continuity does not suffer from this problem, as the triangle inequality is not required to drop perpendiculars. Of course, as Gupta showed, one can construct dropped perpendiculars without mentioning circles at all, so there is no formal result that one continuity axiom is better for I.7 than another, as none at all is actually needed. We merely say that Euclid's proof can be repaired with line–circle, but not with segment-circle.

[15]If we do not specify this in an axiom, nothing allows to prove any betweenness statement, as the first deduction of a betweenness statement cannot come from the betweenness axioms (which have betweenness hypotheses), nor from Pasch (which has betweenness in the hypothesis about a positive angle), nor from the extension axiom, which requires betweenness in the hypothesis about distinct points, nor from circle–circle or line–circle continuity, where the hypotheses about "inside" are expressed using betweenness. Nothing would prevent the plane from collapsing, with all points not not equal to each other. Hence, it is reasonable to call this a "lower dimension axiom".
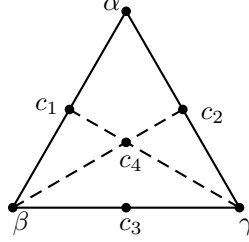
Figure 9: The lower dimension axiom

sphere–sphere continuity. The notions "same side" and "opposite side" of a line still make sense if properly defined. See [4] for the definitions.

5.11   List of axioms for reference

In this section, we give the complete list of axioms of EG. First, we specify the language. It is first-order predicate calculus with equality. There is a 3-ary relation symbol $\mathbf{B}$ and a 4-ary symbol E. E is the official version of the 4-ary relation that we write informally as $ab = cd$; first-order predicate calculus as found in textbooks does not permit that syntax, so $ab = cd$ must remain as an informal abbreviation for the more formal syntax $E(a, b, c, d)$.

A number of other informal abbreviations are used in stating the axioms. The actual axioms are the result of replacing these abbreviations by the right-hand side of their definitions (recursively), so the axioms involve only equality, $\mathbf{B}$, and E. The following is a complete list of these abbreviations. Since each right-hand side involves only definitions earlier in the list, the recursive replacement mentioned does terminate. That is, there is no circularity in this list of definitions.

$$
\begin{aligned}
a\#b \quad &:= \quad \exists c\,(\mathbf{B}(c,a,b) \vee \mathbf{B}(a,c,b) \vee \mathbf{B}(a,b,c)) \\
ab = cd \quad &:= \quad E(a,b,c,d) \\
R(a,b,c) \quad &:= \quad \exists d\,(\mathbf{B}(a,b,d) \wedge ab = bd \wedge ac = dc \wedge a\#b \wedge c\#b \wedge a\#c) \\
Ray(a,b,x) \quad &:= \quad \exists d\,(\mathbf{B}(d,a,b) \wedge \mathbf{B}(b,a,x)) \\
apex(a,b,c) \quad &:= \quad \exists u,v\,(Ray(b,a,u) \wedge Ray(b,c,v) \wedge bu = bv \wedge u\#v) \\
\angle abc = \angle ABC \quad &:= \quad \exists U,V,u,v\,(Ray(b,a,u) \wedge Ray(b,c,v) \\
&\qquad\qquad \wedge Ray(B,A,U) \wedge Ray(B,C,V) \wedge uv = UV) \\
0 < abc \quad &:= \quad apex(a,b,c) \vee R(a,b,c) \vee \\
&\qquad\qquad \exists A,B,C\,(\angle abc = \angle ABC \wedge R(B,A,C) \vee R(A,C,B)) \\
abc < \pi \quad &:= \quad \exists d\,(\mathbf{B}(d,b,a) \wedge 0 < dbc) \\
0 < abc < \pi \quad &:= \quad 0 < abc \wedge abc < \pi \\
\mathbf{T}(a,b,c) \quad &:= \quad \neg(a \neq b \wedge \neg\mathbf{B}(a,b,c) \wedge b \neq c)
\end{aligned}
$$

Here are the betweenness axioms:

$\neg\mathbf{B}(a,b,a)$ (A6-i)

$\mathbf{B}(a,b,c) \to \mathbf{B}(c,b,a)$ (A14-i), symmetry of betweenness

$\mathbf{B}(a,b,d) \wedge \mathbf{B}(b,c,d) \to \mathbf{B}(a,b,c)$ (A15-i), inner transitivity

$\mathbf{B}(a,b,d) \wedge \mathbf{B}(a,c,d) \wedge$
$\neg\mathbf{B}(a,b,c) \wedge \neg\mathbf{B}(a,c,b) \to b = c$ (A17-i), connectivity

Here are the axioms concerning congruence and betweenness:

$A\#B \wedge ab = AB \to \exists e\,(\mathbf{T}(a,b,e) \wedge be = cd)$ (A4-i1) extension

$A\#B \wedge ab = AB \wedge C\#D \wedge cd = CD$
$\to \exists e\,(\mathbf{B}(a,b,e) \wedge be = cd)$ (A4-i2) strict extension

$a\#b \wedge ab = AB \wedge bc = BC \wedge ad = AD$
$\wedge\, bd = BD \to cd = CD$ (A5-i) five-segment axiom

$\mathbf{B}(a,p,c) \wedge \mathbf{B}(b,q,c) \wedge (0 < acb < \pi \vee 0 < qpa < \pi)$
$\to \exists x\,(\mathbf{B}(p,x,b) \wedge \mathbf{B}(a,x,q))$ (A7-i1) inner Pasch

$\mathbf{B}(a,p,c) \wedge \mathbf{B}(b,c,q) \wedge (0 < baq < \pi \vee 0 < abq < \pi)$
$\to \exists x\,(\mathbf{B}(b,p,x) \wedge \mathbf{B}(a,x,q))$ (A7-i2) outer Pasch

Here are the lower dimension axioms:

$$\alpha\beta = \beta\gamma \wedge \alpha\beta = \alpha\gamma \wedge \alpha \neq \beta$$
$$\mathbf{B}(\alpha, c_1, \beta) \wedge \alpha c_1 = c_1\beta$$
$$\mathbf{B}(\alpha, c_2, \gamma) \wedge \alpha c_2 = c_2\gamma$$
$$\mathbf{B}(\beta, c_3, \gamma) \wedge \beta c_3 = c_3\gamma$$
$$\mathbf{B}(\beta, c_4, c_2) \wedge \mathbf{B}(\gamma, c_4, c_1).$$

We have used constants c_i and α, β, γ; as a result the lower dimension axiom is quantifier-free and for convenience can be broken into several formulas, written on different lines without being connected by $\wedge$. This is of no significance—we might as well have used existentially quantified variables, in which case the dimension axiom would need to be one long formula inside the scope of the existential quantifiers.

There is no upper dimension axiom.

The first line–circle continuity axiom says that if point a is inside the circle with center c and radius pq (as witnessed by a lying on a diameter uv), and b is a point distinct from a (so that ba determines a line), then there are two points x and y on the circle, with a between them.

$$\mathbf{B}(u,a,v) \wedge cu = pq \wedge cv = pq \wedge\ \mathbf{B}(u,c,v)$$
$$\to \exists x,y\,(cx = pq \wedge cy = pq \wedge \mathbf{B}(x,a,y))$$

The second line–circle continuity axiom is similar, but with $\mathbf{T}$ instead of $\mathbf{B}$.

$$\mathbf{T}(u,a,v) \wedge cu = pq \wedge cv = pq \wedge\ \mathbf{T}(u,c,v)$$
$$\to \exists x,y\,(cx = pq \wedge cy = pq \wedge \mathbf{T}(x,a,y))$$

Here is the circle–circle continuity axiom, which says that if the circle with center c and radius pq has a point a non-strictly inside the circle with

center C and radius PQ, and a point b non-strictly outside that circle, then the two circles meet.

$$ca = pq \wedge \mathbf{T}(u, a, v) \wedge \mathbf{T}(u, C, v) \wedge Cu = PQ \wedge Cv = PQ$$
$$cb = pq \wedge \mathbf{B}(C, w, b) \wedge Cw = PQ$$
$$\rightarrow \exists e \, (ce = pq \wedge Ce = PQ)$$

Here is Euclid 5. See Fig. 4 for an illustration.

$$pt = qt \wedge \mathbf{B}(p, t, q) \wedge st = rt \wedge \mathbf{B}(s, t, r) \wedge pr = qs$$
$$\wedge \, \mathbf{B}(q, a, r) \rightarrow \exists e \, (\mathbf{B}(p, a, e) \wedge \mathbf{B}(s, q, e))$$

Logical form of the axioms. The occurrences of disjunction in the hypotheses of Pasch's axiom can easily be eliminated, using two formulas for inner Pasch instead of one, and two for outer Pasch. If the defined symbols are replaced by their definitions, that creates existential quantifiers in the hypotheses (which can be eliminated by simple logic) and in the conclusions (which can be eliminated by introducing function symbols). Thus, there is an equivalent formulation in which the axioms are disjunction-free and quantifier free. This observation is important for certain meta-theorems discussed below.

6 Development of non-Markovian geometry

In this section, we check the development of constructive geometry in [4] and [3], searching for uses of Markov principle, and try to eliminate them when found. In [3], once having found an axiom system that supports the double-negation interpretation for geometry, (for which the use of Markov's principle is crucial), we were able to use it to "import" negative theorems from [26]. If Markov's principle is rejected, we can no longer do that. In order to reduce theorems to negative form, we had to find and prove the correctness of numerous "uniform" constructions, to eliminate arguments by cases; those proofs necessitated the direct formal development of a certain amount of geometry. If we now want to eliminate Markov's principle, it at first appears that the entire development of the two cited papers has to be checked, and in addition, the parts of [26] that could, in the presence of Markov's principle, be "imported" via the double negation interpretation.

However, we found a way to avoid some of that work. Namely, as mentioned above, we include as axioms of EG both inner and outer Pasch, and both line–circle and circle–circle continuity. With Markov's principle, this is overkill, as either version of Pasch implies the other, and either continuity axiom implies the other. Perhaps these implications can also be proved without Markov's principle, but verifying that would add nothing to the philosophical points of this paper. Gupta proved (constructively and without Markov's principle), using inner Pasch, that the base of an isosceles triangle has a midpoint. Having circle–circle continuity allows us to prove Euclid I.1, so every segment is the base of an isosceles triangle

34

and hence has a midpoint. From there we can proceed to perpendiculars and thence to the main body of [4], eliminating the need to check [26]. In particular, we need to check neither Gupta's proof of outer Pasch nor his construction of midpoints without using circles, both of which are long and difficult. Why then, do we need outer Pasch? Aside from its innate interest, outer Pasch is the key tool to prove the plane separation theorem, Theorem 2.5 of [3].[16]

6.1   Midpoints

Euclid's own midpoint construction (Prop. I.10) is to construct an isosceles triangle on pq and then bisect the vertex angle. But Euclid's proof is defective; a correct proof has to rely on Pasch's axiom. Gupta showed in 1965 [11] how to construct midpoints without using circles. His proof is complicated, but we are allowed to use circles, so Gupta's construction is irrelevant for this paper.

One of Gupta's "simpler" theorems enables us to justify the second part of this Euclidean midpoint construction, and we present the main construction of that lemma next, for the reader's enjoyment (it is short and beautiful), and because we need to check (in the next lemma) that it relies only on a version of Pasch that is acceptable when Markov's principle is rejected.

Theorem 3 (Gupta) *Assuming Markov's principle, the base of an isosceles triangle always has a midpoint.*
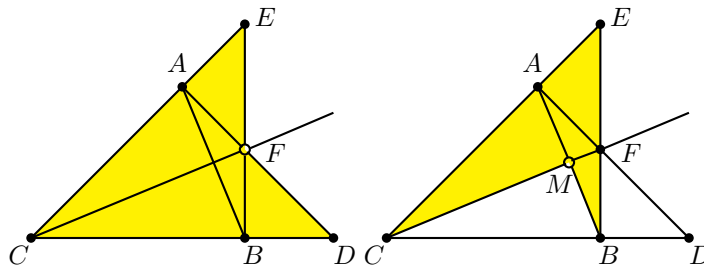
Figure 10: To construct the midpoint M of AB, given C with AC equal to BC, constructing first F and then M by inner Pasch.

Proof. See page 56 of [11]. We repeat the proof here, because below we will adapt it to work without Markov's principle. First we just repeat it as Gupta gave it. We are not claiming that this proof works in EG.

Let ABC be an isosceles triangle with AC equal to BC. It is desired to find the midpoint M of AB. Let α and β be any two distinct points,

---

[16]Gupta's proof of outer Pasch from inner Pasch does not use the parallel axiom. That raises the possibility that it might be easy to prove outer Pasch if we allow the use of the parallel axiom. But then, we have to ask which version of the parallel axiom. In [3], Theorem 9.2, we show that Euclid 5 implies Tarski's parallel axiom, but we used outer Pasch in the proof. It is not difficult to prove that Tarski's parallel axiom plus inner Pasch implies outer Pasch. But we do not have a short proof of outer Pasch from Euclid 5. We do have a short proof of inner Pasch from outer Pasch and Euclid 5.

and extend both CB and CA by $\alpha\beta$ to produce points D and E as shown in Fig. 10. As shown in the figure, two applications of inner Pasch produce points F and M. The point M is the desired midpoint. The proof that M is in fact the desired midpoint is not quite straightforward, but it appears in [11], p. 56, so we do not repeat it here.

We now consider whether Gupta's proof still works without Markov's principle. The point is that to use non-Markovian Pasch we would need to prove that the angles in question are positive and less than π. In general there seems to be no way to justify that, since we do not yet have available midpoints and perpendiculars. However, we do have circle–circle continuity, and hence Euclid I.1, so we only need to make Gupta's proof work for equilateral triangles, not for isosceles triangles in general. And, in Gupta's proof, we get to choose the points α and β, which Gupta needed only to be any distinct points.

Here is Gupta's theorem for equilateral triangles, proved without Markov's principle.

Theorem 4 *[Gupta without Markov] Intuitionistic Tarski geometry (without any continuity axioms and without Markov's principle) proves that if A and B are distinct points, and ABC is an equilateral triangle, then AB has a midpoint.*

Proof. We choose $\alpha = A, \beta = B$. Then A is the midpoint of CE and B is the midpoint of CD. Because ABC is an equilateral triangle, angle $ACB = 60°$ (in the sense precisely defined in §5.5). Therefore by Lemma 6, $ACB > 0$. Let P be the reflection of B in C, so $\mathbf{B}(P, C, B)$ and $PC = CB$. Then $PCA = 120°$, so by Lemma 6, $PCA > 0$ and $ACB < \pi$. That justifies the first application of inner Pasch, so point F in Gupta's proof exists.

To justify the second application of inner Pasch, we need to show that $0 < CEB < \pi$. That is, $0 < AEB < \pi$. This is where we use $CA = AE$, which gives us (using points A, B, C and E) the conclusion $AEB = 30°$. Hence, by Lemma 6, $AEB > 0$. Let Q be the reflection of A in E, so $\mathbf{B}(A, E, Q)$ and $EQ = AE$. Then $BEQ = 150°$, so by Lemma 6, $BEQ > 0$, so $AEB < \pi$. Then the second application of inner Pasch is justified, and the point M exists. The fact that M, once constructed, is indeed the desired midpoint can be proved exactly as in Gupta's thesis. That completes the proof.

Theorem 5 *Every positive segment has a midpoint.*

Proof. Since circle–circle continuity enables us to construct an equilateral triangle on any segment (via Euclid I.1), and EG has circle–circle continuity, the corollary follows from Gupta's theorem. We remark that Euclid I.1, like inner Pasch, does not depend on a dimension axiom, since in R^n, so-called circle–circle continuity is really sphere-sphere continuity, so Euclid I.1 holds without using a dimension axiom.

Lemma 7 *Let a, b, and c be three distinct points with $0 < abc < \pi$, and suppose $\mathbf{B}(b, e, c)$. Then $a \# e$.*

Proof. See Fig. 11. Let d be the midpoint of ab. Since angle $0 < abc < \pi$, we can apply inner Pasch to the configuration $adbce$. The result is a point

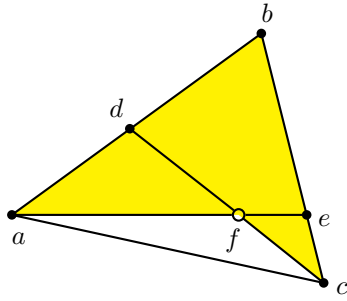Figure 11: Given $a\#b$, $a\#c$, $B(b, e, c)$, and $0 < abc < \pi$, construct f showing $a\#e$.

f with $\mathbf{B}(a, f, e)$. Hence, by the third clause in the definition of $\#$, $a\#e$. That completes the proof.

6.2   Uniform perpendicular

In [3], we gave two different constructions of the uniform perpendicular (to a line L through a point p, without assuming that p is or is not on line L). One construction assumes the parallel axiom, but not line–circle continuity. The other assumes line–circle continuity, but not the parallel axiom. (It is an open problem to do it without assuming either of the two.) In this paper, we assume both those two hypotheses, so we can disregard the comparatively difficult construction that avoids line–circle continuity. The other construction is fairly straightforward: just draw a "large enough" circle C about point p (large enough that C meets L in two points a and b with $a\#b$), and then bisect segment ab to find the foot f of the perpendicular, and then erect the perpendicular to L at f, using the construction of Euclid I.1 to construct an equilateral triangle over ab and connect its vertex to f. The only tricky part of this construction is to find the radius to use to draw C. Line L is given by two points, say u and v, with $u\#v$. So segment uv can be extended; extend it by pu (which may be null or not, we do not care) and again by pv. The resulting segment is at least uv longer than pu or pv, so it can serve as the radius of C.

6.3   Euclid does not need Markov

Does Markov's principle, or the stability of congruence, actually play an important role in Euclidean reasoning? We searched for a theorem in Euclid Books I-IV in which Markov's principle is needed. There is an *a priori* constraint, which we consider first.

First, the double negation interpretation. For simplicity we consider the version of EG based on Tarski's language and axioms, with Skolem symbols *ext* for the segment-extension axiom, *ip* for inner Pasch, and *e*5 for Euclid 5. As discussed at the end of §5.11, the axioms are then quantifier-free and can be put in the form $A \to B$, where A and B are conjunctions of atomic formulas. Let ϕ^- be the double-negation interpretation of ϕ. Since double negation commutes with implication and conjunction, each axiom ϕ satisfies $\phi^- \leftrightarrow \neg\neg\, \phi$. Each theorem ψ of Euclid is also of the form just described. Hence, if the conclusion of ψ mentions only congruence and equality (and not betweenness), then we will have $\psi^- \to \psi$, in EG plus the stability of congruence. Therefore, any theorem of Euclid that actually requires Markov's principle must mention betweenness in its (formalized) conclusion.

We therefore examined the propositions of Euclid looking for theorems whose conclusions involve collinearity, angle ordering, or segment ordering. This investigation is slightly complicated, because Euclid contains well-known errors, and because Euclid has never yet been formalized faithfully (i.e., in a theory that permits correcting Euclid's proofs, while assuming not too much more than Euclid did).[17] Nevertheless, those errors can be

---

[17]There have certainly been many correct formal theories of geometry capable of proving versions of the propositions of Euclid. But that is not quite the same thing as formalizing Euclid. For example, to get the two circles in Proposition I.1 to intersect, clearly Euclid is implicitly assuming some kind of circle–circle continuity. But Euclid has no dimension axiom, so circles are really "spheres". Did Euclid intend to assume that two circles intersect on a

corrected, so the search for a possible use of Markov's principle is possible, even in this rough terrain.

The first one is Euclid I.14, which says

> If with any straight line, and at a point on it, two straight lines not lying on the same side make the adjacent angles equal to two right angles, the two straight lines will be in a straight line with one another.

The collinearity statement in the conclusion is naturally expressed using betweenness. For a diagram, refer to your copy of Euclid. Our point about this example is that stability of betweenness is not actually needed. Let E (in Euclid's figure) lie on CB extended, with $BE = BC$. Let $PB \perp BC$; then CBP and PBE are right angles, so CBA and ABE are equal to two right angles. Hence angles CBA and ABE together equal angles CBA and ABD. Hence angle ABE equals angle ABD. It follows that E and D lie on the same line CBE on the same side of B. (That follows in Tarski's system from the definition of equality of angles.) We do not need the argument using Markov's principle that Euclid gives in I.14.

6.4 The exterior angle theorem and its consequences

The next potential example is I.16, the exterior angle theorem. Euclid's proof constructs a crucial point F, but he left a gap in failing to prove that F lies in the interior of a certain angle.

The proof is also instructive, for those not yet completely familiar with Tarski's treatment of angles, in that it shows how to "unwind" a theorem about angles, to see what it "really states" when angles are eliminated in favor of a points-only formulation.

Lemma 8 (Exterior angle theorem, Euclid I.16) *Suppose* $\mathbf{B}(B, C, D)$ *and* $A \# B$ *and* $0 < BAC < \pi$. *Then* $ACD > BAC$.

Remark. The hypotheses are weaker than "ABC is a triangle", which by definition requires all three angles to be positive and have positive supplements.

Proof. Euclid's diagram is extended by Fig. 12. We show how to complete Euclid's proof. To prove $ACD > BAC$, we must construct a point F in the interior of ACD such that $ACF = BAC$. Euclid knew what angle ordering means: he constructs F as shown in the figure. We can construct E (and hence F) without Markov's principle, since by Lemma 7, $B \# E$, so segment BE can be extended as required. The lemma is applicable by the hypothesis $0 < BAC < \pi$.

given side of the line connecting their centers, or just that they intersect *somewhere*? If only the latter, then his proof that an angle could be bisected is wrong. If the former, then he would not need to bisect an angle first, but could bisect a segment directly by constructing two equilateral triangles on opposite sides of the segment. Proposition 7 patently fails in three-space, so perhaps Euclid did mean to have a dimension axiom. But eventually he works on the Platonic solids, so he needs to *not* have a dimension axiom.
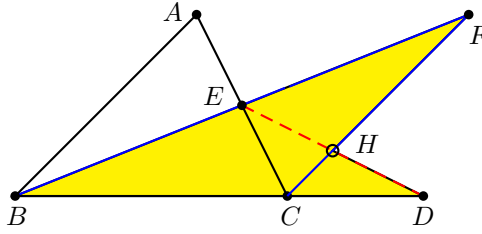
Figure 12: Triangle ABC has exterior angle ACD. E is the midpoint of AC and also of BF. H exists by inner Pasch.

Euclid neglects to prove that F lies in the interior of ACD. To fill this gap in Euclid, we need point H in the figure to exist. We wish to obtain H from inner Pasch, applied to the five-point configuration shaded in Fig. 12. But to use inner Pasch in Markov-free geometry, we need an appropriate angle to be positive. It would suffice to show that $0 < FBC < \pi$, but there seems to be no way to do that until the theory of positive angles is developed, which cannot precede the exterior angle theorem. However, our formulation of inner Pasch allows it to be applied if $0 < ECF < \pi$.

Since E is the midpoint of both AC and BF, and vertical angles are equal (Lemma 3), triangles AEB and CEF are congruent, and triangles AEF and CEB are congruent, by the SAS congruence theorem (Euclid I.2); that is, $AF = CB$. Then triangles ABC and CFA are also congruent; hence angle BAC is equal to angle FCA. Since $0 < BAC < \pi$, we have $0 < FCA < \pi$. But angle FCA is equal to angle FCE, since $\mathbf{B}(C, E, A)$. Angle FCE is equal to angle ECF. Therefore $0 < ECF < \pi$, justifying the desired application of inner Pasch. That completes the proof that F is in the interior of angle ACD, which in turn completes the proof of the exterior angle theorem.

Remark. It is an open question whether the exterior angle theorem is still provable if inner Pasch is formulated with only the hypothesis $0 < acb < \pi$ instead of the hypothesis $0 < acb < \pi \lor 0 < pqb < \pi$. We tried various ways of applying outer Pasch without success.

The immediate corollary (Euclid I.17) is that any two angles of a triangle, taken together, are less than two right angles. In particular, no triangle contains two right angles.

Lemma 9 *In a right triangle, the hypotenuse is greater than either leg.*

Proof. First prove Euclid I.18 and I.19. Then apply them as indicated in Exercise 22, p. 198 of [10]. Alternately, see Satz 11.46 of [26]; the proof there uses nothing but elementary betweenness and congruence, and the existence of perpendiculars.

6.5   Crossbar theorem

The "crossbar theorem" proved in this section is an easy consequence of outer Pasch, but I do not know how to derive it from inner Pasch (without Markov's principle).

Theorem 6 *Let a, b, and c be distinct points with $0 < abc < \pi$ and $0 < buv < \pi$. Suppose $\mathbf{B}(a, e, c)$. Let u and v be points with $\mathbf{B}(b, a, u)$ and $\mathbf{B}(b, c, v)$. Then there exists a point w with $\mathbf{B}(u, w, v)$ and $\mathbf{B}(b, e, w)$.*
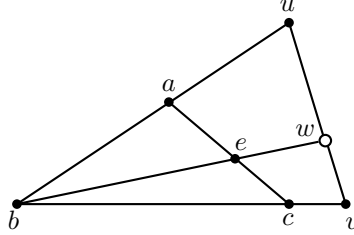
Figure 13: Crossbar theorem. $Ray(b, e)$ meets the crossbar uv.
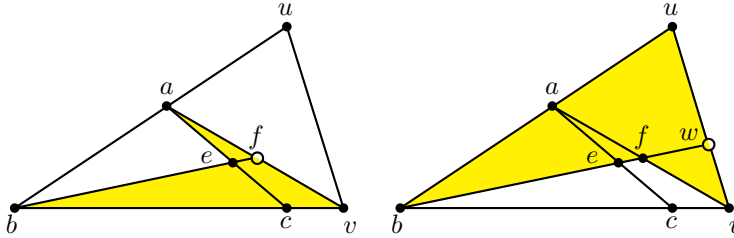
Figure 14: Proof of the crossbar theorem. Apply outer Pasch twice.

Proof. Because we are given $0 < abc < \pi$, we can apply non-Markovian outer Pasch to the configuration *bcvae*. The conclusion is the existence of a point f with $\mathbf{B}(a, f, v)$ and $\mathbf{B}(b, e, f)$. Now we apply outer Pasch again, this time to the configuration *bauvf*. Again, the required hypothesis is fulfilled, since $0 < abv < \pi$, and *abv* is the same angle as *abc*. The result is a point w such that $\mathbf{B}(u, w, v)$ and $\mathbf{B}(b, f, w)$. Now we have $\mathbf{B}(b, f, w)$ and $\mathbf{B}(b, e, f)$. By the inner transitivity of betweenness, we have $\mathbf{B}(b, e, w)$ as desired. That completes the proof.

6.6 Further properties of angle ordering

Euclid took the congruence of all right angles as his Postulate 4. Hilbert ([15], p. 20) remarks that this was "unjustified", and says that the proof of it goes back to Proclus.

Lemma 10 *All right angles are congruent. In other words, if abc and ABC are right angles with $ab = AB$ and $bc = BC$ then $ac = AC$.*

Proof. This is Satz 10.12 in [26]. However, the proof appeals only to the definition of angle congruence and simple theorems, such as the fact that reflections in points and in lines are isometries.

Lemma 11 *(i) If a, b, and c are not collinear, the triangle inequality holds: $ac < ab + bc$.*

(ii) Whether or not a, b, and c are collinear, we have $ac \leq ab + bc$.

Proof. The proof we gave in [3], Lemma 8.14, used Markov's principle. But Euclid's proof in I.20, relying on I.5 and I.19, is perfectly constructive as it stands, not requiring Markov's principle or any argument by cases. Part (ii) has a negative conclusion, since $\leq$ can be expressed negatively; hence its provability without Markov's principle follows from the double negation interpretation.

6.7   A lemma about two perpendiculars

The next lemma is used in [3] to prove that every Lambert quadrilateral (plane quadrilateral with three right angles) is a rectangle, which is a key step in establishing a coordinate system on a geometric basis. In [3], Lemma 8.15, we used the exterior angle theorem and Markov's principle to prove the lemma. Here we give a proof without Markov's principle, relying on inner Pasch and Euclid 5. The proof in [3] does not use Euclid 5, i.e., it is a proof in neutral geometry, but it does use Markov's principle.

Lemma 12 *Let a, b, c, and d be distinct points (pairwise distinct) with $ab \perp ad$ and $cd \perp ad$. If bc meets the line containing ad, then the intersection point m is between a and d.*
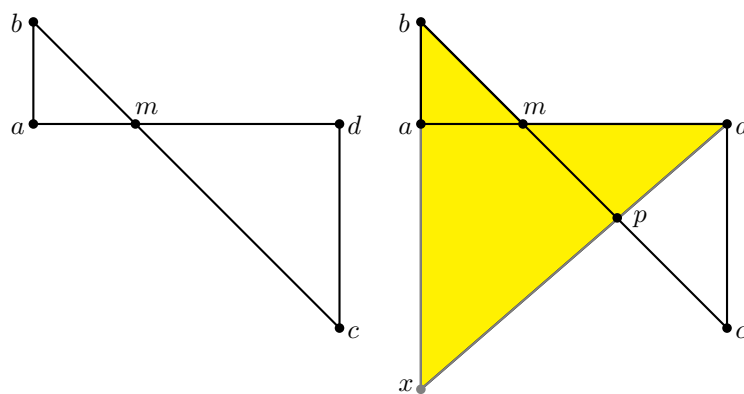
Figure 15: To show $\mathbf{B}(a, m, d)$, construct p, then construct x by Euclid 5, then apply inner Pasch to the shaded configuration.

Proof. Please refer to Fig. 15. The first part of the figure illustrates the theorem, and the second part shows the construction used for the proof. The lines containing ab and cd are parallel, since both are perpendicular to ad. Let p be any point between m and c, for example, the midpoint of mc. We have $\mathbf{B}(b, m, p)$ and $\mathbf{B}(m, p, c)$. Hence $\mathbf{B}(b, p, c)$, by the outer transitivity of betweenness (which has been discussed above).

Then Euclid 5 applies, and yields the existence of a point x such that $\mathbf{B}(d, p, x)$ and $\mathbf{B}(b, a, x)$. Then $bxd > 0$, because it is an angle of the right triangle axd, which has $a \# d$ by hypothesis and $a \# x$ because $\mathbf{B}(b, a, x)$. Now we can apply inner Pasch to the configuration *dpxab*. That yields a point j such that $\mathbf{B}(b, j, c)$ and $\mathbf{B}(a, j, d)$. Then j is the intersection

of ad and bc. Since those segments are not collinear, their intersection is unique, so $j = m$. Hence $\mathbf{B}(a, m, d)$ as desired. That completes the proof.

6.8   Parallelograms

We have not yet defined the word "parallel." It may seem obvious that two lines are parallel just when they do not intersect, but that is only true in plane geometry; skew lines in space are not considered parallel, and "parallel" can be defined in a way that does not depend on the presence of a dimension axiom. Intuitively, two lines are parallel if they lie in the same plane and do not meet. Precisely, we define ab to be parallel to cd if a and b are on the same side of cd and there is no point x such that $L(a, b, x)$ and $L(c, d, x)$. This definition assumes that $a\#b$ and $c\#d$. Then it can be proved, without an upper dimension axiom, that if ab is parallel to cd then cd is parallel to ab. The proof requires the plane separation theorem of Hilbert. A version of that theorem is proved in [4], and it can also be derived in non-Markovian geometry using the crossbar theorem (Theorem 6).

The traditional Euclidean results about parallel lines and parallelograms offer no difficulties in non-Markovian intuitionistic geometry, provided we define a *transversal* of two lines L and K to be a line J that makes angles with L and K, both of which are between 0 and π.

We omit the proofs of the following facts:

Lemma 13 *If lines L and K are cut by a transversal meeting them at distinct points (and "transversal" is defined as above), then*

(i) if L and K are parallel, then alternate interior angles are equal, corresponding angles are equal, and interior angles on the same side of the transversal are together equal to two right angles, and

(ii) if any of the conditions in (i) hold, then L and K are parallel.

Definition 5 *A* parallelogram *is a quadrilateral with opposite sides parallel and adjacent vertices distinct, whose diagonals each are transversals of each pair of opposite sides.*

In other words, the diagonals form angles between 0 and π with the sides. It is not required by this definition that diagonally opposite vertices be distinct.

The following lemmas can all be proved without Markov's principle.

Lemma 14 *Opposite sides of a parallelogram are congruent.*

Lemma 15 *If opposite sides of a quadrilateral are congruent, and adjacent vertices distinct, and the diagonals meet, then the quadrilateral is a parallelogram.*

Lemma 16 *The diagonals of a parallelogram meet, and they bisect each other.*

Proof. Let $ABCD$ be a parallelogram. Extend BC to E with $CB = BE$. Then E is on the opposite side of AB from C. By definition of parallel, C and D are on the same side of AB. By the plane separation theorem, E is on the opposite side of AB from D. Therefore there is a point F
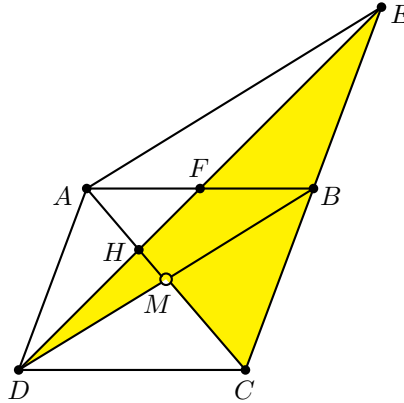
Figure 16: The diagonals of parallelogram $ABCD$ meet at their common midpoint M. F exists by the plane separation theorem. H exists by the crossbar theorem. M exists by inner Pasch.

collinear with AB with $\mathbf{B}(D, F, E)$. Then F is in the interior of angle AEB. Note that $0 < DCB < \pi$, since $ABCD$ is a parallelogram. By Lemma 13 and Lemma 14, and the SAS congruence theorem, triangle ABE is congruent to triangle DCB. Hence angle AEB is equal to angle DBC. Hence $0 < AEB < \pi$. Then the crossbar theorem can be applied to angle AEB, and since $\mathbf{B}(A, F, B)$, the ray FE meets the crossbar AC. Let H be the intersection point. Now we can apply inner Pasch to the configuration $DHECB$ (shaded in the figure), because $0 < BDC < \pi$. The result is a point M such that $\mathbf{B}(C, M, A)$ and $\mathbf{B}(D, M, B)$. That point M is the desired intersection of the two diagonals of $ABCD$. Now by vertical angles (Lemma 3) and Lemma 13, the four triangles into which the diagonals divide $ABCD$ are each congruent to their reflections in M. Hence M is the midpoint of each diagonal. That completes the proof.

Corollary 1 *All four vertices of a parallelogram are distinct.*

Lemma 17 *The lines connecting the midpoints of opposite sides of a parallelogram meet at their common midpoint, which is also the point of intersection of the diagonals.*

Proof. This follows from Lemma 16 and the fact that reflection in a point preserves betweenness and congruence.

6.9 Angle copying

Hilbert took as an axiom the proposition that a given angle can be copied to a new position, specified by a line L, a point p on L, and a point q not on L; the angle should be copied so that its vertex is at p and its third point t is on the opposite side of L from q. That is, a point x such that $\mathbf{B}(t, x, q)$ is asserted to exist.

In our development, angle copying is a theorem. We will prove that any angle can be copied, not just any positive angle. Then evidently the conclusion $\mathbf{B}(t, x, q)$ needs to be weakened to $\mathbf{T}(t, x, q)$.

To specify a side of a line L given by two distinct points p and s, we have to give a point q which is distinct from the foot of the perpendicular to L passing through q. It is not enough just to assume $\neg L(p, s, q)$. We abbreviate this property by "q is distinct from L."

Theorem 7 (Angle copying) *Let p and s be any two distinct points, and let q be any point distinct from the line containing p and s. Assume a and c are distinct from b (but not necessarily $a \# c$). Then angle abc can be copied to the line containing ps, with vertex and p and the third point lying on the opposite side of ps from q. More precisely, there exist points a', c', and x such that angle abc is congruent to angle $a'pc'$, $L(p, s, c')$, and $\mathbf{T}(q, x, a')$. Moreover, if $a \# c$, then $\mathbf{B}(q, x, a')$.*

Proof. Let $c' = ext(s, p, b, c)$. Then $pc' = bc$ and $L(p, s, c')$. Let c' be a point with $bc = pc'$ on $Ray(s, p)$. Erect a perpendicular line K to pc' at c', on the opposite side of ps from q.

Then there are two points a' on K such that $a'c' = ac$. Either choice will make triangle $a'pc'$ congruent to abc, since they are right triangles with corresponding legs equal. We need to choose the correct a' to be on the opposite side of pc' from q. By construction of K, one of the two rays of K emanating from p is on the opposite side of ps from q. Choose a' on that side of ps. That completes the proof.

6.10   Positive angles revisited

Recall that we defined a positive angle by two cases: a positive angle is either an angle of a right triangle (with positive sides) or an apex angle.

Theorem 8 *Suppose abc is any angle with $0 < abc < \pi$. Then $a \# c$.*

Remarks. (1) The special case when abc is a right angle is already interesting. That special case can be stated loosely as, "every right angle has a positive hypotenuse." Note that abc does not count as a right triangle until all its sides are positive; to be a right *angle* it only needs positive legs. That is why we speak of the hypotenuse of an angle here, rather than the hypotenuse of a triangle.

(2) The proof makes use of the parallel postulate. But our theory of non-Markovian geometry has already depended heavily on the parallel postulate. Our point in this paper is simply that non-Markovian geometry is a coherent conception; we make no attempt to develop it without the parallel axiom, in contrast to [3], where we did develop constructive neutral geometry to some extent (with Markov's principle).

Proof. Since abc is an angle, we have $a \# b$ and $c \# b$. We construct a network of copies of abc that "tile" part of the plane near that triangle, as shown in Fig. 17. The idea is that we can draw lines parallel to ab and cb, because $a \# b$ and $c \# b$, but we cannot necessarily draw lines parallel to ac because we only know $a \neq c$. Nevertheless we can construct the tiling shown in the figure. Here are the details:

Extend ab by ab to point d. Copy angle abc so ab goes to bd, and a goes to a point e on the same side of ab as a. Then triangle abc is congruent to triangle bde. Therefore angle cab is congruent to angle ebd. Therefore

Figure 17: Starting with abc with $a\#b$ and $b\#c$, construct the midpoint m of ab and apply Pasch to prove $\mathbf{B}(a, c, f)$. Hence $a\#c$.

segments ac and be are parallel (in the sense that if there are lines through those segments, they cannot meet, because that would contradict the exterior angle theorem). Angle cba is congruent to angle edb, so ed is parallel to cb. Since $c\#b$ and $ed = cb$, we have $e\#b$, so we can extend de by de to a point f with $ef = bc$.

We have $cb = ed$ and $eb = be$. Since ed is parallel to bc, the alternate interior angles cbe and bed are equal. Hence, by SAS, triangle ceb is congruent to triangle dbe. Then $cedb$ is a quadrilateral with opposite sides equal. Hence, by Lemma 15, it is a parallelogram. Hence, by Lemma 14, $cd = bd = ab$. By construction of f, $ef = ed = bc$. Hence, by SAS, triangle cef is congruent to triangle abc. Note, however, that we have not proved that a, c, and f lie on a line, and indeed the betweenness relation $\mathbf{B}(a, c, f)$ is what we are trying to prove, as that is one way to show $a\#c$. We construct two more copies of triangle abc, by constructing a line vf through f parallel to ad and ce, with $vf = ab$ and v on the same side of f as a, and extending df through f by bc to point u. Let m be the midpoint of segment ab, which exists since $a\#b$. Let ℓ be the midpoint of vf, which exists since $vf = ab$ and $a\#b$, so $v\#f$.

Angle bde is congruent to angle abc, because triangle cba is congruent to triangle edb. Since $0 < abc < \pi$ by hypothesis, we have $0 < bde < \pi$. Hence we can apply inner Pasch to the configuration $ufdam$. The result is a point k such that $\mathbf{B}(u, k, m)$ and $\mathbf{B}(a, k, f)$. In fact $k = c$, although we have not proved that. Extending segments eb and uv, we get a large parallelogram $pueq$ as shown in the figure. The intersection of its diagonals is at the midpoint of each diagonal, namely c. Hence $\mathbf{B}(u, c, q)$. Now a and f are midpoints of opposite sides of the parallelogram, so by Lemma 17, $\mathbf{B}(a, c, f)$. Then by definition, $a\#c$. That completes the proof.

Theorem 9 *Every positive angle is an apex angle.*

Proof. A positive angle, by definition, is either an apex angle, or a right angle, or an angle of a right triangle. If it is an apex angle, there is nothing

to prove.

A right angle is an apex angle, since given a right angle, we can lay off the same distance on each of its sides, constructing a right isosceles triangle. By Theorem 8, the base of this triangle has distinct endpoints, and hence, it has a midpoint. That makes the right angle an apex triangle.

It remains to prove that every angle of a right angle is an apex angle. Let acb be a right triangle with the right angle at c. Fig. 18 illustrates the proof.
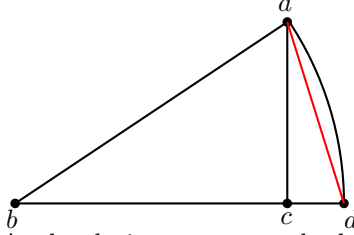

Figure 18: Angle abc is an apex angle, because $a\#d$ as the hypotenuse of right angle acd.

According to Lemma 9, $ac < ab$. By definition of triangle, $a\#c$. Let e be the reflection of c in b, so $\mathbf{B}(e, b, c) \wedge eb = ec$. ($e$ is not shown in the figure.) Then by Lemma 4, there is a unique point d with $\mathbf{B}(e, b, d)$ and $bd = ab$. By Lemma 9, $bc < ab$. Hence $bc < bd$. By definition of $<$, there exists a point c' with $\mathbf{B}(b, c', d)$ and $bc' = bc$. We have $\mathbf{B}(b, c', d)$ and $\mathbf{B}(e, b, d)$. Applying inner transitivity (Axiom A15-i), we have $\mathbf{B}(e, b, c')$. Now we have $\mathbf{B}(e, b, c') \wedge bc' = bc$, and also $\mathbf{B}(e, b, c) \wedge bc = bc$. By the uniqueness part of Lemma 4, $c = c'$. Since $\mathbf{B}(b, c', d)$ we have $\mathbf{B}(b, c, d)$. Therefore $c\#d$. Now acd is a right triangle since acb is a right angle, and both legs cd and ac have distinct endpoints. Applying Theorem 8 to right angle acd, we obtain $a\#d$. Since $ba = bd$, this proves abc is an apex angle, as desired. That completes the proof.

Theorem 10 (Angle bisection) *Every positive angle can be bisected.*

Proof. Let abc be a positive angle. Then by Theorem 9, abc is an apex angle. That is, there are points u, v on $Ray(b, a)$ and $Ray(b, c)$ respectively such that $bu = bv$ and $u\#v$. Since $u\#v$, Theorem 5 implies that uv has a midpoint, say m. By Theorem 8, since $b\#u$ and $u\#m$, we have $b\#m$. Hence bmu and bmv are triangles. Since $um = vm$, angle ubv is bisected by $Ray(b, m)$. Since angle abc is the same angle as ubv, angle abc is bisected by $Ray(b, m)$. That completes the proof.

Remark. Euclid's Proposition I.9 purports to prove that every angle can be bisected; but the proof has more than one flaw. First, if the angle to be bisected is equilateral, the two points determining Euclid's bisector might coincide, and thus fail to determine a line. Second, without any dimension axiom, the equilateral triangle produced by applying I.1 in the proof of I.9 might lie in a different plane from the angle to be bisected, thus the "bisector" would fail to lie in the plane of the angle. These difficulties are not easy to repair, especially without using a dimension axiom; but

it is done as above, by using Gupta to find the midpoint of the base of an equilateral triangle, and then using inner Pasch. In other words, the above approach, which we use here in non-Markovian geometry, is (as far as I know) actually the simplest way to repair Euclid I.9. Of course, if we do not care about avoiding Markov's principle, some of the complications fall away.

7 Geometric arithmetic without Markov's principle

In [4, 3], we showed that in EG plus Markov's principle, one can define coordinates, and one can define addition and multiplication of points on a fixed line. The constructions differ from the classical ones because we need to avoid a case distinction according to the signs of the arguments. Two ways of defining addition and multiplication constructively are studied in the cited paper. One is purely geometrical, in which addition is defined using uniform rotation and reflection, and multiplication is defined using a circle, following Hilbert. We have checked that Markov's principle is not used in these constructions and proofs in any essential way. Actually, we found only one place where it was used, and we fixed that by providing in this paper a non-Markovian proof of Lemma 12. In the discussion below, we will point out how that lemma is used.

7.1 Coordinates

Once we have the uniform perpendicular, it is possible to define coordinates. Pick two perpendicular lines to serve as the x-axis and the y-axis. Then to each point p we can assign the feet a and b of the perpendiculars through p to the x-axis and y-axis, respectively. Then a is the x-coordinate of p. To get the y-coordinate, however, we must rotate b by $90°$, because the y-coordinate is supposed to be a point on the x-axis, the same distance from the origin as b. Therefore we need to define uniform rotation; "uniform" in the sense that it works without a case distinction as to the sign of b. A rotation by $90°$ is equivalent to two reflections in the line $y = x$; so it suffices to define uniform reflection.

The final step in the construction of coordinates is to show that, given points x and y on the x-axis, we can construct a point $P(x, y)$ whose coordinates are (x, y). Namely, erect a perpendicular J to the x-axis at x on the opposite side of the x-axis from $(0, -1)$. (The theorem about erecting perpendiculars requires a point not on the line, and then the perpendicular is constructed on the opposite side of the line from that point. This guarantees that J lies in the plane of the two axes. Rotate y counterclockwise about the origin by $90°$ (in the plane of the two axes, i.e., using the two axes to define the angle of rotation) to a point q. Drop a perpendicular from q to J. The foot of that perpendicular is $P(x, y)$. But why are the coordinates of $P(x, y)$ equal to (x, y)? Let $p = P(x, y)$ and let 0 be the origin. By construction the quadrilateral $qpx0$ has right angles at p,x, and 0; it is therefore a Lambert quadrilateral (three right

angles). It lies in the plane determined by the two axes since the points p and q are on the opposite side of the x-axis from $(0, -1)$.

Using the parallel postulate, we prove every Lambert quadrilateral in a plane is a rectangle. Hence the angle at q is also a right angle; hence q is the foot of the perpendicular from p to the y-axis; hence y is the y-coordinate of p, as claimed.

Thus the construction of coordinates depends on the theorem that every planar Lambert quadrilateral is a rectangle (a quadrilateral with four right angles). We claim that this theorem can be proved, without using Markov's principle, from the weakest form of the parallel postulate (Playfair's axiom). It is mentioned in Lemma 5.24 of [4], but the proof is not given there. In [3], Lemma 10.2, it is proved in intuitionistic Tarski geometry that a planar Saccheri quadrilateral is a rectangle. (A Saccheri quadrilateral is a quadrilateral $abcd$ with right angles at b and c and $ab = cd$.) Half a Saccheri quadrilateral is a Lambert quadrilateral, and a Lambert quadrilateral can be doubled to make a Saccheri quadrilateral, so the cited lemma implies that a planar Lambert triangle is a rectangle, as desired. The proof of that theorem given in [3] is almost free of Markov's principle, but it does appeal to Lemma 12 (in the numbering of the present paper), which in the cited paper is proved using Markov's principle. However, in this paper we proved Lemma 12 without using Markov's principle. Hence, the uniqueness of coordinates is proved without using Markov's principle.

7.2 Euclidean fields without Markov's principle

In [4], we gave axioms for (intuitionistic) Euclidean fields. The language has constants 0 and 1, function symbols $+$, $-$, $\cdot$, $1/x$ or x^{-1} for multiplicative inverse, and $\sqrt{\ }$, and a predicate $P(x)$ for "x is positive." The axioms are the field axioms, the axioms for order (involving P), and the axiom that non-negative elements have square roots. In addition we took Markov's principle $\neg\neg P(x) \to P(x)$ as an axiom.

Now we want to drop Markov's principle. When Markov's principle is allowed, non-zero elements have multiplicative inverses if and only if positive elements have inverses, since $1/x = x/|x|$, where $|x| = \sqrt{x^2}$. This argument works if $|x|$ is positive when x is nonzero, but that is equivalent to Markov's principle. In the absence of Markov's principle, then, the axiom that positive elements have multiplicative inverses does not imply that all nonzero elements have multiplicative inverses, and we leave that axiom unmodified.

Models of the resulting theory will be called "weakly Euclidean fields", since we already used "Euclidean fields" for models that satisfy Markov's principle. To be precise, the axioms are the usual axioms of commutative ring theory, plus the following:

$$\neg x \neq y \to x = y \qquad \text{stability of equality}$$
$$0 \neq 1 \qquad \text{EF0}$$
$$P(x) \to \exists y\, (x \cdot y = 1 \wedge P(y)) \qquad \text{EF1}$$
$$P(x) \wedge P(y) \to P(x + y) \wedge P(x \cdot y) \qquad \text{EF2}$$

$$x + y = 0 \rightarrow \neg(P(x) \wedge P(y)) \qquad \text{EF3}$$
$$x + y = 0 \wedge \neg P(x) \wedge \neg P(y) \rightarrow x = 0 \qquad \text{EF4}$$
$$x + y = 0 \wedge \neg P(y) \rightarrow \exists z(z \cdot z = x) \qquad \text{EF5}$$

Equivalently one may express EF1 and EF5 using the function symbols $1/x$ and $\sqrt{x}$, achieving thereby a quantifier-free axiomatization; and if one does not like using a function symbol for a not-everywhere-defined function, one may use the Logic of Partial Terms (LPT), as explained in [3]. Note that EF1 requires that the inverse positive element not only exist but also be positive itself. If Markov's principle is allowed, one can prove that the inverse of a positive element is positive, but since we are not assuming Markov's principle, that cannot (as it turns out) be proved, and we assume it as part of EF1.

Let $\mathbb{F}$ be an ordered field. Then "the plane over $\mathbb{F}$", denoted by $\mathbb{F}^2$, is a geometrical structure determined by defining relations of betweenness and equidistance in $\mathbb{F}^2$, using the given order of $\mathbb{F}$. Namely, b is between a and c if it lies on the interior of the segment ac. That relation can be expressed more formally in various ways, for example (using the cross product) by $(c - b) \times (b - a) = 0$ and $(b - c) \cdot (a - b) > 0$.

The equidistance relation $E(a, b, c, d)$, which means that segment ab is congruent to segment cd, can be defined by $(b - a)^2 = (d - c)^2$. Note that no square roots were used, so these definitions are valid in any ordered field.

The cross product $u \times v$ is defined using components as usual. If the lines through uv and st intersect (and do not coincide), then using Cramer's rule, the formula for the intersection point has denominator given by the cross product $D = (t - s) \times (v - u)$. (Details are in §9.3 of [4].) So, in verifying inner Pasch and outer Pasch, we will need to know that the cross products for the lines that need to intersect are positive. Thus what has to be proved is the following lemma:

Lemma 18 *In $\mathbb{F}^2$, the cross product (in the sense defined above) of the vectors giving the sides of angle abc with $0 < abc < \pi$ is positive. In particular, if d lies on ab, then $|d - c| > 0$, and if ab lies on the x-axis, then $|c_2| > 0$.*

Remark. At the end, the proof requires the modified EF1, which says that inverses of positive elements are positive.

Proof. By a rotation and translation, we can assume that $b = (0, 0)$ and a lies on the positive x-axis. (All three cases in the definition of "positive angle" are invariant under rotation and translation.) Then the cross product in question is $a \times c = c_2 a_1$. Since $a_1 > 0$, the cross product has positive absolute value if and only if $|c_2| > 0$. The penultimate statement follows from $c_2 > 0$, since $|d - c|$ is the hypotenuse a right triangle with a vertical leg of length c_2, and in $\mathbb{F}^2$ the theorem that a leg of a right triangle is shorter than the hypotenuse holds.

Assume $0 < abc < \pi$. If angle abc is a right angle or an angle of a right triangle, $|v_2| > 0$ is immediate. So we may assume abc is an apex angle. Then there are distinct points $u = (u_1, 0)$ and v on the ray from $(0, 0)$ through c, such that $u_1 > 0$ and $|u| = |v|$. We have $|u - v| > 0$ since

50

u and v are distinct. The cross product of the two sides of the angle is $u \times v = u_1 v_2$, since $u_2 = 0$. We must show $|u_1 v_2| > 0$. Since $u_1 > 0$, that is equivalent to $v_2 > 0$. We have not yet used the assumption $abc < \pi$, which we will now need. Let $d = (-a_1, 0)$. Then $0 < dbc$, by definition of $abc < \pi$. There are three cases: dbc is an apex angle, a right angle, or an angle of a right triangle. If either of the latter two then $c_2 > 0$, so $v_2 > 0$ and we are done. So we may assume dbc is an apex angle. Then there are distinct points p and q on the negative x-axis and bc, respectively, such that $|p| = |q|$ and $|p| > 0$. Extending the positive-length vectors u and v by $|p|$ and the vectors p and q by $|u|$, we can assume without loss of generality that $|u| = |p| = |v| = |q$. Then $v = q$ and angle $pqu = pvu$ is inscribed in the semicircle pvu with center at $b = (0, 0)$. Since classically $\mathbb{F}^2$ satisfies the theorem that an angle inscribed in a semicircle is a right angle, and being a right angle is equationally described, pvu is a right angle. $|p_2|$ is the altitude of that right triangle. By Lemma 18, $\mathbb{F}^2$ satisfies that the altitude of a right triangle (with positive legs) is positive. Hence $|p_2| > 0$. But $c_2/|c| = p_2/|p|$, and $1/|c|$ and $1/|p|$ are both positive by EF1, since $|c| > 0$. Hence c_2 is a positive quantity times p_2. Hence $|c_2| > 0$, as desired. That completes the proof of the lemma.

The following theorem verifies that we have the axioms for weakly Euclidean fields right.

Theorem 11 *Every plane $\mathbb{F}^2$ over a weakly Euclidean field $\mathbb{F}$ is a model of Euclidean geometry EG.*

Remark. This theorem depends on the modifications we made to the extension axiom and to inner and outer Pasch; had we not changed the formulations of inner and outer Pasch to require an angle to be positive and have a positive supplement, we would need Markov's principle to verify them. Although these calculations are elementary, we need to verify that Markov's principle is not necessary, and that we only need to divide by elements that are known to be positive, and we need square roots only of positive elements (as opposed to nonzero elements). The fact that these calculations succeed without Markov is a strong indication that we have the axioms of EG right.

Proof. We define $x > y$ to abbreviate $P(y - x)$ and $x \geq y$ to abbreviate $\neg x < y$, and similarly for $<$ and $\leq$. We prove the transitivity of $<$. Suppose $x < y$ and $y < z$. Then $P(y - x) \wedge P(z - y)$. By EF2, $P((y - x) + (z - y))$. Hence $P(z - x)$. Hence $x < z$.

We define absolute value in $\mathbb{F}^2$:

$$|(x, y)| := \sqrt{x^2 + y^2}.$$

We have $x^2 \geq 0$ for all x (for example, since $\leq$ is stable, we can prove that by contradiction). Hence $|u|$ is always defined.

The interpretation of betweenness in $\mathbb{F}^2$ has to be given without using cases. Collinearity $L(u, v, w)$ can be defined by

$$L(u, v, w) := (w - u) \times (w - v) = 0.$$

Then betweenness is defined by

$$\mathbf{B}(u, v, w) := L(u, v, w) \wedge |v - u| + |w - v| = |w - u| \wedge |v - u| > 0 \wedge |w - v| > 0.$$

Since translations and rotations are given by linear maps with determinant one, they preserve cross and dot products, and hence they preserve distance and collinearity, and hence they preserve betweenness as well.

We claim that two points u and v in $\mathbb{F}^2$ are distinct if and only if $|v - u| > 0$. According to the definition of $u \# v$, u and v are distinct if there is a point x bearing any of the three possible betweenness relations with u and v. For example, suppose $\mathbf{B}(u, x, v)$. Then by the definition of betweenness, $|x - u| > 0$ and $|v - x| > 0$ and $L(u, x, v)$. We must show $|v - u| > 0$. By a rotation and translation, we may assume $u = (0, 0)$ and $x = (x_1, 0)$ and $y = (y_1, 0)$. Then we have $x_1 > 0$ and $y_1 > x_1$, and we must show $y_1 > 0$; but that is the transitivity of $>$, which we proved above. Similarly for the other two possible betweenness relations. Thus $u \# v$ if and only if $|v - u| > 0$.

We turn to the betweenness axioms. Consider the axiom $\neg \mathbf{B}(a, b, a)$. In $\mathbb{F}^2$, we have $\mathbf{B}(a, b, c) \rightarrow L(a, b, c)$, by the definition of $\mathbf{B}(a, b, c)$. Suppose $\mathbf{B}(a, b, a)$. Then $L(a, b, a)$. By a rotation and translation, we can assume a and b lie on the x-axis. Then $\mathbf{B}(a, b, a)$ becomes $|b_2 - a_2| + |a_2 - b_2| = |a_2 - a_2|$. But the last expression is 0. With $x = b_2 - a_2$, we have $2|x| = 0$. Hence $x = 0$; that is, $b_2 = a_2$. Hence $a = b$. But that contradicts $\mathbf{B}(a, b, a)$. Hence our assumption $\mathbf{B}(a, b, a)$ is contradictory. That is, $\neg \mathbf{B}(a, b, a)$, which is what had to be proved.

Consider the symmetry of betweenness, $\mathbf{B}(a, b, c) \rightarrow \mathbf{B}(c, b, a)$. Suppose $\mathbf{B}(a, b, c)$; after a rotation and translation, we can assume a, b, and c lie on the x-axis. For simplicity we drop subscripts, writing just a instead of a_1, etc. Then we have $|c - b| + |b - a| = |c - a|$ and $|c - b| > 0$ and $|b - a| > 0$. What has to be proved is an equation representing non-strict betweenness $\mathbf{T}(c, b, a)$ and the two inequalities $|a - b| > 0$ and $|b - c| > 0$. By the stability of equality, we can argue classically for the non-strict betweenness, which we take as proved. The two inequalities follow from the easy lemma $|-x| = |x|$. That completes the verification of the symmetry of betweenness.

Consider the inner transitivity of betweenness, axiom A15-i:

$$\mathbf{B}(a, b, d) \wedge \mathbf{B}(b, c, d) \rightarrow \mathbf{B}(a, b, c).$$

Suppose the hypothesis. Again we can apply a rotation and translation to bring all the points to the x-axis. As before the non-strict inequality $\mathbf{T}(a, b, c)$ is expressed by an equation, so by the stability of equality, we can prove it classically, which we assume done. For simplicity we drop the subscripts, writing just a instead of a_1, etc. Then $a \le b \le c \le d$. We have $a < b < d$ and $b < c < d$ by hypothesis. We need to prove $a < b < c$. But we have both inequalities already; there is no existential assertion to prove. That completes the verification of the inner transitivity axiom.

Next we turn to the inner Pasch axiom. Please refer to Fig. 8. By a rotation and translation, and possibly renaming some of the points, we may assume aq lies on the x-axis and $c_2 \ge 0$ and $q_1 \ge a_1$. We have to show that $|(q - a) \times (b - p)| > 0$. That is, $|(b_2 - p_2)(q_1 - a_1)| > 0$. The hypothesis of inner Pasch includes the specification that either angle bca or angle pqb is between 0 and π. First assume $0 < bca < \pi$. By Lemma 18, applied to angle bca, any point on bc is at a positive distance

from a. Hence $|a - q| > 0$. Hence $q_1 - a_1 > 0$. Therefore by EF1, $1/(q_1 - a_1) > 0$. Hence it suffices to show $|b_2 - p_2| > 0$. Because $\mathbf{B}(c, q, b)$ holds by hypothesis, we have $|c - q| > 0$, so acq is an angle, and the same angle as acb. Since being positive respects the same-angle relation, and even angle congruence, $0 < acq < \pi$. Then by Lemma 18, c lies at positive distance from aq. That is, $c_2 > 0$. Therefore $b_2 < 0$, since $\mathbf{B}(b, q, c)$ and $q_2 = 0$. Because $\mathbf{B}(a, p, c)$ holds by hypothesis and $c_2 > 0$ and $a_2 = 0$, we have $p_2 > 0$. Now we have $b_2 < 0 < p_2$. Hence $|b_2 - p_2| > 0$, as desired. That completes the verification of inner Pasch, in the case $0 < bac < \pi$. We note that it would not work without Markov's principle if we had only $\mathbf{T}(c, q, b)$ for a hypothesis instead of $\mathbf{B}(c, q, b)$.

It remains to verify inner Pasch under the assumption $0 < qpa < \pi$ instead of $0 < bca < \pi$. Then Lemma 18 tells us that $|(q - p) \times (a - p)| > 0$. We are free to choose the origin of coordinates so that $p_1 = 0$ and $a_2 = q_2 = 0$. Then

$$
\begin{aligned}
|(q_1 - p_1)(a_2 - p_2) - (q_2 - p_2)(a_1 - p_1)| &> 0 \\
|-q_1 p_2 - (-p_2)(a_1)| &> 0 \\
|p_2(a_1 - q_1)| &> 0.
\end{aligned}
$$

By Lemma 18, $|a_1 - q_1)| > 0$. Hence $p_2 > 0$. Since $\mathbf{B}(a, p, c)$, that implies $c_2 > 0$. That in turn implies $|(a - c) \times (q - c)| > 0$, which in turn implies $|(a - c) \times (q - \ell)| > 0$, for any point ℓ on the ray cq; in particular for the point ℓ such that $|\ell - c| = |a - c|$. Hence $0 < acb < \pi$. But we have already verified that under that hypothesis, inner Pasch holds, so that completes the verification of inner Pasch.

Now we turn to outer Pasch. We may assume $a = (0, 0)$ and $q = (q_1, 0)$, so aq lies on the x-axis. As with inner Pasch, the crucial issue is to verify that the lines that need to intersect have positive cross product. For outer Pasch that is $|(q - a) \times (b - p)|$, which is $|q_1 - a_1||b_2 - p_2|$. Since $0 < baq < \pi \vee 0 < abq < \pi$, q and a are distinct points, so $q_1 > a_1$, and $b_2 > 0$. Hence $|q_1 - a_1| > 0$, and to finish the verification it suffices to prove $|b_2 - p_2| > 0$. Since $\mathbf{B}(b, c, q)$, we have $0 < c_2 < b_2$, and since $\mathbf{B}(a, p, c)$ we have $0 < p_2 < c_2$. Hence $0 < p_2 < b_2$. That completes the verification of outer Pasch.

Any axiom whose analytic-geometry interpretation does not involve a positive occurrence $P(x)$ will be automatically verifiable, by Gödel's double-negation interpretation. That is, starting with the classical verification, we just push double negations inwards, and when they reach an equality, we drop them by the stability of equality, and the result is an intuitionistic verification. If the statement had a positive occurrence of $P(x)$ we would need Markov's principle to drop the double negation, but a negative occurrence causes no problem. If the statement does not involve $P(x)$, the need does not arise. On these grounds, the 5-segment axiom, line–circle continuity and circle–circle continuity (which have occurrences of $P(x)$, counting as negative by virtue of being in the hypotheses) hold in $\mathbb{F}^2$. (See [4] for a more detailed discussion.)

The dimension axioms (once properly formulated) are easily verified in $\mathbb{F}^2$. That completes the proof of the theorem.

8 Geometric arithmetic

We have already discussed the construction of coordinates relative to a pair of perpendicular lines chosen as axes. The definitions of multiplication and addition for positive arguments then offers no difficulty. For example, the multiplicative inverse of a point $(a, 0)$ on the x-axis is found by erecting a perpendicular to the x-axis at the point 1, and using Euclid 5 to prove that it meets the line containing the line through $(0, 0)$ and $(a, 1)$. That point has coordinates $(1, 1/a)$. To use the non-Markovian version of Euclid 5, we need the hypothesis that point $(a, 0)$ is distinct from $(0, 0)$. The passage from signed to unsigned multiplication is discussed in [4], where two solutions are given. For example, the solution by using Hilbert's definition of multiplication using a circle works in EG (but it does require the non-strict version of line–circle continuity, which we included partly for that reason). It follows that, loosely speaking, "every model of EG is a plane over a Euclidean field."

9 Some results about EG

In this section we present some metamathematical results about non-Markovian geometry EG.

9.1 A Kripke model of non-Markovian geometry

We want to demonstrate that non-Markovian geometry is "about something" by giving a model of it in which Markov's principle fails. Of course, we cannot give a model in the classical sense, since the double negation of Markov's principle is a theorem. But we can give a two-node Kripke model. (For an introduction to Kripke models and a proof of the completeness theorem, see [29], Part V, pp. 324*ff.* Since our model is particularly simple, you may be able to understand it without a background in Kripke models.)

In view of the fact that non-Markovian geometry is interpretable in the theory of Euclidean fields (without Markov's principle), it suffices to give a Kripke model that does not satisfy Markov's principle of Euclidean field theory in the form $\neg\neg P(x) \to P(x)$. Then a corresponding model of EG can be constructed by replacing each field $\mathbb{F}_\alpha$ in the first Kripke model by the plane $\mathbb{F}_\alpha^2$ over $\mathbb{F}_\alpha$, with betweenness and congruence interpreted as in the previous section.

Theorem 12 *Euclidean field theory does not imply Markov's principle, and geometry EG does not imply Markov's principle.*

Proof. The second claim follows from the first and Theorem 11. We will construct a Kripke model of Euclidean field theory in which MP fails. Let $\mathbb{F}$ be any non-Archimedean field (so it contains some "infinite" element, larger than any positive integer). Let $\mathbb{F}_0$ be the set of "finitely bounded" elements of $\mathbb{F}$, i.e., those elements bounded by some positive integer. $\mathbb{F}_0$ contains "infinitesimal" elements x, by definition those whose reciprocals are not finitely bounded. We interpret $P(x)$ in $\mathbb{F}_0$ to mean that x is

positive in $\mathbb{F}$ and not infinitesimal. Then $\mathbb{F}_0$ is not a classical model of Euclidean ordered field theory, but it is still admissible as a node in a Kripke model.

The definition of a Kripke model includes these two clauses: $\mathcal{M}_\alpha \models \neg\phi$ means that ϕ fails everywhere above α, and $\mathcal{M}_\alpha \models A \to B$ means that if $\mathcal{M}_\beta \models A$ for $\beta > \alpha$, then $\mathcal{M}_\beta \models B$. The soundness theorem (easy to check) for Kripke models says that the satisfaction relation preserves intuitionistic logical consequence. Hence, to prove that Markov's principle is not a theorem of Euclidean field theory, it suffices to exhibit a Kripke model in which Markov's principle is not satisfied. That Kripke model, call it $\mathcal{M}$, will be a two-node model. The root node $\mathcal{M}_0$ has universe $\mathbb{F}_0$, but $P(x)$ is interpreted in $\mathcal{M}_0$ to mean that x is positive in $\mathbb{F}$ and not infinitesimal. The other node, lying "above" $\mathcal{M}_0$, is $\mathcal{M}_1$, which is the classical model $\mathbb{F}$.

We now show that Markov's principle fails in $\mathcal{M}$. Let x be infinitesimal in $\mathbb{F}$. Then the root node $\mathcal{M}_0$ satisfies $\neg\neg P(x)$, since x is positive in $\mathbb{F}$, and satisfaction in $\mathcal{M}_1$ is just classical satisfaction. Suppose, for proof by contradiction, that $\mathcal{M}$ satisfies Markov's principle. Then the root node satisfies the hypothesis $\neg\neg P(x)$ of Markov's principle, so it must satisfy the conclusion $P(x)$. But, by definition of P in $\mathcal{M}_0$, that implies x is not infinitesimal, contradiction. Hence Markov's principle does not hold in $\mathcal{M}$.

We still need to check that the axioms of Euclidean field theory hold in $\mathcal{M}$. Since satisfaction at node $\mathcal{M}_1$ is just classical satisfaction, we only need to check $\mathcal{M}_0$.

Consider EF0: $\neg x \neq y \to x = y$. Suppose $\mathcal{M}_0$ satisfies $\neg x \neq y$. Then $\mathcal{M}_1$ does not satisfy $x \neq y$. Hence $x = y$ in $\mathbb{F}$. Hence $x = y$ in $\mathbb{F}_0$. Hence $\mathcal{M}$ satisfies EF0.

Consider EF1: $P(x) \to \exists y\, (x \cdot y = 1)$. Suppose $\mathcal{M}_0 \models P(x)$. Then x is positive and not infinitesimal. Hence $y = 1/x$ in $\mathbb{F}$ is finitely bounded, so it belongs to $\mathbb{F}_0$. Hence x has an inverse in $\mathbb{F}_0$. EF1 also requires that inverse to be positive, which it is, since it is positive in $\mathbb{F}$. Therefore $\mathcal{M}_0$ satisfies EF1.

Consider EF2: $P(x) \wedge P(y) \to P(x+y) \wedge P(x \cdot y)$. Suppose $\mathcal{M}_0 \models P(x) \wedge P(y)$. Then x and y are positive and not infinitesimal. Then their sum and product are positive. Since x is not infinitesimal, by definition $1/x$ is finitely bounded. Let n be an integer such that $1/x < n$. Then

$$\frac{1}{x+y} < \frac{1}{x} < n.$$

Hence $1/(x+y)$ is finitely bounded. Hence $x+y$ is positive in $\mathcal{M}_0$. Also

$$\frac{1}{xy} = \frac{1}{x}\frac{1}{y} < nm$$

provided $1/y < m$. Hence xy is also positive in $\mathcal{M}_0$. Hence $\mathcal{M}_0$ satisfies EF2.

Consider EF3: $x + y = 0 \to \neg(P(x) \wedge P(y))$. Suppose $x+y = 0$ holds in $\mathcal{M}_0$ and $P(x) \wedge P(y)$ holds in $\mathcal{M}_0$ or in $\mathcal{M}_1$. Then x and y are positive in $\mathbb{F}$, contradicting the fact that $\mathbb{F}$ satisfies EF3. Hence $\mathcal{M}$ satisfies EF3.

Consider EF4: $x + y = 0 \land \neg P(x) \land \neg P(y) \to x = 0$. Suppose $\mathcal{M}_0 \models \neg P(x)$. Then $\mathcal{M}_1$ does not satisfy $P(x)$, so x is not positive in $\mathbb{F}$. If the same is true of $y = -x$, then since $\mathbb{F}$ classically satisfies EF4, $x = 0$. Hence $\mathcal{M}_0$ satisfies EF4.

Consider EF5, namely $x + y = 0 \land \neg P(y) \to \exists z(z \cdot z = x)$. Suppose $\mathcal{M}_0$ satisfies $x + y = 0 \land \neg P(y)$. Then $y = -x$ is not positive in $\mathbb{F}$, so by EF5 in $\mathbb{F}$, x has a square root z in $\mathbb{F}$. We claim z is finitely bounded, and hence lies in $\mathbb{F}_0$. If $|z| < 1$ we are done. Otherwise, let n be an integer greater that $|x|$; then $|z| \leq \sqrt{n} < n + 1$. That completes the proof of the theorem.

Remark. In the corresponding geometrical Kripke model, whose two nodes are $\mathcal{M}_0 = F_0^2$ and $\mathcal{M}_1 = F_1^2$, with betweenness and congruence defined in the obvious way (details in [4]), two points of $\mathcal{M}_0$ are distinct if and only if the difference of both their coordinates is infinitesimal, and an angle is positive if and only if the rays that form its sides contain points at equal distances from the vertex that are distinct; loosely speaking, the angle is not infinitesimal. This model is in accord with the intuition that if Markov's principle does not hold, one has to imagine that every point has a "cloud" of nearby points, infinitesimally close to it but not equal; so near that the two points cannot determine a line.

Remark. This model, and another more complicated model, were given in [4], to show that Markov's principle does not follow from EF0–EF5. But it was not appreciated at the time that Markov's principle *does* follow from the axioms of constructive geometry, unless the extension axiom and Pasch's axiom(s) are modified by requiring distinct points and positive angles; geometry without Markov's principle was purposely not dealt with in [4], since "sufficient to the day are the tribulations thereof." In other words, the difficulties of non-Markovian geometry then seemed intimidating, and were postponed.

9.2   EG is not too weak

We created the theory EG by starting with constructive geometry from [4], removing Markov's principle, and modifying the segment extension and Pasch axioms. We now show that we have not made EG too weak, in the sense that if we add Markov's principle back to EG, we also automatically get back the unmodified axioms of constructive geometry.

Lemma 19 *EG plus Markov's principle proves that any two unequal points are distinct, and if a, b, c are not collinear, then $0 < abc < \pi$.*

Proof. Let $a \neq b$. Let $\alpha \# \beta$ (two such points exist by the lower dimension axiom of EG). Extend $\alpha\beta$ by ab to point e. Then $T(\alpha, \beta, e)$. Since $a \neq b$ and $ab = \beta e$, we have $\beta \neq e$. With Markov's principle, we have $\mathbf{B}(x, y, z)$ equivalent to $\mathbf{T}(x, y, z) \land x \neq y \land y \neq z$. Hence, with Markov's principle, $\mathbf{B}(\alpha, \beta, e)$. Hence $\beta \# e$. Then ab is congruent to the positive segment βe. Hence it can be extended by cd to a new point f, using the segment extension axiom of EG. Then $\mathbf{B}(a, b, f)$. Then by definition $a \# b$.

Now suppose $\neg L(a, b, c)$. We will prove $0 < abc < \pi$. Since $\neg L(a, b, c)$, we have $a \neq b$ and $b \neq c$, and using Lemma 4, we can assume without

loss of generality that $ab = bc$. Since $\neg L(a, b, c)$ we have $a \neq c$. As shown above, Markov's principle then implies $a \# c$, so $0 < abc$. Let d be the reflection of c in b; then similarly, $0 < abd$, so $abc < \pi$. That completes the proof of the lemma.

Theorem 13 *EG plus Markov's principle proves the axioms of constructive Tarski geometry given in [4].*

Proof. We will prove the segment extension axiom A4, which says that if $a \neq b$ then ab can be extended by any segment cd. By Lemma reflemma:nottooweak, with Markov's principle $a \neq b$ implies $a \# b$, so the segment extension axiom of EG applies. That completes the proof of axiom A4.

Next, we will verify the (unmodified) inner Pasch axiom. In constructive geometry, the hypotheses include $\neg L(a, b, c)$. By Lemma 19, $0 < abc < \pi$. Then the inner Pasch axiom of EG can be applied. That completes the verification of inner Pasch.

The dimension axiom of EG immediately implies the dimension axiom of constructive geometry, and no other axioms were modified. That completes the proof of the theorem.

9.3 EG is strong enough

In this section we advance the following principle

> If, in any theorem of Euclidean geometry of the type found in Books I-IV, we replace "a and b are unequal" by "a and b are distinct", and we require that all angles be positive, then the resulting theorem is provable in non-Markovian geometry EG.

The first difficulty here is "of the type found in Books I-IV." The principle surely fails if the theorem in question is Markov's principle itself. But Markov's principle is not a proposition of the kind found in Euclid. As discussed above, those propositions are "Horn formulas"; that is, their hypotheses are atomic statements (possibly involving defined concepts) and some given points, and the conclusions assert the existence of some points bearing positive relations to the original points.

A second difficulty is that the language of EG does not mention "angles" directly, so what exactly is meant by "require that all angles be positive"? The phrase "angle abc" means (in ordinary geometry) that the three points are distinct and not collinear. We call a formula ϕ "of Euclidean form" if it is expressed as a Horn formula in a language extending EG by a 3-ary predicate *Angle* and a 2-ary predicate D, where D occurs only negatively (that is, in the antecedent of the implication). Then the "ordinary interpretation" of ϕ is obtained by replacing $D(a, b)$ by $a \neq b$ and $Angle(a, b, c)$ by $\neg L(a, b, c)$. The "non-Markovian interpretation" is obtained by replacing $D(a, b)$ by $a \# b$ and $Angle(a, b, c)$ by $0 < abc < \pi$, both of which are abbreviations for existential formulas of EG. The existential quantifiers in the antecedent are then pulled out, so that the non-Markovian interpretation of ϕ has additional free (or universally quantified) variables. Now we are able to state a theorem:

Theorem 14 *Let ϕ be a formula of Euclidean form whose ordinary interpretation is provable in constructive geometry. Then the non-Markovian interpretation of ϕ is provable in EG.*

Remark. By Theorem 13, "provable in constructive geometry" is equivalent to "provable in EG + MP."

Proof. The non-Markovian interpretation of any Euclidean formula is a Horn formula, so it suffices to prove that Markov's principle is conservative over EG for Horn theorems. Let ϕ be a Horn formula provable from a list of axioms Γ and formulas $\neg\neg P \to P$, where P is atomic (Markov's principle is of that form). We write one instance of Markov's principle explicitly and suppose Γ may contain more instances. We will show how to eliminate one instance at a time. By cut-elimination, there is a cut-free derivation in Gentzen's sequent calculus G3 [16] of

$$\Gamma, \neg\neg P \to P \Rightarrow \phi.$$

Since ϕ and Γ do not contain quantifiers or disjunction, by [17], the inferences may be permuted so that the implication on the left is the last-introduced connective, introduced by rule $\to\Rightarrow$. Then there are cut-free derivations of $P \Rightarrow \phi$ and of $\Gamma \Rightarrow \neg\neg P$. But again we can permute the inferences so that the last inference introduces $\neg\neg P$. Hence there is a cut-free derivation of $\Gamma, \neg P \Rightarrow \bot$. Permuting again if necessary, there is a cut-free derivation of $\Gamma \Rightarrow P$. Recall that we have a cut-free derivation of $P \Rightarrow \phi$. Applying the cut rule to these two derivations, we obtain a derivation (with one cut) of $\Gamma \Rightarrow \phi$. Hence we have eliminated one instance of $\neq \neg P \to P$. Applying cut elimination, we now have a cut-free proof of $\Gamma \Rightarrow \phi$, so we have eliminated one instance of Markov's principle. Continuing in this fashion, or more formally, by induction on the number of instances of Markov's principle in the original proof, we eventually eliminate all instances of Markov's principle. That completes the proof of the theorem.

In [4] we concluded, by a combination of geometry and metamathematics, that the essential content of Euclid Books I–III is formalizable in constructive geometry. Now, the above metatheorem allows us to conclude that Euclid Books I–III are provable without Markov's principle, if the hypotheses of unequal points and nonzero angles are strengthened as above to distinct points and positive angles.

10   Conclusions

Much of Brouwer's work was focused on the nature of the continuum. Although most, but not all, of his work preceded the development of recursive analysis and recursive realizability (which he never mentioned, even in later papers), he was clearly aware that "lawlike" or computable real numbers are not enough to "fill out" the geometric continuum; therefore choice sequences or even lawless sequences were needed. Brouwer pointed out the non-constructive nature of the "intersection theorem", and by implication, the classical view of elementary geometry.

Our point in this paper is that Brouwer's criticisms of "elementary geometry" definitely do not apply to geometry as Euclid wrote it. We have shown that there is a coherent and beautiful theory of intuitionistic Euclidean geometry that neither assumes nor implies Markov's principle.

The consistency and coherence of this theory of non-Markovian Euclidean geometry show that Euclidean geometry, the law of the excluded middle, and Markov's principle are three separate issues. One can choose to accept Markov's principle (but not the law of the excluded middle), in which case one gets a very nice constructive geometry [3, 4]; or one can choose not to accept Markov's principle, in which case one is forced to make some finer distinctions, but still gets a coherent intuitionistic geometry, as shown here.

References

[1] J. AVIGAD, E. DEAN, AND J. MUMMA, *A formal system for Euclid's Elements*, Review of Symbolic Logic, 2 (2009), pp. 700–768.

[2] M. BEESON, *Constructive geometry*, in Proceedings of the Tenth Asian Logic Colloquium, Kobe, Japan, 2008, T. Arai, J. Brendle, C. T. Chong, R. Downey, Q. Feng, H. Kikyou, and H. Ono, eds., Singapore, 2010, World Scientific, pp. 19–84.

[3] ——, *A constructive version of Tarski's geometry*, Annals of Pure and Applied Logic, 166 (2015), pp. 1199–1273.

[4] ——, *Constructive geometry and the parallel postulate*, Bulletin of Symbolic Logic, 22 (2016), pp. 1–104.

[5] M. BEESON AND L. WOS, *Finding proofs in Tarskian geometry*, Journal of Automated Reasoning, accepted (2016).

[6] L. E. J. BROUWER, *Essentieel negatieve eigenschappen*, KNAW Proc., 51 (1948), pp. 964–964. Translated as *Essentially Negative Properties*, 1948A in [14], pp. 478-840.

[7] ——, *Contradictoriteit der elementaire meetkunde*, KNAW Proc., 52 (1949), pp. 315–316. Translated as *Contradictority of elementary geometry*, 1949B in [14], pp. 497-498.

[8] ——, *De non-aequivalentie van de constructieve en de negatieve orderrelatie in het continuum*, KNAW Proc., 52 (1949), pp. 122–124. Translated as *The non-equivalence of the constructive and the negative order relation on the continuum*, 1949A in [14], pp. 495-406.

[9] J. A. GELDOF, *Over de arithmetisering van axiomatische meetkunden*, PhD thesis, University of Amsterdam, 1951.

[10] M. J. GREENBERG, *Euclidean and non-Euclidean Geometries: Development and History*, W. H. Freeman, New York, fourth ed., 2008.

[11] H. N. GUPTA, *Contributions to the Axiomatic Foundations of Geometry*, PhD thesis, University of California, Berkeley, 1965.

[12] A. HEYTING, *Zur intuitionistischen Axiomatik der projectiven Geometrie*, Mathematische Annalen, 98 (1927), pp. 491–538.

[13] ———, *Axioms for intuitionistic plane affine geometry*, in The axiomatic method with special reference to geometry and physics, Proceedings of an international symposium, held at the University of California, Berkeley, December 26, 1957-January 4, 1958., L. Henkin, P. Suppes, and A. Tarksi, eds., vol. 27 of Studies in Logic and the Foundations of Mathematics, North-Holland, 1959, pp. 160–173.

[14] ———, ed., *L. E. J. Brouwer, Collected Works, volume 1: Philosophy and Foundations of Mathematics*, North-Holland, Amsterdam, New York and Oxford, 1975.

[15] D. HILBERT, *Foundations of Geometry (Grundlagen der Geometrie)*, Open Court, La Salle, Illinois, 1960. Second English edition, translated from the tenth German edition by Leo Unger. Original publication date, 1899.

[16] S. C. KLEENE, *Introduction to Metamathematics*, van Nostrand, Princeton, 1952.

[17] ———, *Permutability of inferences in Gentzen's calculi LK and LJ*, Memoirs of the American Mathematical Society, 10 (1952), pp. 1–26.

[18] S. C. KLEENE AND R. E. VESLEY, *The Foundations of Intuitionistic Mathematics*, North-Holland, Amsterdam, 1965.

[19] M. LOMBARD AND R. VESLEY, *A common axiom set for classical and intuitionistic plane geometry*, Annals of Pure and Applied Logic, 95 (1998), pp. 229–255.

[20] M. MANDELKERN, *Constructive coordinatization of Desarguesian planes*, Beiträge zur Algebra und Geometrie (Contributions to Algebra and Geometry), 48 (2007), pp. 547–589.

[21] ———, *The common point problem in constructive projective geometry*, Indagationes Mathematicae, 24 (2013), pp. 111–114.

[22] M. MANDELKERN, *A constructive real projective plane*, Journal of Geometry, 107 (2016), pp. 19–60.

[23] M. MANDELKERN, *Constructive projective extension of an incidence plane*, Transactions of the American Mathematical Society, 366 (2104), pp. 691–706.

[24] J. MOLLERUP, *Die Beweise der ebenen Geometrie ohne Benutzung der Gleichheit und Ungleichheit der Winkel*, Mathematische Annalen, 58 (1904), pp. 479–496.

[25] M. PASCH, *Vorlesung über Neuere Geometrie*, Teubner, Leipzig, 1882.

[26] W. SCHWABHÄUSER, W. SZMIELEW, AND A. TARSKI, *Metamathematische Methoden in der Geometrie: Teil I: Ein axiomatischer Aufbau der euklidischen Geometrie. Teil II: Metamathematische Betrachtungen (Hochschultext)*, Springer–Verlag, 1983. Reprinted 2012 by Ishi Press, with a new foreword by Michael Beeson.

[27] A. TARSKI AND S. GIVANT, *Tarski's system of geometry*, The Bulletin of Symbolic Logic, 5 (1999), pp. 175–214.

[28] A. Troelstra, *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis*, vol. 344 of Lecture Notes in Mathematics, Springer-Verlag, Berlin Heidelberg New York, 1973.

[29] ———, *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis*, no. 344 in Lecture Notes in Mathematics, Springer, Berlin, Heidelberg, New York, 1973.

[30] D. van Dalen, *Extension problems in intuitionistic plane projective geometry*, Indagationes Mathematicae, 25 (1963), pp. 349–383.

[31] ———, *Heyting and intuitionistic geometry*, in Mathematical Logic, P. Petrov, ed., Plenum Press, New York, 1990, pp. 19–27.

[32] ———, *"Outside" as a primitive notion in constructive projective geometry*, Geom. Dedicata, 60 (1996), pp. 107–111.

[33] G. Veronese, *Fondamenti di geometria a più dimensioni e a più specie di unità rettilinee esposti in forma elementare. Lezioni per la scuola di magistero in matematica*, Padova: Tipografia del Seminario. Digitized by Cornell University Library. Italian text available at `http://ebooks.library.cornell.edu/cgi/t/text/text-idx?c=math;idno=01460002`, 1891.

[34] J. von Plato, *The axioms of constructive geometry*, Annals of Pure and Applied Logic, (1995), pp. 169–200.