Constructive Geometry

Michael Beeson

July 30, 2009

Abstract

Euclidean geometry, as presented by Euclid, consists of straightedge-and-compass constructions and rigorous reasoning about the results of those constructions. A consideration of the relation of the Euclidean "constructions" to "constructive mathematics" leads to the development of a first-order theory **ECG** of the "Euclidean Constructive Geometry", which can serve as an axiomatization of Euclid rather close in spirit to the *Elements* of Euclid. **ECG** is axiomatized in a quantifier-free, disjunction-free way. Unlike previous intuitionistic geometries, it does not have apartness. Unlike previous algebraic theories of geometric constructions, it does not have a test-for-equality construction. We show that **ECG** is a good geometric theory, in the sense that with classical logic it is equivalent to textbook theories, and its models are (intuitionistically) planes over Euclidean fields. We then apply the methods of modern metamathematics to this theory, showing that if **ECG** proves an existential theorem, then the object proved to exist can be constructed from parameters, using the basic constructions of **ECG** (which correspond to the Euclidean straightedge-and-compass constructions). In particular, objects proved to exist in **ECG** depend continuously on parameters. We also study the formal relationships between several versions of Euclid's parallel postulate, and show that each corresponds to a natural axiom system for Euclidean fields.[1]

1 Introduction

Euclid's *Elements* [7], written down about 300 BCE, has been extraordinarily influential in the development of mathematics, and prior to the twentieth century was regarded as a paradigmatic example of pure reasoning. But during those 2300 years, most people thought that Euclid's theory was about something. What was it about? Some may have answered that it was about points, lines, and planes, and their relationships. Others may have said that it was about methods for constructing points, lines, and planes with certain specified relationships to given points, lines, and planes, for example, constructing an equilateral triangle with a given side. In these two answers, we see the viewpoints of pure (classical) mathematics and of algorithmic mathematics represented. Hilbert's 1899 reworking of the theory [10] gave another answer, surprising at the time: Euclid's theories were *not about anything at all.* Instead of "points, lines, and planes", one should be able to read "tables, chairs, and beer mugs." All the reasoning should still be valid. The names of the "entities" were just place holders. That was the viewpoint of twentieth-century axiomatics.

In the late twentieth century, contemporaneously with the flowering of computer science, there was a new surge of vigor in algorithmic, or constructive, mathematics, beginning with Bishop's book [4]. In algorithmic mathematics, one tries to reduce every "existence theorem" to an assertion that a certain algorithm has a certain result. In the terminology of computer

[1] We would like to thank Jeremy Avigad and Freek Wiedijk for productive and interesting discussions, and for careful reading and useful suggestions.

science, existence theorems should become correctness proofs of algorithms. The proof theory of arithmetic has provided many beautiful theorems to show that indeed, existence theorems in number theory (when constructively proved) contain algorithms that can be "extracted" from the proofs. In particular we mention the techniques of recursive realizability, the Dialectica interpretation of Gödel, and the extraction of algorithms from cut-free proofs as well-known examples of the phenomenon.

In this paper, we re-examine Euclidean geometry from the viewpoint of constructive mathematics. The phrase "constructive geometry" suggests, on the one hand, that "constructive" refers to geometrical constructions with straightedge and compass. On the other hand, the word "constructive" may suggest the use of intuitionistic logic. We investigate the connections between these two meanings of the word. Our method is to proceed by analogy with the extensive body of work that has been done on number theory and analysis, applying the relevant methodologies to the weaker theories of geometry. The basis for the work described here is the idea that in geometry, we can take "algorithm" in the restricted sense of "geometric construction." That is, we pursue the analogy

$$\frac{\text{formal number theory}}{\text{Turing computable functions}} = \frac{\text{intuitionistic geometry}}{\text{geometric constructions}}$$

To carry out this program, we need a suitable formal theory for intuitionistic geometry. It should be a theory with terms for the geometric constructions, so that there will be terms available to denote the means of constructing things that have been proved to exist. This leads us to look for a quantifier-free axiomatization. In formulating a suitable theory of geometry, another important consideration was the principle that "constructive proof implies continuity in parameters". This principle is familiar to those who have studied constructive mathematics; but it is easy to understand on an intuitive basis. If we wish to allow an interpretation of geometry in which points are given by approximations (as accurate as one may demand), for example, if they are given as pairs of real numbers (x, y), and we (constructively) prove $\forall x \exists y\, A(x, y)$, then we ought to be able to produce approximations to y when we are given approximations to x. This idea led us to look at the continuity of the Euclidean constructions. We looked for (and found) an axiom system **ECG** with intuitionistic logic, such that when an existential theorem is proved in **ECG**, the object asserted to exist can be constructed by straightedge and compass, continuously in parameters.

Decision functions, such as whether two points are equal or whether a given point lies on a given line, are discontinuous. Past work on the algebraic approach to constructions (see for example [13]) has always assumed decision functions, probably because they seem to be needed to define projection of a point on a line, which is needed in order to introduce coordinates in the Euclidean plane and connect geometry to field theory. But projection itself is continuous. In an early version of our theory **ECG**, we took it as a fundamental operation. In the version presented here, that operation is not necessary, because projection can be defined in terms of the operations of **ECG**, and its essential properties proved using the axioms of **ECG**.

Our interest in this subject began with a computer animation of Euclid's constructions that permits the user to drag the starting points, and see how the construction depends on the changed starting points. The results for Euclid's Book I, Proposition 2, were surprising and interesting. In Book I, Proposition 2, Euclid attempted to show that a rigid compass could be simulated by a collapsible compass. This famous construction shows how to use the rigid compass to construct a point $D = e(A, B, C)$ such that whenever $B \neq C$ and $A \neq B$ then $AD = BC$. The first problem with this construction is that it does not work when $A = C$. Of course, in that case we can just take $D = B$; but that case distinction requires classical logic. And the computer animation reveals that when we drag point C close to A, and then around A in a small circle, then the constructed point D moves around A in a circle of radius close to BC. Hence D does not depend continuously on C. This discontinuity, together with the need for a case distinction just mentioned, makes it clear that from the intuitionistic point of view, there are two different versions of Euclid I.2: the version in which (in addition to $B \neq C$) we take $A \neq C$ as a hypothesis, or the "uniform" version in which we do not assume $A \neq C$, but assert that D can be constructed, whatever the values of A, B, and C, as long as $B \neq C$. Which version of Euclid I.2 is provable with intuitionistic logic turns out to depend on whether we take (a formalism corresponding to) a rigid compass, or Euclid's collapsible compass.

Past work on axiomatizations of constructive (intuitionistic) geometry, such [9] and [14], has replaced the law of the excluded middle by "apartness" (which is explained in the body of the paper). One can introduce an apartness "construction", but it is also not continuous.[2] Therefore we work in theories without apartness.

We are able to obtain the main metatheorem we want using cut-elimination; no new proof-theoretic techniques are developed here, so this paper should be accessible to geometers, not only to logicians.

Geometry is an ancient subject, but it is very much alive. There are a number of avenues of current research that are related to this work, but not dealt with in this paper; and rather than undertake to survey them, we prefer to stick to the specific aims outlined above. Similarly, there is a long history of axiomatizations of geometry. For a review of some of these, see [15]. Here we explain only how our axiomatization is different from others: (i) it uses constructive logic, rather than classical, which distinguishes it from all previous axiomatizations except Heyting's, and (ii) it does not use apartness, which enables it to have the property that points proved to exist can be constructed by Euclid's (continuous) constructions.

2 Euclid's constructions as algorithms

Euclid's five books present 48 two-dimensional constructions and about a dozen three-dimensional constructions. (For simplicity, we will not discuss the three-dimensional part of Euclid, contained mostly in Book 5). We consider the 48 two-dimensional constructions to be the world's first systematic collection of algorithms. (We do not say, "the world's first algorithms", because there certainly were a few number-theoretic algorithms known in China and India much earlier.)

If the constructions are considered as algorithms, then Euclid's *Elements* contained the first proofs of correctness of algorithms.

Euclid presents his readers with both "postulates" and "axioms". Modern mathematicians often treat these words as synonyms. For Euclid and his contemporaries, however, they had quite different meanings. Here is the difference, as explained by Pambuccian [16], p. 12.

> For Proclus, who relates a view held by Geminus, a postulate prescribes that we construct or provide some simple or easily grasped object for the exhibition of a character, while an axiom asserts some inherent attribute that is known at once to one's auditors. And just as a problem differs from a theorem, so a postulate differs from an axiom, even though both of them are undemonstrated; the one is assumed because it is easy to construct, the other accepted because it is easy to know. That is, postulates ask for the production, the *poesis* of something not yet given . . ., whereas axioms refer to the *gnosis* of a given, to insight into the validity of certain relationships that hold between given notions.

Euclid's famous "parallel postulate" states that if two lines L and M are traversed by another line T, forming adjacent interior angles on one side of T adding up to less than two right angles, then L and M will intersect on that side of T. Stated this way, the postulate can be viewed as a

[2] Even though apartness is intuitionistically acceptable, an apartness "constructor" must make use of the idea that points are not given "all at once" but by a sequence of approximations. The same point can be given by different sequences, and though an apartness constructor can be continuous in the approximating sequences, it cannot be continuous in the geometric topology on the points. Hence an apartness constructor needs algorithms beyond the Euclidean constructions. In other words, geometry with apartness goes beyond Euclid. We have extended our metamathematical work to theories with apartness, but we have not presented those extensions in this paper.

construction method for producing certain triangles. Nowadays, the parallel postulate is often stated as an axiom: Given a line L, and a point P not on L, there exists exactly one parallel to L through P. (Parallel lines are by definition lines that do not meet.) Written this way, the parallel postulate does not directly assert the existence of any specific point.[3]

3 The algebraic approach to constructions

The geometrical theory that we shall eventually formulate is quantifier-free, with terms to denote the geometrical constructions. A model of such a theory can be regarded as a many-sorted algebra with partial functions representing the basic geometric constructions. Specifically, the sorts include at least *Point*, *Line*, and *Circle*, and they may possibly include *Arc*, *Ray*, *Triangle*, and *Square*. We have constants and variables of each sort. This collection of data types is almost adequate to cover the return types and argument types of the 48 plane Euclidean constructions. The obvious exception is the construction of a regular pentagon or hexagon. More generally, some of the later constructions use the word "figure", which apparently means something like what a modern mathematician would call a "closed polygon". Some small fixed number of sides would suffice for the Euclidean constructions. In Euclid, no figure with more sides than an octagon is constructed, and no figure with more than four sides is an *input* to another construction, except for constructions that work on any "figure." The general concept of a closed polygon of any number of sides may be logically problematic as it drags the concept of integer into geometry.

Our algebras include function symbols for the basic constructors and accessors:

$\mathit{Line}\ (\mathit{Point}\ A, \mathit{Point}\ B)$	A and B lie on this line
$\mathit{pointOn1}(\mathit{Line}\ L)$	The points from which L was originally constructed
$\mathit{pointOn2}(\mathit{Line}\ L)$	
$\mathit{Circle}\ (\mathit{Point}\ A, \mathit{Point}\ B)$	A is the center, and the circle passes through B
$\mathit{center}(\mathit{Circle}\ C)$	
$\mathit{pointOnCircle}(\mathit{Circle}\ C)$	A point on circle C,

and for the "elementary constructions" (each of which has type *Point*):

$$\begin{aligned}&\mathit{IntersectLines}(\mathit{Line}\ K, \mathit{Line}\ L)\\ &\mathit{IntersectLineCircle1}(\mathit{Line}\ L, \mathit{Circle}\ C)\\ &\mathit{IntersectLineCircle2}(\mathit{Line}\ L, \mathit{Circle}\ C)\\ &\mathit{IntersectCircles1}(\mathit{Circle}\ C, \mathit{Circle}\ K)\\ &\mathit{IntersectCircles2}(\mathit{Circle}\ C, \mathit{Circle}\ K)\end{aligned}$$

Each of these has several "overloaded" variants, which can be defined from these using constructors and accessors. For example,

$$\begin{aligned}&\mathit{IntersectLines}(\mathit{Point}\ A, \mathit{Point}\ B, \mathit{Point}\ C, \mathit{Point}\ D)\\ &\quad = \mathit{IntersectLines}(\mathit{Line}\ (A, B), \mathit{Line}\ (C, D))\\ &\mathit{IntersectLineCircle1}(\mathit{Point}\ A, \mathit{Point}\ B, \mathit{Point}\ C, \mathit{Point}\ D)\\ &\quad = \mathit{IntersectLineCircle1}(\mathit{Line}\ (A, B), \mathit{Circle}\ (C, D))\\ &\mathit{IntersectLineCircle1}(\mathit{Point}\ A, \mathit{Point}\ B, \mathit{Circle}\ C)\\ &\quad = \mathit{IntersectLineCircle1}(\mathit{Line}\ (A, B), C)\end{aligned}$$

[3] For example, it is not immediately clear whether this version implies the first version using only intuitionistic (constructive) logic, although of course it does in classical logic. This question is settled (in the negative) in [3].

As these three examples illustrate, one can regard circles and lines as mere intermediaries; points are ultimately constructed from other points. In a lemma at the end of this section, we state and prove this principle precisely.

We can distinguish the study of such many-sorted algebras from the study of axiomatic first-order theories containing symbols for these algebraic operations. Some questions can be taken up without the consideration of axioms or logical inferences. We shall discuss a few of these.

There is a second constructor for circles, which we can describe for short as "circle from center and radius", as opposed to the first constructor above, "circle from center and point." Specifically $\mathit{Circle3}(A, B, C)$ constructs a circle of radius BC and center A, provided $B \neq C$. These two constructors for circles correspond to a "collapsible compass" and a "rigid compass" respectively. The compass of Euclid was a collapsible compass: you cannot use it to "hold" the length BC while you move one point of the compass to A. You can only use it to hold the radius AB while one point of the compass is fixed at A, so in that sense it corresponds to $\mathit{Circle}\,(A, B)$. The second constructor *Circle3* corresponds to a rigid compass.[4] In the next section we will have more to say about the relationship between these two constructors.

We introduce here a first example of a "construction" not considered by Euclid, the *test-for-equality* construction. This "construction" **D** takes four points, and tests its first two arguments for equality, producing the third or fourth argument depending on the outcome:

$$\mathbf{D}(a, b, c, d) = \begin{cases} c & \text{if } a = b \\ \\ d & \text{if } a \neq b \end{cases}$$

The algebraic approach to constructions was pioneered by Kijne [13], but all the systems he considered contained "decision functions" such as test-for-equality or test-for-incidence. In this paper we will not study systems containing decision functions.

We note in the following lemma that, as far as constructing points goes, the other types are mere conveniences. The elementary constructions can be expressed, as we have noted, in several ways using variables of different types. For example, we could have $\mathit{IntersectLines}(K, L)$ where K and L have type *Line* , or $\mathit{IntersectLines}(A, B, C, D)$, where A, B, C, and D have type *Point*, and

$$\mathit{IntersectLines}(\mathit{Line}(A, B), \mathit{Line}(C, D) = \mathit{IntersectLines}(A, B, C, D).$$

Lemma 1 *Let t be a term of type Point, whose variables are all of type Point. Then there is a term t^* with the same variables as t, also of type Point, such that in the standard plane t and t^* determine the same function, and t^* contains only function symbols of type Point having Point arguments.*

Proof. By induction on the complexity of t. Suppose that t has the form $\mathit{IntersectLines}(r, s)$ where r and s are terms of type *Line* . None of the elementary constructions has type *Line* , and r and s cannot be variables (since all variables in t are of type *Point*), so r must have the form $\mathit{Line}\,(p, q)$ for some terms p and q, and s must have the form $\mathit{Line}\,(u, v)$ for some terms u and v. Then t^* can be taken to be $\mathit{Line}\,(p^*, q^*, u^*, v^*)$. The other elementary constructions are treated similarly. The basis case, when t is a variable or constant, is treated by taking $t^* = t$.

4 Models of the elementary constructions

The algebraic approach allows us to consider "models" without (yet) having formulated any axioms or logical theories. There are several interesting models, of which we now mention four. To define these models, we assume there are three constants α, β, and γ of type *Point* .

[4] There is a word in Dutch, *passer*, for this type of compass, which was used in navigation in the seventeenth century. But there seems to be no single word in English that distinguishes either of the two types of compass from the other.

The *standard plane* is $\mathbf{R}^2$. Points, lines, and circles (as well as segments, arcs, triangles, squares, etc., in extended algebras in which such objects are considered) are interpreted as the objects that usually bear those names in the Euclidean plane. More formally, the interpretation of the type symbol "*Point*" is the set of points, the interpretation of "*Line*" is the set of lines, etc. In particular we must choose three specific non-collinear points to serve as the interpretations of α, β, and γ. Let us choose $\alpha = (0,1)$, $\beta = (1,0)$, and $\gamma = (0,0)$. The constructor and accessor functions listed above also have standard and obvious interpretations. It is when we come to the five operations for intersecting lines and circles that we must be more specific. There are three issues to decide:

- when there are two intersection points, which one is denoted by which term?
- In degenerate situations, such as *Line* (P,P), what do we do?
- When the indicated lines and/or circles do not intersect, what do we do about the term(s) for their intersection point(s)?

We take up the last item first. When, for example, line L does not meet circle C, we say that the term *IntersectLineCircle1*(L,C) is *undefined*. In other words, the operations of these "algebras" do not have to be defined on all values of their arguments. The same issue, of course, arises in many other algebraic contexts, for example, division is not defined when the denominator is zero, and $\sqrt{x}$ is not defined (when doing real arithmetic) when x is negative.

Regarding the "which is which" issue, our guiding principle is continuity. We therefore make the following definitions: *IntersectCircles1*(C,K) is the intersection point P such that the angle from *center*(C) to *center*(K) to P makes a "left turn". This is defined as in computer graphics, using the sign of the cross product. Specifically, let $A =$ *center*(C) and $B =$ *center*(K). Then the sign of $(A-B)\times(P-B)$ determines whether angle ABP is a "left turn" or a "right turn". Thus $\alpha\beta\gamma$ is a left turn and $\gamma\beta\alpha$ is a right turn. In case the two intersection points are different, one of these cases must apply. This explanation has used a case distinction as well as the cross product; later we will show how to define "left turn" and "right turn" using only the axioms of Euclidean geometry and intuitionistic logic. For now we simply note that this notion is constructively appealing, because of continuity: there exists a unique continuous function of C and K that satisfies the stated handedness condition for *IntersectCircles1* when C and K have two distinct intersection points, and is defined whenever C and K intersect (at all).

The principle of continuity leads us to make *IntersectCircles1*(C,K) and *IntersectCircles2*(C,K) undefined in the "degenerate situation" when circles C and K coincide, i.e. have the same center and radius. Otherwise, as the center of C passes through the center of K, there is a discontinuity. It makes sense, anyway, to have them undefined when C and$K$ coincide, as the usual formulas for computing them get zero denominators, and there is no natural way to select two of the infinitely many intersection points.

We still need to settle the "which is which" issue for *IntersectLineCircle1*(*Line* $(A,B),C)$ and *IntersectLineCircle2*(*Line* $(A,B),C)$. Here the rule is that these two points must occur in the same order on *Line* (A,B) as A and B do. Again, we require continuity of the function

$$\textit{IntersectLineCircle1}(A,B,P,Q)$$

with four *Point* variables, i.e.

$$\textit{IntersectLineCircle1}(\textit{Line}\ (A,B), \textit{Circle}\ (P,Q)).$$

There is a unique continuous extension from the set of (A,B,P,Q) where there are two intersection points to the set where there is at least one intersection point; that extension is the interpretation of *IntersectLineCircle1*.

Regarding the degenerate forms *Line* (P,P), we say that the term is undefined. With respect to degenerate circles, *Circle* (A,A), continuity and computability do not present the same obstacles as in the case of degenerate lines *Line* (A,A). Thus we have a choice to allow degenerate segments and circles, without destroying the continuity of the elementary constructions.

We may choose to allow them or not. In the "standard model" R^2 we take *Circle* (A, A) to be defined, i.e. circles of zero radius are allowed. There is only one point on such a circle. Then *IntersectLineCircle1*$(L, Circle(A, A))$ is defined if A is on line L, and is equal to A. We also consider (briefly) the model R^{2-} in which degenerate circles are not allowed, so *Circle* (A, A) is undefined.

The *recursive plane* $\mathbf{R_{rec}}^2$ consists of points in the plane whose coordinates are given by "recursive reals". We write $\{e\}(n)$ for the result, if any, of the e-th Turing machine at input n. Rational numbers are coded as certain integers, and modulo this coding we can speak of recursive functions from $\mathbf{N}$ to $\mathbf{Q}$. A "recursive real number" x is the index of a Turing machine e such that $|\{e\}(n) - x| \le 1/n$ for each $n \in \mathbf{N}$. The real number to which the approximations $\{x\}(n)$ converge is sometimes also called a "recursive real number", but we call it the "value of x". It is a routine exercise to show that the recursive points in the plane are closed under the Euclidean constructions. In particular, given approximations to two circles (or to a circle and a line), we can compute approximations to their "intersection points", even though it may turn out that when we compute better approximations to the circles, we see that they do not intersect at all.

In the recursive plane, there is no computable test-for-equality function, that is, no computable function D that operates on two Turing machine indices x and y, and produces 0 when x and y are recursive real numbers with the same value, and 1 when they are recursive real numbers with different values. Proof, if we had such a D, we could solve the halting problem by applying D to the point $(E(x), 0)$, where $\{E(x)\}(n) = 1/n$ if Turing machine x does not halt at input x in fewer than n steps, and $\{E(x)\}(n) = 1/k$ otherwise, where x halts in exactly k steps. Namely, $\{x\}(x)$ halts if and only if the value of $E(x)$ is not zero, if and only if $D(Z, E(x)) \neq 0$, where Z is an index of the constant function whose value is the (number coding the) rational number zero.

Readers familiar with recursion theory may realize that there are several ways to define computable functions of real numbers. The model we have just described is essentially the plane version of the "effective operations". It is a well-known theorem of Tseitin, Kreisel, LaCombe, and Shoenfield, known traditionally as KLS (and easily adapted to the plane) that effective operations are continuous. Of course, in the case at hand we can check the continuity of the elementary constructions directly.

The *algebraic plane* consists of points in the plane whose coordinates are algebraic. Since intersection points of circles and lines are given by solutions of algebraic equations, the algebraic plane is also closed under these constructions. Since algebraic numbers can be computed, this is a submodel of the recursive plane.

In the algebraic plane, there *is* a computable test-for-equality function D. We assume algebraic numbers are given by means of a rational interval (a, b) and a square-free polynomial $f \in \mathbf{Q}[x]$ such that f has only one root in (a, b). To determine if (a, b) and f determine the same or a different real number than (p, q) and g, first check if the two rational intervals overlap. If not, the two reals are different. If so, let (r, s) be the intersection. Now we have to determine if f and g have a common zero on (r, s). There is a simple recursive algorithm to do that: Say g has degree greater than or equal to that of f. Then write $f = gh + r$ with r of lower degree than g. Then f and g have a common zero on (r, s) if and only if f and r have a common zero. Recurse until both polynomials are linear, when the decision is very easy to make.

The *Tarksi model* is $K \times K$, where $K = \mathbf{Q}(\sqrt{\ })$ is the least subfield of the reals containing the rationals and closed under taking the square root of positive elements. This is a submodel of the algebraic plane.

In such models, the elementary constructions are interpreted as functions from M to M, and of course M must be closed under those functions, but they may also have other interesting properties. For example, in the examples given above, the (interpretations of the) elementary constructions are all computable functions; indeed they are algebraic functions of low degree. In particular, they are all continuous on their domains. Hence, for example, there is no algebraic test-for-equality function in the algebraic plane, even though there is a computable one (since a test-for-equality function would enable us to define a function $f(x)$ such that $f(0) = 0$ but $f(x) = 1$ when $x \neq 0$; since f is not continuous, it cannot be algebraic).

5 Euclid's Book I, Proposition 2

Book I, Proposition 2 has been discussed in the introduction. The question it addresses concerns the constructor *Circle3*(A, B, C), which constructs a circle with center A and radius BC. As discussed in the introduction, Euclid gives a term that accomplishes this aim under the assumptions, not only that $B \neq C$, but also that $A \neq C \vee B \neq C$. We consider the stronger theorem "uniform Euclid I.2", which asserts that for every A and BC with $B \neq C$, there exists a D with $AD = BC$. The "uniformity" refers to the missing assumption $A \neq B$. To "realize" uniform Euclid I.2, we would need a term $e(A, B, C)$ that produces D uniformly, whether $A = B$ or not. If we had such an e, then of course we could define

$$\mathit{Circle3}(A, B, C) = \mathit{Circle}\ (A, e(A, B, C)).$$

Conversely, such an e can be defined from *Circle3* like this:

$$e(A, B, C) = \mathit{pointOnCircle}(\mathit{Circle3}(A, B, C))$$

But the mere fact that Euclid's own construction does not suffice to define *Circle3* does not show that some other construction won't do the job. However, we are able to prove that no other construction can define *Circle3*, as a corollary to the following theorem, which shows that without *Circle3*, no total function is definable, at least if we insist that *Circle* (x, x) is undefined.

Theorem 1 (with Freek Wiedijk) *No total unary point-valued function (other than the identity) is definable in the standard model* R^{2-} *(with degenerate circles undefined) from the elementary geometrical constructions excluding Circle3. More precisely, let* t *be a compound term containing exactly one variable* A *of type Point, and no other variables, but possibly containing some constants* $\alpha, \beta, \ldots$ *of type Point. Suppose that* t *does not contain Circle3. Let the constants be interpreted as certain (fixed) distinct points in the standard plane. Then for some value of* A*,* t *is not defined in the standard plane. In fact, we can make* t *undefined by assigning* A *the same value as any constant occurring in* t*, or if* t *has no constants,* t *is always undefined.*

Proof. We start by eliminating the "overloaded" versions of the elementary constructions from t. For instance, if t contains a subterm $IntersectLine(a, b, c, d)$, we replace it by a term using the "fundamental" form of the construction, *IntersectLines*(*Line* (a, b), *Line* (c, d)). The result of such replacements is a term with the same value as t under any assignment of a value to the variable A, and containing no variables of types other than *Point* .

We proceed by induction on the complexity of terms t as in the theorem, but also containing only the fundamental versions of the elementary constructions (no overloaded versions).

Since the theorem only applies to compound terms, the basis case occurs when t has only variables or constants for arguments. We note that *Circle* (A, A), *Line* (A, A), *Ray* (A, A), *Segment* (A, A), *Arc* (A, A, A) are undefined. But we also have to consider the possibility that the constants β or γ occupy one of the argument places. For example, *Circle* (A, β) is undefined when A takes the value $\bar{\beta}$ that is assigned to the constant β. In the rest of the proof we shorten this kind of statement to "is undefined when $A = \beta$." Similarly, *Circle* (β, A) is undefined when $A = \beta$; *Line* (A, β) and *Line* (β, A) are undefined when $A = \beta$, and the same for *Ray* , *Arc* , and *Segment* . Note that this argument does not work for *Circle3*, since *Circle3*(A, β, γ) is always defined, but by hypothesis, t does not contain *Circle3*.

Since t does not contain any overloaded constructions, the basis case is finished, as there is no fundamental construction that takes only arguments of type *Point*. Specifically, there are just five fundamental constructions for producing the intersections of lines and lines, lines and circles, or circles and circles, and they each need arguments of type *Line* or *Circle*.

Now consider the induction step. If the main symbol of t is a constructor, such as *Line* , then t has the form *Line* (a, b). One of a or b must contain a variable, and hence be somewhere undefined. Hence t is also somewhere undefined, and indeed the same assignment of a value to A that makes a or b undefined will work. Similarly for the other constructors (since *Circle3*, to which this argument does not apply, is not allowed).

Next consider the case when t is *IntersectLines*(p, q). Then p and q have type *Line*. The only terms of type *Line* are those of the form *Line* (a, b), or variables of type *Line*. But t is not allowed to contain variables of type *Line*. Therefore t must have the form,

$$\mathit{IntersectLines}(\mathit{Line}\ (a, b), \mathit{Line}\ (c, d))$$

Since t contains a variable, one of $s = \mathit{Line}\ (a, b)$ or $s = \mathit{Line}\ (c, d)$ must contain a variable, and by induction hypothesis the term $s\sigma$ is undefined when substitution σ assigns the variable of s to one of the constants in s (or any constant if s has no constants). Hence $t\sigma$ is also undefined.

Similarly for the other elementary constructions. That completes the proof.

Corollary 1 *Let $e(A, B, C)$ be any term built up from the elementary constructions, not containing Circle3, having type Point, and containing exactly three variables A, B, and C of type Point. Then it is not the case that in R^{2-}, whenever $B \neq C$ then $e(A, B, C)$ is defined and is a point D such that $AD = BC$.*

Proof. Let us invent two constants β and γ, and interpret them as two distinct points $\bar{\beta}$ and $\bar{\gamma}$ (fixed for the rest of the proof). Then let $f(A) = e(A, \beta, \gamma)$. Suppose, for proof by contradiction, that $e(A, B, C)$ is defined whenever $B \neq C$ and is a point D such that $AD = BC$. Then $f(A)$ is defined for all A and is always different from A, contradicting the previous theorem.

Corollary 2 *Circle3 is not definable in the model R^{2-} from the (other) elementary constructions.*

Proof. *Circle3* is a term e fulfilling the hypotheses of the previous corollary.

Circle3 is intimately connected with the "compass" of "straightedge and compass constructions". Euclid's compass is supposed to be "collapsible", so that you cannot use it to measure BC and then move it to draw a circle of that same radius centered at A. Therefore Euclid proved Euclid I.2, showing how you can accomplish this with a collapsible compass; but for that we need to assume $A \neq B$. In effect we need a test-for-equality construction. When Proclus criticized Euclid for omitting arguments by cases, perhaps this is what he had in mind. The theorem "for every A, B, and C with $B \neq C$ there exists a point D with $AD = BC$" is (apparently) not "realized" by an elementary construction $D = f(A, B, C)$.

The theorem above shows that without *Circle3*, we cannot even define a construction f such that for each point A, $f(A)$ is a point different from A. We also cannot define a construction $f(A, L)$ that takes a point A and line L into a point on L, different from A.

The fact that *Circle3* is not definable in R^{2-} means that, if we do not include *Circle3* as a primitive construction, we shall not be able to define it in any axiomatic theory that has R^{2-} for a model. It seems clear that Euclid's book does have R^{2-} as a model; so the uniform version of Book I, Proposition 2 is essentially non-constructive, using Euclid's non-rigid compass. We therefore add *Circle3* as a fundamental construction (rendering I.2 a triviality) and give a constructive theory that works for the rest of Euclid.

There is still an unresolved technical issue here: is *Circle3* undefinable in R^2? The above theorem does not extend to R^2, because, for example, *IntersectLineCircle2*(*Line* (A, B), *Circle* (B, x)) is a total function if *Circle* is total. We conjecture that *Circle3* is undefinable in R^2 as well as in R^{2-}.

6 Circles of zero radius and *Circle3*

In this section, we address the two issues of whether our basic theory ought to include *Circle3* and whether we ought to allow circles of zero radius. We conclude that both should be allowed. Our approach is to consider what is required to achieve a formalization of Euclid using intuitionistic logic, with no test-for-equality construction needed.

Here are some possible constructions we wish to consider.

Euclidean Extension. One of the Euclidean axioms says that we can extend a given segment AB by a segment CD. More precisely, we can construct a point $P = \mathit{Extend}(A, B, C, D)$ such that $BP = CD$ and B is between A and P. The assumptions here are that $A \neq B$ and $C \neq D$, but it is not assumed that $B \neq C$ or $B \neq D$.

Strong Extension. $\mathit{ExtendA}(A, B, C, D)$ is the unique continuous extension of $\mathit{Extend}(A, B, C, D)$ that is defined when $A \neq B$, i.e. without assuming $C \neq D$. When $C = D$, we have

$$\mathit{ExtendA}(A, B, C, D) = B.$$

We will show below how to define *ExtendA* in terms of *Circle3*.

Projection. The construction $\mathit{project}(P, L)$ takes a point P and line L and produces a point Q on L such that P lies on the perpendicular to L at Q. The well-known Euclidean construction for the projection applies only if P is known not to be on L. To define *project* using that construction, we would require a test-for-incidence that allows us to test whether point P is on line L or not. But no such test-for-incidence construction is computable over the computable plane, so the Euclidean projection construction does not lead in any obvious way to a definition of *project*. (That does not, however, constitute a proof that *project* is not definable in terms of the elementary constructions, which we give below.)

Projection is absolutely necessary in order to reduce geometry to algebra. We want to pick a line, call it X, and erect a perpendicular Y to X, and project each point P onto its coordinates $x = \mathit{project}(P, X)$ and $y = \mathit{project}(P, Y)$.

Lemma 2 *Strong extension and projection are definable from Circle3 (with circles of zero radius allowed). Also one can, using those primitives, construct the perpendicular to line L passing through point P, without conditions as to whether P is or is not on L.*

Proof. First we define

$$\mathit{ExtendA}(A, B, C, D) = \mathit{IntersectLineCircle2}(\mathit{Line}\ (A, B), \mathit{Circle3}(B, C, D))$$

Since circles of zero radius are allowed, we do not need to assume $C \neq D$.

To construct the projection of point P on line L, we just need some circle with center P that intersects L in two distinct points Q and R. Then the projection of P on L is the midpoint of segment QR. If L is $\mathit{Line}\ (A, B)$, then a suitable radius would be the sum of the lengths of AB and PA. Thus the circle we need can be constructed as $\mathit{Circle3}(P, A, \mathit{ExtendA}(A, B, P, A))$. Note that we need strong extension, not just Euclidean extension, because we cannot rule out $P = A$. That completes the definition of projection.

Now, to construct the perpendicular to L at P, we simply erect the perpendicular to L at the projection of P on L, using the usual Euclidean construction. That completes the proof of the lemma.

The following lemma helps make the case for allowing circles of zero radius and for allowing *Circle3*.

Lemma 3 *In R^{2-} (where circles of zero radius do not exist), projection is not definable in terms of the elementary constructions without Circle3.*

Proof. Let α and β be constants, whose values will be two distinct (fixed) points. Let $t(A)$ be the term $\mathit{project}(A, \mathit{Line}\ (\alpha, \beta))$. Then $t(A)$ is defined for all values of A. By Theorem 1, $t(A)$

is not definable in terms of the elementary constructions without *Circle3*, in the model R^{2-}. That completes the proof.

There are also some "constructions" that go "beyond Euclid." We have already mentioned the test-for-equality function **D**. In addition, there is a notion of *apartness* introduced by Heyting [9]. The apartness axiom says that if $B \neq C$, then for any other point A, either $A \neq B$ or $A \neq C$. An "apartness construction" would be an operation $\#$ such that if $B \neq C$ then $\#(A, B, C)$ is defined[5] , and is equal to B or to C, and satisfies $A \neq \#(A, B, C)$. In other words, when B and C are different, A must be different from one of them, and this constructor picks one from which A is different.

If we have an apartness construction, then $\#(\#(A, B, C), B, C)$ is the *other* point of B and C, so it would not be necessary to add a second apartness construction.

The following lemma just summarizes what the recursive model R^{2-} tells us about apartness and test-for-equality.

Lemma 4 *(i) No apartness construction is definable in terms of the elementary constructions and Circle3. (ii) Test-for-equality is not definable in terms of the elementary constructions and Circle3, even with the aid of an apartness construction.*

Proof. Ad (i): No apartness construction is definable in terms of the elementary constructions and *Circle3*, since all such terms define continuous functions, and # is not continuous. Ad (ii): We have shown in a previous section that no test-for-equality function exists in the recursive plane. However, the recursive plane does have *Circle3* and an apartness construction. Here is how to compute a point apart from one of two distinct recursive points B and C. Namely, let n be a positive integer such that $1/n$ is less than the length of segment BC, and compute rational approximations a, b, and c to A, B, and C within $1/(4n)$. Then a cannot be within $1/(4n)$ of both b and c, and the answer is C if $|a - b| < 1/(4n)$ and B otherwise. That completes the proof.

We investigated a system in which *Circle3* is included, but circles of zero radius are not allowed, and instead *project* is taken as primitive. This is a workable system, but it is a bit more complex, and there seems to be no intuitive justification for projection in terms of the compass that does not also suffice to justify allowing circles of zero radius. In fact it can be argued that there is no good reason *not* to allow the two points of the compass to coincide. We therefore chose to allow *Circle3*, and to allow circles of zero radius by requiring that *Circle3*(A, B, C) is always defined.

7 Continuous Coordinatization and Arithmetic

Nowadays we usually think of analytic geometry as coordinatizing a plane and translating geometrical relations between points and lines into algebraic equations and inequalities. But the converse is also possible: translating algebra into geometry, and this is important for lower estimates on the power of geometric constructions, for example for showing that the models of the geometry of constructions are planes over Euclidean fields.

In modern books (such as [5]) arithmetic is geometrized as operations on congruence classes of segments. We operate instead on points on some fixed line $X = Line\,(0, 1)$, where 0 and 1 are two arbitrarily fixed points. As far as I can tell, past work on coordinatization has always assumed some discontinuous constructions, such as test-for-equality or at least apartness. Since coordinatization itself is patently computable and continuous, it is in some sense "overkill" to appeal to discontinuous and non-computable "constructions" to achieve coordinatization and arithmetization. Although coordinatization is standard, old, and not complicated, we need to check that it can in fact be done from the specified primitives, without using apartness or test-for-equality, by definitions that apply without (for example) case distinctions as to whether

[5] Strict constructivists will want to assume B is apart from C in the definition of an apartness construction, but we haven't defined that yet and the discussion here makes sense with classical logic.

numbers being multiplied are equal to 0 or 1 or not. We note that it is crucial that circles of zero radius be allowed (or else we need to take projection as another primitive).

In this section, we give the definitions of constructions that serve to implement coordinatization and the arithmetic operations in a continuous way. We show that, in the models discussed above, these operations are defined and behave as desired. In a later section, we will give an axiomatic theory capable of formalizing these correctness proofs without reference to models.

Recall that the concepts of "right turn" and "left turn" have been introduced and discussed in Section 4. For example, in Fig. 1, AOZ is a "right turn", because the sign of the cross product $OA \times OB$ is positive. Intuitively, traveling from A to O to Z requires one to turn right at O. This definition of "right turn" is adequate for this section, since we are only concerned with models where cross product makes sense. (Later, we will expend considerable effort defining "right" and "left" in an axiomatic context.)

To perform addition geometrically we suppose given a line $L = \mathit{Line}\ (R, S)$ and an "origin", a point O on L with S between R and O. We need to define a construction $\mathit{Add}(A, B)$, which also depends, of course, on S, R, and O, such that $\mathit{Add}(A, B)$ is a point C on L representing the (signed) sum of A and B, with O considered as origin. In particular if A and B are on the same side of O then $\mathit{Add}(A, B) = \mathit{ExtendA}(O, A, O, B)$, but that does not suffice to define Add. It may be instructive to see another failed attempt to define $\mathit{Add}(O, A, B)$:

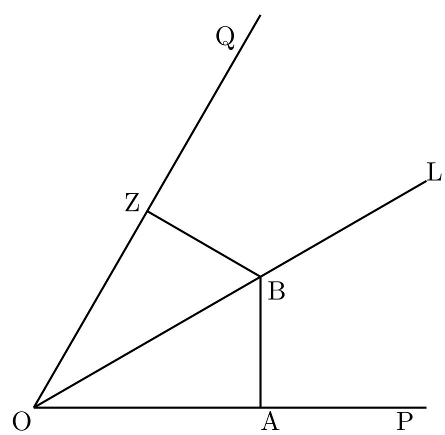
$$\mathit{IntersectLineCircle2}(\mathit{Line}\ (O, B), \mathit{Circle}3(A, 0, B))$$

which is correct when $O \neq B$, but is undefined when $O = B$.

Lemma 5 *Given line L = Line(R,S), and a point O on L with S between R and O, we can construct a point $\mathit{Add}(A, B)$ on L representing the signed sum of A and B, with O considered as origin, using the elementary constructions and Circle3.*

Proof. We first define an auxiliary construction *rotate*, which requires as inputs three distinct, non-collinear points P, O, and Q (think of angle POQ), as well as a point A on $\mathit{Line}\ (O, P)$. The desired result of $\mathit{rotate}(P, O, Q, A)$ is a point Z on $\mathit{Line}\ (O, Q)$ such that $OZ = OA$ and if $A \neq O$ then AOZ is a right turn. (See Fig. 1.) The point is that Z is defined even when

Figure 1: $Z = \mathit{rotate}(P, O, Q, A)$



$A = O$ (in which case it is just O, of course), and if A moves along $\mathit{Line}\ (O, P)$ through O, then Z moves along $\mathit{Line}\ (O, Q)$, passing through O as A does. To construct Z, we first bisect the angle POQ (by the usual Euclidean construction, which is not problematic since the three points are not collinear). Let the angle bisector be line L. Then let line K be the perpendicular to *Line* (O, P) at A, let B be the intersection point of K and L, and let Z be the projection of B on *Line* (O, Q) (which is defined no matter whether $O = A$ or not).

Note that there are, if $A \neq O$, two points Z on OQ such that $OZ = OA$. The one constructed by $rotate(P, O, Q, A)$ is such that, if POQ is a right turn, then AOZ is a right turn when $A \neq O$, regardless of whether A is between O and P or not. Similarly, if POQ is a left turn, so is AOZ.

With *rotate* in hand, we can give a construction for $Add(A, B)$ (depending also on R, S, and O). (The construction is illustrated in Fig. 2.) First, we replace R and S with new points on $L = Line\ (R, S)$, farther away from O, so that O, A, and B are all on the same side of R and S, and the new R and S are in the same order on line L as before. (This can be done using *Extend*.) Now erect the perpendicular K to L at O, and the perpendicular H to L at B. In the process of erecting these perpendiculars, we will have constructed points C on K and D on H such that ROC is a right turn. Then let

$$\begin{aligned} U &= rotate(R, O, C, A) \\ V &= project(U, H) \\ W &= rotate(D, B, R, V) \end{aligned}$$

We set $Add(A, B) = W$. Then $Add(A, B)$ is defined for all A, B. Suppose $A \neq O$. Then UV is perpendicular to both K and H. Then U and V are on the same side of L, since if UV meets L at a point X, then XU and XO are both perpendicular to K, which implies $U = O$, which implies $A = O$, contradicting $A \neq O$. It then follows from the property of *rotate* that B and W occur on line L in the same order that O and A occur. Refer to Fig. 3 for an illustration of the case when A is negative. This implies that $Add(A, B)$ represents the algebraic sum of A and B, since in magnitude $BW = OA$.

Figure 2: Signed addition without test-for-equality

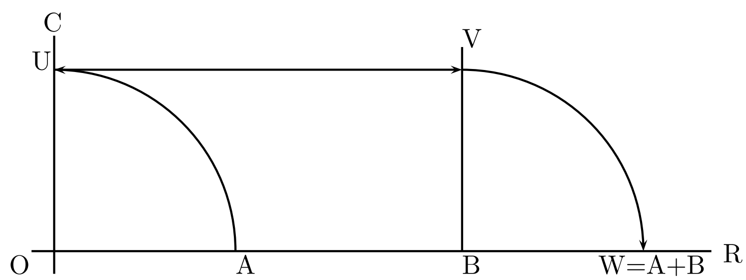
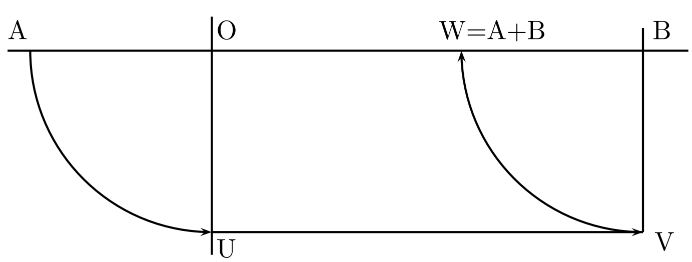


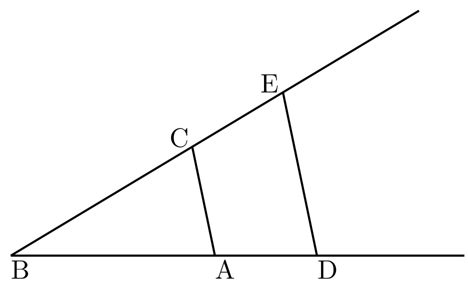
Figure 3: Signed addition when A is negative



Having defined addition, we now turn to multiplication, division, and square root. The geometrical definitions of these operations go back to Descartes. On the second page of *La*

Geometrie [6], he gives constructions for multiplication and square roots. We reproduce the drawings found on page 2 of his book [6] in Figures 4 and 5. Here is Descartes' explanation of these figures:

Figure 4: *La Multiplication* according to Descartes



1. *Let AB be taken as unity.*
2. *Let it be required to multiply BD by BC. I have only to join the points A and C, and draw DE parallel to CA; then BE is the product of BD and BC.*
3. *If it be required to divide BE by BD, I join E and D, and draw AC parallel to DE; then BC is the result of the division.*
4. *If the square root of GH is desired, I add, along the same straight line, FG equal to unity; then, bisecting FH at K, I describe the circle FIH about K as a center, and draw from G a perpendicular and extend it to I, and GI is the required root.*

From the point of view of constructive geometry, there is a problem with the construction. Namely, Descartes has only told us how to multiply two segments with non-zero lengths, and at least one of whose lengths is not 1 (the length of unity–he needs this when constructing AC parallel to DE), while we want to be able to multiply in general, without a test-for-equality construction. To solve this problem, we recall from Lemma 2 that we can define $perp(P, L)$, the perpendicular to L passing through P, without regard to whether P is or is not on L. Then we can define a construction *para* such that, for any line L and any point P (which may or may not be on L), $para(P, L)$ passes through P, and if P is not on L then $para(P, L)$ is parallel to L, while if P is on L, then $para(P, L)$ has the same points as L. The definition of *para* is

$$para(P, L) = perp(P, perp(P, L)).$$

In words: First find the perpendicular to L passing through P. Then erect the perpendicular to that line at P.

Using the *para* construction where Descartes calls for "drawing DE parallel to CA", we no longer have a problem multiplying numbers near 1 or 0. We now give a construction for multiplication (which of course could be written as a single, much less readable, term). The construction assumes that 0 and 1 are two distinct points on line X, and D and Q are points on line X to be multiplied, and Y is another line through X, meeting X at 0 and distinct from X. We could, for example, take Y to be the perpendicular to X at 0, although that does not match the illustration from Descartes' book.

```
Multiply(Point D, Point Q)
{ I = IntersectLineCircle1(Y,Circle(0,1))
  C = rotate(1,0,I,Q)    // f is as in the lemma, so OC = OQ
```

```
  // now we have to multiply BD by BC  per Descartes
  L = Line(1,C)       // AC in Descartes' diagram.  A there is 1 here; B is 0.
  K = para(1,C,D)     // parallel to AC through D  (or AC itself if D=1)
  E = IntersectLines(K,Y)  // defined because K is not parallel to Y
  N = rotate(I,0,1,E)        // rotate length OE back to line X from line Y
  return N
}
```

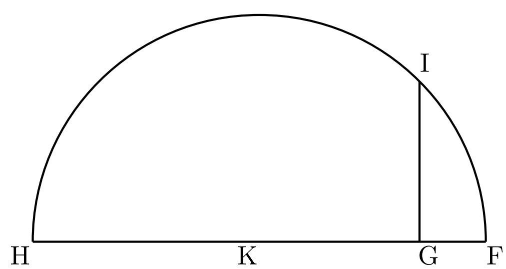
Note that since I is defined by *IntersectLineCircle1*, which side of 0 it lies on (on Y) is determined arbitrarily, by the unknown order on Y of the two points *pointOn1*(Y) and *pointOn2*(Y). That side of 0 on Y becomes the "positive" side. If Q is positive (lies on the same side of $Zero$ as *One* does) then C is positive, i.e. lies on the same side of 0 as I, and vice-versa, if Q is negative (lies on the opposite side of 0 as 1 does), then C lies on the opposite side of $Zero$ as U. This makes multiplication of signed numbers come out correctly without needing a test-for-equality constructor.

It remains, of course, to prove in some geometrical theory that multiplication and addition satisfy the field laws. We do not take that up at this point since we have not yet discussed theories and axioms.

Descartes' division method is handled similarly, using *para* where Descartes constructs a parallel. Turning to Descartes' square root construction, we show that it can be carried out uniformly, without needing to know that the point whose square root is required is different from $Zero$. We carry out Descartes's construction, but then we need to find a point on X whose distance from 0 is IG, without dividing into cases as to whether $G = 0$ or not.

Now we take up the geometrical construction of square roots. Fig. 5 shows Descartes' construction.

Figure 5: Square roots according to Descartes




```
SquareRoot(Point G)
{ // H in Descartes' diagram is 0
  F = Add(0,G,0,1)  // so FG has unit length
  K = Midpoint(F,0)
  C = Circle(K,F)
  L = perp(G,Line(0,F))
  I = IntersectLineCircle1(L,C)
  U = IntersectLineCircle1(L,Circle(G,F))  // rotate unit length to line L
  R = rotate(U,G,F,I) // so now RG = IG but R is on X, on the same side of G as F
  // now we need N so that N0 = RG
  MinusOne = IntersectLineCircle2(1,0,Circle(0,1))
  N = ExtendA(MinusOne,0,G,R)
  return N
}
```

The point here is that we do not need to assume $G \neq 0$ for this construction to work; we only need that 0 is not between G and 1, i.e. loosely speaking $G \geq 0$. This works because *perp* is total.

8 Euclidean Constructive Geometry ECG

In this section we develop a first order axiomatic theory of geometry as close as possible to Euclid. We call it **ECG**, for "Elementary Constructive Geometry". Euclid worked with the following data types: *Point*, *Line*, *Circle*, *Segment*, and *Arc*.[6] For foundational purposes, it seems simplest to use only *Point*, *Line*, and *Circle*, and that is what we do in **ECG**. We therefore choose a multi-sorted theory, with "sorts" corresponding to those types. We use the words "sort" and "type" synonymously in this paper. It is, of course, not difficult to add sorts *Ray* , *Arc* and *Segment* , and axioms making the extended theories conservative over **ECG**, but we do not do so here. We take function symbols corresponding to constructors and accessors for those types, described in detail below. The relation symbols we use are standard in axiomatic geometry, **B** for (strict) betweenness and δ for equidistance. We emphasize that **B** is used for strict betweenness; as Hilbert put it, if C is between A and B, then A, B, and C are three distinct points.

We use $on(P, L)$ for the incidence of point P on line L, and $On(P, C)$ for the incidence of point P on circle C. There is a complete list of the axioms of **ECG** in the Appendix, for reference. In this section, we introduce these axioms one by one, with discussion and explanation.

ECG has five basic function symbols, shown here with arguments:

$$\begin{array}{l} \textit{IntersectLines}(L, K) \\ \textit{IntersectLineCircle1}(L, C) \\ \textit{IntersectLineCircle2}(L, C) \\ \textit{IntersectCircles1}(C, K) \\ \textit{IntersectCircles2}(C, K) \end{array}$$

The intuitive meaning of these symbols has been discussed above. **ECG** does not have "overloaded" versions of these functions; in other words, we just write

$$\textit{IntersectLines}(\textit{Line}\ (A, B), \textit{Line}\ (P, Q))$$

instead of having an overloaded version of *IntersectLines* that takes four points.

Our underlying logic is intuitionistic. We first give the specifically intuitionistic parts of our theory, which are very few in number. We do not adopt decidable equality, nor even the substitute concept of "apartness" introduced by Heyting (and discussed above), primarily because we aim to develop a system in which definable terms (constructions) denote continuous functions, but also because we wish to keep our system closely related to Euclid's geometry, which contains nothing like apartness. Our first four axioms express our intuition that there is nothing asserting existence in the meaning of equality or incidence; hence assertions of equality or incidence can be constructively proved by contradiction.

$\neg\neg x = y \supset x = y$ (Axiom 1)
$\neg\neg\delta(A, B, C, D) \supset \delta(A, B, C, D)$ (Axiom 2)
$\neg\neg on(P, L) \supset on(P, L)$ (Axiom 3)
$\neg\neg On(P, C) \supset On(P, C)$ (Axiom 4)

[6] Euclid also worked with triangles, squares, pentagons, hexagons, and "figures". By "figure" he seems to have meant, "closed polygon". One cannot work with arbitrary figures without introducing variables for integers, which in the modern view takes us beyond geometry. We therefore view those theorems of Euclid that mention "figure" as geometrical theorem *schemata*, which result in a theorem about polygons of N sides, for each particular integer N.

We will take care to formulate our axioms without quantifiers and without disjunction, which will be key to our applications of proof theory. What we aim to do in this section is to formulate such a theory, which we feel is quite close in spirit to Euclid. In formulating this theory, we made use of the famous axioms of Hilbert [10], which have been given a more modern and detailed formulation in the textbook of Greenberg [8]. Of course, we do not take the full continuity axioms of Hilbert, but only the line-circle and circle-circle continuity axioms. Where possible, we formulate our axioms as correctness statements about the constructions; in that form they are automatically quantifier-free. Some axioms, which are not about constructions, are inherently quantifier-free. The only question of serious interest is whether disjunction can be completely avoided. It can, as it turns out. The details of this axiomatization may seem somewhat tedious, but the system must of course be specified in complete detail in order to use it in metamathematical proofs. Moreover, some of the details, as far as I can determine, are actually new. In particular, we show how to define the relations "ABC is a left turn" and "ABC is a right turn" in our theory; the experts we consulted thought this was new. We need this in order to distinguish the two intersection points of two circles.

Since many of our function symbols denote "partial functions", i.e. functions that are not always defined, we will use the "logic of partial terms" **LPT**in our theories. This is a modification of first-order logic, in which the formation rules for formulas are extended by adding the following rule: If t is a term then $t\downarrow$ is a formula. Then in addition the quantifier rules are modified so instead of $\forall x(A(x) \supset A(t))$ we have $\forall x(t\downarrow \wedge A(x) \supset A(t))$, and instead of $A(t) \supset \exists x\, A(x)$ we have $A(t) \wedge t\downarrow \supset \exists x\, A(x)$. Details of LPT can be found in [2], p. 97.

We could try to deal with partial terms, such as $\sqrt{x}$, by simply using an ordinary function symbol for $\sqrt{}$, but not saying anything in the axioms about $\sqrt{}$ of negative numbers. Thus $\sqrt{-1}$ would some real number, but we would not know or care which one, and we would not be able to prove that its square is -1. This approach rapidly becomes awkward when complicated terms involving square roots of different quantities are used, and you must add extra hypotheses to every theorem asserting that what is under every square root is positive, and we choose to use **LPT** instead.

LPT includes axioms $c\downarrow$ for all constants c of any theory formulated in **LPT**; this is in accordance with the philosophy that terms denote things, and while terms may fail to denote (as in "the King of France"), there is no such thing as a non-existent thing. Thus $1/0$ can be undefined, i.e. fail to denote, but if a constant ∞ is used in **LPT**, it must denote something.

The meaning of $t = s$ is that t and s are both defined and they are equal. We write $t \cong s$ to express that if one of t or s is defined, then so is the other, and they are equal. **LPT** contains the axioms of "strictness", which are as follows (for each function symbol f and relation symbol R in the language):

$$f(t_1, \ldots, t_n)\downarrow \supset t_1\downarrow \wedge \ldots \wedge t_n\downarrow$$
$$R(t_1, \ldots, t_n) \supset t_1\downarrow \wedge \ldots \wedge t_n\downarrow$$

The following axioms express the meaning of the five main function symbols of **ECG**. They do not, however, make any assertions of geometrical content, nor do they distinguish one intersection point (of a line and circle, or of two circles) from the other.

$P = \mathit{IntersectLines}(L, K) \supset \mathit{on}(P, L) \wedge \mathit{on}(P, K)$ (Axiom 5)
$\mathit{IntersectLines}(L, K) \cong \mathit{IntersectLines}(K, L)$ (Axiom 6)
$P = \mathit{IntersectLineCircle1}(L, C) \supset \mathit{on}(P, L) \wedge \mathit{On}(P, C)$ (Axiom 7)
$P = \mathit{IntersectLineCircle2}(L, C) \supset \mathit{on}(P, L) \wedge \mathit{On}(P, C)$ (Axiom 8)
$P = \mathit{IntersectCircles1}(C, K) \supset \mathit{On}(P, C) \wedge \mathit{On}(P, K)$ (Axiom 9)
$P = \mathit{IntersectCircles2}(C, K) \supset \mathit{On}(P, C) \wedge \mathit{On}(P, K)$ (Axiom 10)
$\mathit{on}(P, L) \wedge \neg \mathit{on}(P, K) \supset \mathit{IntersectLines}(L, K) \downarrow$ (Axiom 11)
$\mathit{IntersectLines}(L, K) \downarrow \wedge \mathit{on}(P, L) \wedge \mathit{on}(P, K) \supset P = \mathit{IntersectLines}(L, K)$ (Axiom 12)
$\mathit{on}(P, L) \wedge \mathit{On}(P, C) \supset \mathit{IntersectLineCircle1}(L, C) \downarrow$ (Axiom 13)
$\mathit{on}(P, L) \wedge \mathit{On}(P, C) \supset \mathit{IntersectLineCircle2}(L, C) \downarrow$ (Axiom 14)
$\mathit{On}(P, C) \wedge \mathit{On}(P, K) \supset \mathit{IntersectCircles1}(C, K) \downarrow$ (Axiom 15)
$\mathit{On}(P, C) \wedge \mathit{On}(P, K) \supset \mathit{IntersectCircles2}(C, K) \downarrow$ (Axiom 16)

Axioms to distinguish between the two intersection points in each case will be given below, but that must await further developments.

In order to rule out "degenerate" lines, we need an axiom saying that they don't exist; but we *do* allow degenerate circles. The following axioms also provide for lines and circles to exist when they ought to.

$\mathit{Line}\ (A, B) \downarrow \leftrightarrow A \neq B$ (Axiom 17)
$\mathit{Circle}\ (A, B) \downarrow$ (Axiom 18)

There are functions symbols corresponding to the constructor and accessor functions for each of the sorts. The argument and value types of these symbols are obvious, and hence not specified here. Here are the axioms (20 through 27) relating the constructors and accessors.

$\mathit{Line}\ (\mathit{pointOn1}(L), \mathit{pointOn2}(L)) = L$ (Axiom 19)
$A \neq B \wedge \mathit{pointOn1}(\mathit{Line}\ (A, B)) = A$ (Axiom 20)
$A \neq B \wedge \mathit{pointOn2}(\mathit{Line}\ (A, B)) = B$ (Axiom 21)
$\mathit{pointOn1}(L) \neq \mathit{pointOn2}(L)$ (Axiom 22)
$\mathit{Circle}\ (\mathit{center}(C), \mathit{pointOnCircle}(C)) = C$ (Axiom 23)
$\mathit{center}(\mathit{Circle}\ (A, B)) = A$ (Axiom 24)
$\mathit{pointOnCircle}(\mathit{Circle}\ (A, B)) = B$ (Axiom 25)
$\mathit{center}(C) \neq \mathit{pointOnCircle}(C)$ (Axiom 26)

For readers unfamiliar with the logic of partial terms, we point out that Axiom 20 could have been written

$$\mathit{pointOn1}(\mathit{Line}\ (A, B)) \cong A$$

where the relation $t \cong s$ means that if either side is defined, both sides are defined, and they are equal. Also, written the way it is written, with equality instead of $\cong$, Axiom 20 implies that $\mathit{Line}\ (A, B)$ is defined when $A \neq B$, making half of Axiom 17 superfluous.

There are two incidence relations, for $\mathit{on}(P, L)$ for points lying on lines, and $\mathit{On}(P, C)$ for points lying on and circles. There are three constants α, β, and γ with axioms saying that these three points are non-collinear. Specifically

$\neg \mathit{on}(\alpha, \mathit{Line}\ (\beta, \gamma))$ (Axiom 27)
$\neg \mathit{on}(\beta, \mathit{Line}\ (\alpha, \gamma))$ (Axiom 28)
$\neg \mathit{on}(\gamma, \mathit{Line}\ (\alpha, \beta))$ (Axiom 30)

We do not have to say explicitly that $\alpha \downarrow$, because it is part of the logic of partial terms that every constant is defined–it is nothing special to any particular theory.
The other axioms of incidence are

$\mathit{on}(A, \mathit{Line}\ (A, B))$ (Axiom 30)
$\mathit{on}(B, \mathit{Line}\ (A, B))$ (Axiom 31)
$\mathit{on}(P, L) \wedge \mathit{on}(Q, L) \wedge \mathit{on}(R, \mathit{Line}\ (P, Q)) \supset \mathit{on}(R, L)$ (Axiom 32)

We do not need $\mathit{On}(B, \mathit{Circle}(A, B))$ because that will follow axiom 44 below.

We use the equality symbol between points to mean "identically equal". Between lines, equality means "intensional equality". In the spirit of constructive mathematics, lines "come equipped" with two associated distinct points. Thus, *Line* $(A, B) =$ *Line* (P, Q) if and only if $A = B$ and $P = Q$. This does not need to be assumed, as it follows from the axioms given above for the accessor and constructor functions. It may, however, be confusing to those not accustomed to constructive mathematics. The notion of "extensional equality" refers to the defined relation between two lines, that the same points are on both lines. In practice, to avoid confusion, we rarely if ever mention equality between lines. It should be noted, however, that our theory does depend on this view of lines, since the order of the two intersection points of line *Line* (A, B) with circle C is opposite to the order of the two intersection points of *Line* (B, A) with C. In other words, when considering *IntersectLineCircle1*(L, C), it is essential that L is given by two points.

The basic relations of our theories are equidistance and betweenness, which have been recognized as fundamental at least since Hilbert's famous 1899 book [10]. All the arguments of these two relations have sort *Point*. The (strict) betweenness relation is written $\mathbf{B}(a, b, c)$. We read this "b is between a and c". It implies that the three points are collinear. The first betweenness axiom is

$$\mathbf{B}(a, b, c) \supset \mathbf{B}(c, b, a) \qquad \text{(Axiom 33)}$$

Before giving a constructive version of the remaining betweenness axioms, we discuss a related principle. By "Markov's principle for betweenness" we mean

$$\neg\neg\mathbf{B}(A, B, C) \supset \mathbf{B}(A, B, C) \qquad \text{(Axiom 34)}$$

Markov's principle expresses the idea that by computing two points P and Q etc. to greater and greater accuracy, if they are not identical we will eventually find that out. We want it to be provable in **ECG** for several reasons: it is used in Euclid (e.g. I.6 and I.26, as is discussed below); it is needed for some fundamental theorems (see for example Lemma 7); and it makes for a smooth metatheory. While some may consider Markov's principle in number theory to be of questionable constructivity, we consider that geometry without Markov's principle is awkward. As we shall see in other sections, Euclid does use it, and including it does not harm our ability to construct things that are proved to exist. In terms of order, it expresses the principle that $\neg x \leq 0 \supset x > 0$.

Hilbert's second axiom for betweenness is, "given three distinct points, one and only one of the points is between the other two." That formulation is too strong, constructively. (For example its translation into the recursive plane is not provable in **HA** plus Markov's principle.) Instead, we consider the following version, which says that for three distinct points, if two of the alternatives fail then the third must hold, and no two can hold.

$$\begin{array}{l} a \neq b \wedge a \neq c \wedge b \neq c \supset \qquad \text{(Axiom 35*)} \\ \quad (\neg\mathbf{B}(a, b, c) \wedge \neg\mathbf{B}(b, c, a) \supset \mathbf{B}(c, a, b)) \wedge \\ \quad (\neg\mathbf{B}(b, c, a) \wedge \neg\mathbf{B}(c, a, b) \supset \mathbf{B}(a, b, c)) \wedge \\ \quad (\neg\mathbf{B}(c, a, b) \wedge \neg\mathbf{B}(a, b, c) \supset \mathbf{B}(b, c, a)) \wedge \\ \quad \neg(\mathbf{B}(a, b, c) \wedge \mathbf{B}(b, c, a)) \wedge \neg(\mathbf{B}(a, b, c) \wedge \mathbf{B}(b, a, c)) \wedge \\ \quad \neg(\mathbf{B}(b, c, a) \wedge \mathbf{B}(b, a, c)) \end{array}$$

In the presence of Markov's principle for betweenness (Axiom 34), we can weaken Axiom 35* by replacing the unnegated betweenness formulas by their double negations. We obtain the following:

$$\begin{array}{l} a \neq b \wedge a \neq c \wedge b \neq c \supset \qquad \text{(Axiom 35)} \\ \quad (\neg\mathbf{B}(a, b, c) \wedge \neg\mathbf{B}(b, c, a) \supset \neg\neg\mathbf{B}(c, a, b)) \wedge \\ \quad (\neg\mathbf{B}(b, c, a) \wedge \neg\mathbf{B}(c, a, b) \supset \neg\neg\mathbf{B}(a, b, c)) \wedge \\ \quad (\neg\mathbf{B}(c, a, b) \wedge \neg\mathbf{B}(a, b, c) \supset \neg\neg\mathbf{B}(b, c, a)) \wedge \\ \quad \neg(\mathbf{B}(a, b, c) \wedge \mathbf{B}(b, c, a)) \wedge \neg(\mathbf{B}(a, b, c) \wedge \mathbf{B}(b, a, c)) \wedge \\ \quad \neg(\mathbf{B}(b, c, a) \wedge \mathbf{B}(b, a, c)) \end{array}$$

In fact, Axioms 34 and 35 together are exactly equivalent to Axiom 35*, as the following lemma shows. We separate Axiom 35* into two axioms to facilitate the discussion of different verions of Euclid's parallel postulate, whose relations depend on Markov's principle.

Lemma 6 *Markov's principle for betweenness (Axiom 34) is provable from Axiom 35*.*

Proof. Suppose $\neg\neg\mathbf{B}(a,x,b)$. We want to prove $\mathbf{B}(a,x,b)$. From $\neg\neg\mathbf{B}(a,x,b)$ we immediately have $a \neq b$ and $a \neq x$ and $x \neq b$. By Axiom 35* it suffices to prove $\neg\mathbf{B}(x,a,b)$ and $\neg\mathbf{B}(a,b,x)$. To prove $\neg\mathbf{B}(x,a,b)$, suppose $\mathbf{B}(x,a,b)$. Then $\neg\mathbf{B}(a,x,b)$, by Axiom 35*. But that contradicts $\neg\neg\mathbf{B}(a,x,b)$. That proves $\neg\mathbf{B}(x,a,b)$. Similarly we have $\neg\mathbf{B}(a,b,x)$. That completes the proof of the lemma.

There is a betweenness axiom that says, in Hilbert's formulation, that *Line* (P,Q) contains a point between P and Q. We call this the "density" axiom. Since we want a quantifier-free axiomatization, we would like to specify the point asserted to exist. The natural candidate for a point between P and Q is the result of Euclid's segment-bisection construction. We therefore take the following axiom:

$$\mathbf{B}(P, \mathit{IntersectLines}(\mathit{Line}(P,Q), \mathit{Line}(\mathit{IntersectCircles1}(\mathit{Circle}(P,Q), \mathit{Circle}(Q,P)), \mathit{IntersectCircles2}(\mathit{Circle}(P,Q), \mathit{Circle}(Q,P)))), Q) \quad \text{(Axiom 36)}$$

Note that, in view of the strictness axioms of **LPT**, this axiom implies that the circles and intersection points involved are defined. Similarly, there is a betweenness axiom that asserts that *Line* (A,B) contains points outside the segment AB. Again, we want to specify such points so that our axiomatization is quantifier-free:

$$\mathbf{B}(\mathit{IntersectLineCircle1}(\mathit{Line}(P,Q), \mathit{Circle}(P,Q)), P, Q) \quad \text{(Axiom 37)}$$
$$\mathbf{B}(P, Q, \mathit{IntersectLineCircle2}(\mathit{Line}(P,Q), \mathit{Circle}(Q,P))) \quad \text{(Axiom 38)}$$

The remaining betweenness axiom is called the "plane separation axiom". To make its statement more readable, we introduce the usual definitions of two points P and Q being on opposite sides of, or on the same side of, line L:

$$\mathit{OppositeSide}(P,Q,L) := \mathbf{B}(P,Q, \mathit{IntersectLines}(\mathit{Line}\ (P,Q), L) \quad \text{(Definition 39)}$$
$$\mathit{SameSide}(P,Q,L) := \neg\mathbf{B}(P,Q, \mathit{IntersectLines}(\mathit{Line}\ (P,Q), L)) \quad \text{(Definition 40)}$$

When we use the symbol :=, we mean that the symbol on the left is regarded as an abbreviation at the meta-level, rather than a symbol of the formal language. When it is used in subsequent formulas, it stands for the formal equivalent given by the right hand side.

Note that if *Line* (P,Q) does not meet L, then the argument of B is undefined, so by the strictness axioms P and Q are on the same side of L. This formulation, however, does not require us to be able to decide whether L is or is not parallel to *Line* (P,Q). Using these definitions we can give the plane separation axiom(s):

$$\mathit{SameSide}(A,B,L) \wedge \mathit{SameSide}(B,C,L) \supset \mathit{SameSide}(A,C,L) \quad \text{(Axiom 41)}$$
$$\mathit{OppositeSide}(A,B,L) \wedge \mathit{OppositeSide}(B,C,L) \supset \mathit{SameSide}(A,C,L) \quad \text{(Axiom 42)}$$

Rays. Although we have not included rays in **ECG**, we do want to support our claim that a conservative extension including rays can easily be introduced; and also, we sometimes make informal arguments using rays with the implication that they can be formalized in **ECG**. We now show that the use of intuitionistic logic does not cause a problem about incidence on rays or segments. Using betweenness, we can define incidence for rays. However, there is a technicality: the origin O of *Ray* (O,B) is considered to lie on the ray, i.e. "rays are closed", while betweenness means "strictly between." It is thus easier to define the "opposite ray": P is on the opposite ray to *Ray* (O,B) if P is on *Line* (O,B) and O is between P and B. Then Q is on *Ray* (O,B) if it is on *Line* (O,B) but not on the opposite ray:

$$\mathit{on}(Q, \mathit{Ray}\ (O,B)) := \mathit{on}(Q, \mathit{Line}\ (O,B)) \wedge \neg\mathbf{B}(P,O,B) \quad \text{(Definition 43)}$$

This definition can be used to express informal arguments about rays in **ECG** without needing to introduce an explicit sort and axioms for rays.

Segments. In a similar way we can define incidence for segments, so that "segments are closed".

$$On(P, Segment\,(Q,R)) \leftrightarrow On(P, Line\,(Q,R)) \wedge \neg \mathbf{B}(P,Q,R) \wedge \neg \mathbf{B}(Q,R,P)$$

Recall that $Segment\,(R,R)$ and $Line\,(R,R)$ are both undefined, so the fact that $\neg\mathbf{B}(P,P,P)$ does not make P lie on $Segment\,(P,P)$. We do not number this definition, since it is not used in any further axioms or proofs.

The equidistance relation is written $\delta(A,B,C,D)$. We will often express this using the informal notation $AB = CD$. Axioms for equidistance are sometimes called "congruence axioms" since equidistance can be thought of as congruence of segments. Sometimes, following a tradition that goes back to Euclid, we write $AB = CD$ instead of $\delta(A,B,C,D)$.

Using equidistance, we define incidence for circles:

$on\,(P, Circle\,(A,Q)) \leftrightarrow \delta(A,P,A,Q) \wedge A \neq Q$ (Axiom 44)

The $A \neq Q$ part is needed to avoid conflict with our axiom that $Circle\,(A,Q)$ is undefined if $A = Q$.

Remark. Again, we want to show that, if desired, arcs can be correctly handled in a natural extension of **ECG**:

$$\begin{aligned} on\,(P, Arc\,(A,C,Q)) \quad := \quad & On(P, Circle\,(C,A)) \;\wedge\; \delta(A,C,C,Q) \\ & \wedge\; on\,(IntersectLines(Line\,(A,Q), Line\,(C,P)), Segment\,(A,Q)) \end{aligned}$$

If any two of the three points A, C, and Q are equal, we have undefined terms both on the left and the right. Note that this definition makes arcs "closed", in that A and Q will be on $Arc\,(A,C,Q)$, because segments are closed. No further use of arcs will be made in this paper.

Greenberg's first congruence axiom, paraphrased from [8], is closely related to the uniform version of Euclid's Book I, proposition 2:

$$A \neq B \wedge C \neq D \supset \exists R(on\,(R, Ray\,(A,B)) \wedge \delta(A,R,C,D) \qquad \text{(not an axiom of } \mathbf{ECG}\text{)}$$

This axiom permits us to "lay off" segment CD along $Ray\,(A,B)$. Since we are seeking a quantifier-free axiomatization, we want to specify the point R. This we do by taking

$$R(A,B,C,D) = IntersectLineCircle1(A,B, Circle3(A,C,D)).$$

Our version of the axiom is thus

$$A \neq B \wedge C \neq D \supset (on\,(R(A,B,C,D), Ray\,(A,B)) \wedge \delta(A, R(A,B,C,D), C, D)$$

Note that the strictness axioms then will imply that $R(A,B,C,D)$ is defined when $A \neq B$ and $C \neq D$. The official version, with R replaced by is definition, is Axiom 45:

$$\begin{aligned} & A \neq B \wedge C \neq D \supset \\ & \quad (on\,(IntersectLineCircle1(Line\,(A,B), Circle3\,(A,C,D)), Ray\,(A,B)) \wedge \\ & \quad \delta(A, IntersectLineCircle1(Line\,(A,B), Circle3\,(A,C,D)), C, D) \end{aligned}$$

The second congruence axiom is

$\delta(A,B,C,D) \wedge \delta(A,B,E,F) \supset \delta(C,D,E,F)$ (Axiom 46)

The third congruence axiom can be thought of as saying that addition is well-defined on congruence classes of segments:

$\mathbf{B}(A,B,C) \wedge \mathbf{B}(P,Q,R) \wedge \delta(A,B,P,Q) \wedge \delta(B,C,Q,R) \supset \delta(A,C,P,R)$ (Axiom 47)

With the aid of the congruence axioms considered so far, we can formalize the notion "$AB < CD$". This is defined to mean that the point B' on ray CD such that $CB' = AB$ lies between C and D. More formally, we have to define B' by a term:

$$AB < CD := \mathbf{B}(C, \mathit{IntersectLineCircle2}(\mathit{Line}\ (C,D), \mathit{Circle3}(C,A,B)), D) \qquad \text{(Definition 48)}$$

In constructive mathematics, we cannot define $x \le y$ as $x < y \vee x = y$. Instead, we define $x \le y$ as $\neg y < x$. Writing out the definition of $y < x$ to make the definition of $x \le y$ directly, we have

$$CD \le AB := \neg\mathbf{B}(C, \mathit{IntersectLineCircle2}(\mathit{Line}\ (C,D), \mathit{Circle3}(C,A,B)), D) \quad \text{(Definition 49)}$$

and as remarked earlier we also use the traditional abbreviation

$$AB = CD := \delta(A,B,C,D) \qquad \text{(Definition 50)}$$

We have not included angles as a fundamental data type. Instead, statements about angles can be formalized as statements about two non-collinear rays with the same origin, or about three distinct points. Angles are thus always less than π–there is no such thing as a "straight angle". In this we follow Greenberg [8]. But Greenberg takes congruence of angles as a fundamental notion. Instead, we define it, essentially using the principle SAS to do so. Given three points (thought of as an angle) ABC, and three points PQR, by the first congruence axiom we can find P' on ray QP and R' on ray QR with $QP' = BA$ and $QR' = BC$. Then we define angle ABC to be congruent to angle PQR if and only if $AC = P'Q'$. This is a 6-ary relation between points. Note that it can be expressed in quantifier-free, disjunction-free form, since P' and R' are given by terms.

From the definition of congruence it follows that if A' distinct from B lies on $\mathit{Ray}\ (B,A)$ and C' distinct from B lies on $\mathit{Ray}\ (B,C)$, then angle $A'BC'$ is congruent to angle ABC. This is essentially the reflexivity of congruence viewed as relation between rays; that is the first half of Greenberg's Congruence Axiom 5, which becomes unnecessary. The same observation can be used to prove the symmetry of angle congruence. The second half of Congruence Axiom 5 is the transitivity of congruence of angles, which we take as our fifth congruence axiom, formulated with nine point variables instead of three angle variables. It really comes down to the transitivity of the congruence of triangles. We want to say that if angle ABC is congruent to triangle PQR and triangle PQR is congruent to triangle UVW then triangle ABC is congruent to triangle UVW. The congruence of triangle ABC and triangle PQR is expressed by $AB = PQ \wedge BC = QR \wedge AC = PR$. Hence the transitivity axiom we need is

$$\begin{aligned} &AB = PQ \wedge BC = QR \wedge AC = PR \wedge PQ = UV \wedge QR = VW \wedge PR = UV \supset \qquad \text{(Axiom 51)} \\ &\quad AB = UV \wedge BC = VW \wedge AC = UW \end{aligned}$$

We want to show that congruence of angles really only depends on the four rays involved, not on the six points. To that end we suppose that A' distinct from B lies on $\mathit{Ray}\ (B,A)$ and C' distinct from B lies on $\mathit{Ray}\ (B,C)$, angle ABC is congruent to angle PQR, and P' distinct from Q lies on $\mathit{Ray}\ (Q,P)$, and R' distinct from R lies on $\mathit{Ray}\ (Q,R)$. We wish to show that angle $A'BC'$ is congruent to angle $P'QR'$. Let P_1 on $\mathit{Ray}\ (Q,P)$ have $BA = QP_1$, and R_1 on $\mathit{Ray}\ (Q,R)$ have $BC = QR_1$. Then angle $A'BC'$ is congruent to angle ABC (as observed above), which is congruent to angle PQR by hypothesis, which is congruent to angle $P'QR'$. Hence by the transitivity of congruence, we are finished.

The sixth congruence axiom is the SAS criterion for triangle congruence. With our definition of congruence for angles, this axiom is provable. In Greenberg's system, it simply serves in place of a definition of angle congruence.

Greenberg's fourth congruence axiom states that for any angle BAC and any ray $A'B'$ there is a unique ray C' on a given side of $A'B'$ such that angle $B'A'C'$ is congruent to angle BAC. Of course rays are not necessary here: this is a statement that for any five points satisfying certain conditions, there exists another point C' satisfying a certain condition. To make precise the part about "on a given side of", we have to mention another point P not on $A'B'$ and demand that PC' should not meet $\mathit{Line}\ (A',B')$. To express this in a quantifier-free way, we need to construct the point C' in question. This we do as follows:

$C = \mathit{Circle}(A', A, B)$
$B'' = \mathit{IntersectLineCircle1}(\mathit{Line}\ (A', B'), C)$ so B'' lies on the ray $A'B'$
$K_1 = \mathit{Circle3}(B'', B, C)$
$K_2 = \mathit{Circle3}(A'', A, C)$
$C' = \mathit{IntersectCircles1}(K_1, K_2)$

Now the desired point C' is one of the intersection points of K_1 and K_2. There seems to be no reason, based on the axioms given so far, why these circles intersect. We take as our fourth congruence axiom, the assertion that both their intersection points are defined. Writing it out formally we have

$$A \neq B \wedge A \neq C \wedge B \neq C \wedge A' \neq B' \wedge \qquad \text{(Axiom 52)}$$
$$B'' = \mathit{IntersectLineCircle1}(\mathit{Line}\ (A', B'), C) \wedge$$
$$K_1 = \mathit{Circle3}(B'', B, C)) \wedge K_2 = \mathit{Circle3}(A'', A, C)) \supset$$
$$\mathit{IntersectCircles1}(K_1, K_2) \downarrow \wedge \mathit{IntersectCircles2}(K_1, K_2) \downarrow \wedge$$
$$\mathit{OppositeSide}(\mathit{IntersectCircles1}(K_1, K_2), \mathit{IntersectCircles2}(K_1, K_2), \mathit{Line}\ (A', B'))$$

Of course one can use fewer variables and more complicated terms to express this axiom, eliminating the variables B'', K_1, and K_2, at the cost of human legibility. Now one of the two intersection points asserted to be defined will be the point needed to verify Greenberg's fourth congruence axiom, since the last line of Axiom 51 says that the two points are on opposite sides of $A'B'$.[7] Hence, by Betweenness Axiom 3, one of them is on the same side as P.

Having defined $AB < CD$, we are in a position to formulate the axioms of line-circle continuity. The first two of these just tell us when a line and circles intersect–namely, when there is a point on the line closer (or equally close) to the center than the radius of the circle.

$$AP \leq AB \wedge \mathit{on}(P, L) \supset \mathit{IntersectLineCircle1}(L, \mathit{Circle}\ (A, B)) \downarrow \qquad \text{(Axiom 53)}$$
$$AP \leq AB \wedge \mathit{on}(P, L) \supset \mathit{IntersectLineCircle2}(L, \mathit{Circle}\ (A, B)) \downarrow \qquad \text{(Axiom 54)}$$

Our next axiom says that the intersection points depend "extensionally" on the circle. That is, if two circles contain the same points (which is guaranteed if they have the same center and radius), then their first and second intersection points with any line are the same. Note that the intersection points depend "intensionally" on the line, because the first and second intersection points of *Line* (A, B) with circle C are the second and first intersection points of *Line* (B, A) with C. But the intersection points depend extensionally on the circle:

$$A = \mathit{center}(C) \wedge A = \mathit{center}(K) \wedge \mathit{On}(P, C) \wedge \mathit{On}(Q, K) \wedge AP = AQ \supset \qquad \text{(Axiom 55)}$$
$$\mathit{IntersectLineCircle1}(L, C) \cong \mathit{IntersectLineCircle1}(K, C) \wedge$$
$$\mathit{IntersectLineCircle2}(L, C) \cong \mathit{IntersectLineCircle2}(K, C)$$

We next give the basic axioms about intersections of two circles.

$$\mathit{On}(P, C) \wedge AP \leq AB \supset \qquad \text{(Axiom 56)}$$
$$\mathit{IntersectCircles1}(C, \mathit{Circle}\ (A, B)) \downarrow \wedge\ \mathit{IntersectCircles2}(C, \mathit{Circle}\ (A, B)) \downarrow$$

Our next axiom specifies that the intersection points of two circles depend extensionally on the circles:

$$A = \mathit{center}(C_1) \wedge A = \mathit{center}(C_1) \wedge \mathit{On}(P, C_1) \wedge \mathit{On}(Q, C_1) \wedge AP = AQ \supset \qquad \text{(Axiom 57)}$$
$$\mathit{IntersectCircles1}(C_1, K) \cong \mathit{IntersectCircles1}(C_2, K) \wedge$$
$$\mathit{IntersectCircles2}(C_1, K) \cong \mathit{IntersectCircles2}(C_2, K) \wedge$$
$$\mathit{IntersectCircles1}(K, C_1) \cong \mathit{IntersectCircles1}(K, C_2) \wedge$$
$$\mathit{IntersectCircles2}(K, C_1) \cong \mathit{IntersectCircles2}(K, C_2) \wedge$$

We now come to the expression of Euclid's parallel postulate. We first define $\mathit{Parallel}(L, K)$ for lines L and K to mean that the lines do not meet:

$$\mathit{Parallel}(L, K) := \neg \mathit{IntersectLines}(K, L) \downarrow .$$

[7] We do not know if it is necessary to take this assertion as part of the axiom. Perhaps it can be proved.

Of course, in view of the other axioms for *IntersectLines*, we have

$$Parallel(L,K) \leftrightarrow \forall x \neg(on(x,L) \wedge on(x,K)),$$

but the form we took as the definition has the advantage of being quantifier-free. Most modern treatments of geometry formulate the parallel axiom in this way: if two lines K and M are parallel to L through point p, then $K = M$. In symbols:

$$Parallel(K,L) \wedge Parallel(M,L) \wedge on(p,K) \wedge on(p,M) \supset K = M \quad \text{(Playfair's Postulate)}$$

We call this the "Playfair's postulate", or for short just "Playfair", after John Playfair, who published it in 1795, although (according to Greenberg [8], p. 19) it was referred to by Proclus. 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 do not, however, take this form of Euclid's parallel postulate as an axiom. Instead we take the following axiom:

$$\neg IntersectLines(K,L) \downarrow \wedge on(p,K) \wedge on(p,M) \wedge M \neq K \supset IntersectLines(L,M)$$

(Axiom 58, the parallel postulate)

In other words, if K is parallel to L through P, then any other line M through P must meet L. This differs from Euclid's version in that we are not required to know in what *direction* M passes through P; but also the conclusion is weaker, in that it does not specify *where* M must meet L. The relationships between these different parallel axioms are discussed in section 13 below and in [3].

To illustrate our reasons for including Markov's principle in **ECG**, we exhibit the following lemma.

Lemma 7 *(In* **ECG***) Suppose neither point A nor point B lies on line T. (i) If A is not on the same side of T as B, then A is on the opposite side of T from B. (ii) If A is not on the opposite side of line T from B, then A is on the same side of T as B.*

Proof. Ad (i). Suppose A is not on the same side of line T as B. Then segment AB not not meets T. By Markov's principle for line intersections (which is a theorem of **ECG**), line AB meets T in some point q. Since segment AB not not meets T, point q is not not in segment AB. Since A and B do not lie on T, q is not not between A and B. Then, by Markov's principle for betweenness, the intersection point of AB and T lies between A and B. That is, A is on the opposite side of T from B, proving (i).

Ad (ii). Suppose that A is not on the opposite side of line T as B. To show A is on the same side of T as B, we must show that segment AB does not meet T. Suppose that segment AB does meet T. Since A and B do not lie on T, the point of intersection Q is not equal to A or B. Hence, by Markov's principle for betweenness, Q lies between A and B. Hence A and B are on opposite sides of T, contradiction. That completes the proof of the lemma.

We now have given a quantifier-free, disjunction-free axiom system that enables us to verify Greenberg's axioms, after translating angles as triples of points. But this axiom system has a shortcoming: It does not distinguish which of the two intersection points of two circles are defined by *IntersectCircles1* and *IntersectCircles2*, and we have not completely specified which intersection point of a line and circle is which, either, although we have done so when the line is a diameter of the circle, by specifying that $IntersectLineCircle1(Line(A,B), Circle(A,B))$ is on the same side of A as B. Hence the axioms given up to now do not suffice to prove the continuity of *IntersectCircles1* and *IntersectCircles2*. To put this matter another way, the axioms given so far have models in which *IntersectLineCircle1* and *IntersectCircles1* are discontinuous; indeed arbitrarily discontinuous. Given one model, we can arbitrary switch some of the values of $IntersectLineCircle1(P,Q)$ and $IntersectLineCircle2(P,Q)$, and we still have a model.

We wish to add new axioms in such a way that the constructions defined by these terms are continuous. Our first approach to this was based on the observations that explicit moduli of continuity can be defined for the intersection points of circles and lines, and the intersection points of circles and circles. The latter involve $\delta = \sqrt{\epsilon}$, so they are constructible using straightedge and compass. However, the axioms required are neither short nor elegant, and we discovered an axiomatization that, while still not very short, does qualify as "elegant". The axiomatization is based on a concept that is fundamental to computer graphics, although it does not receive much attention in classical geometry. Namely, the concept that an angle PQR is a "left turn" or a "right turn". Analytically, the quantity in question is the sign of the vector cross product of the vectors QP and QR. What we will exhibit is a geometric axiomatization of this concept.

We could introduce a new predicate *Left* (A, B, C), but there is no need to do that, since we can just define, for A, B, and C not collinear,

$$\mathit{Left}\ (A,B,C) := C = \mathit{IntersectCircles1}(\mathit{Circle}\ (A,C), \mathit{Circle}\ (B,C)) \quad \text{(Definition 59)}$$
$$\mathit{Right}\ (A,B,C) := C = \mathit{IntersectCircles2}(\mathit{Circle}\ (A,C), \mathit{Circle}\ (B,C)) \quad \text{(Definition 60)}$$

If we did introduce new symbols for *Left* and *Right* then these would be axioms instead of definitions. We turn to the axioms about *Left* and *Right*. Recall that α, β, and γ are (constant symbols for) three arbitrary distinct points.

$$\mathit{Left}\ (\alpha,\beta,\gamma) \quad \text{(Axiom 61)}$$
$$\mathit{Right}\ (\alpha,\gamma,\beta) \quad \text{(Axiom 62)}$$

These axioms arbitrarily specify the orientation of the plane. The next axiom says that "handedness" is a property of the rays involved, not just the points:

$$P \neq P' \wedge R \neq R' \wedge on(P', \mathit{Ray}\ (Q,P)) \wedge on(R', \mathit{Ray}\ (Q,R)) \wedge \mathit{Left}\ (P,Q,R) \supset \mathit{Left}\ (P',Q',R') \quad \text{(Axiom 63)}$$

The next axiom says that if PQR is a left turn, and we move P (in any direction) without crossing line QR, it is still a left turn:

$$\mathit{Left}\ (P,Q,R) \wedge \neg\mathbf{B}(P, \mathit{IntersectLines}(\mathit{Line}\ (Q,R), \mathit{Line}\ (P,P')), P') \supset \mathit{Left}\ (P',Q,R) \quad \text{(Axiom 64)}$$

and similarly if we move R without crossing PQ:

$$\mathit{Left}\ (P,Q,R) \wedge \neg\mathbf{B}(R, \mathit{IntersectLines}(\mathit{Line}\ (Q,P), \mathit{Line}\ (R,R')), P') \supset \mathit{Left}\ (P,Q,R') \quad \text{(Axiom 65)}$$

Together these axioms permit us to rotate the sides of a left turn PQR as long as they do not coincide or become opposite, and it remains a left turn.

The next axiom permits us to perform a translation:

$$\mathit{Left}\ (A,B,C) \wedge AB = PQ \wedge BC = QR \wedge AC = PR \wedge AP = BQ \wedge AP = CR \supset \mathit{Left}\ (P,Q,R) \quad \text{(Axiom 66)}$$

Those are all the axioms for *Left* . Here are similar axioms for *Right* :

$$P \neq P' \wedge R \neq R' \wedge on(P', \mathit{Ray}\ (Q,P)) \wedge on(R', \mathit{Ray}\ (Q,R)) \wedge \mathit{Right}\ (P,Q,R) \supset \mathit{Right}\ (P',Q',R') \quad \text{(Axiom 67)}$$
$$\mathit{Right}\ (P,Q,R) \wedge \neg\mathbf{B}(P, \mathit{IntersectLines}(\mathit{Line}\ (Q,R), \mathit{Line}\ (P,P')), P') \supset \mathit{Right}\ (P',Q,R) \quad \text{(Axiom 68)}$$
$$\mathit{Right}\ (P,Q,R) \wedge \neg\mathbf{B}(R, \mathit{IntersectLines}(\mathit{Line}\ (Q,P), \mathit{Line}\ (R,R')), P') \supset \mathit{Right}\ (P,Q,R') \quad \text{(Axiom 69)}$$
$$\mathit{Right}\ (A,B,C) \wedge AB = PQ \wedge BC = QR \wedge AC = PR \wedge AP = BQ \wedge AP = CR \supset \mathit{Right}\ (P,Q,R) \quad \text{(Axiom 70)}$$

Repeated applications of Axioms 64 and 65 permit us to perform an arbitrary rotation on a left turn PQR, preserving the fact that it is a left turn. The reader who wishes to understand the motivation for these axioms about *Right* and *Left* should see Lemma 8 and its proof, below.

With the aid of *Right* and *Left* , we define the concept "P and Q have the same order on line L as A and B", constructively and without needing case distinctions. Of course, we assume

$P \neq Q$ and $A \neq B$. First, we construct point E such that ABE is a left turn. By Definition 59, $E = \mathit{IntersectCircles1}(\mathit{Circle}\ (A,B), \mathit{Circle}\ (B,A))$ is such a point. Then P and Q have the same order as A and B if and only if $\mathit{Left}\ (P,Q,E)$. Formally,

$$\begin{aligned}&\mathit{SameOrder}\ (A,B,P,Q) := \qquad \text{(Definition 71)}\\ &\quad A \neq B \wedge P \neq Q \wedge \mathit{on}(P, \mathit{Line}\ (A,B)) \wedge \mathit{on}(Q, \mathit{Line}\ (A,B)) \wedge\\ &\quad \mathit{Left}\ (P,Q,\mathit{IntersectCircles1}(\mathit{Circle}\ (A,B), \mathit{Circle}\ (B,A)))\end{aligned}$$

With *SameOrder* in hand, it is easy to distinguish the two intersection points of a circle. However, we must be careful to allow for the case when the two intersection points coincide.

$$\begin{aligned}&P = \mathit{IntersectLineCircle1}(\mathit{Line}\ (A,B), C) \wedge \qquad \text{(Axiom 72)}\\ &Q = \mathit{IntersectLineCircle2}(\mathit{Line}\ (A,B), C) \wedge P \neq Q\\ &\supset \mathit{SameOrder}\ (A,B,P,Q)\end{aligned}$$

Next we give the remaining axioms for *IntersectCircles1* and *IntersectCircles2*. We want to say essentially that if P and Q are the two intersection points of circles C and K with centers A and B respectively, then ABP is a left turn and ABQ is a right turn. But there is also the possibility that the two circles are tangent, and the two intersection points coincide. Then neither ABP nor ABQ is a left or right turn. Therefore, instead of saying ABP is a left turn and ABQ is a right turn, we say that ABP is not a right turn, and ABQ is not a left turn. Here are the axioms in question:

$$R = \mathit{IntersectCircles1}(\mathit{Circle}\ (A,P), \mathit{Circle}\ (B,Q)) \supset \neg \mathit{Right}\ (A,B,R) \qquad \text{(Axiom 73)}$$
$$R = \mathit{IntersectCircles2}(\mathit{Circle}\ (A,P), \mathit{Circle}\ (B,Q)) \supset \neg \mathit{Left}\ (A,B,R) \qquad \text{(Axiom 74)}$$

This completes our list of axioms of **ECG**. Note that these axioms are all quantifier-free and disjunction-free. We will consider one more axiom, which does contain disjunction, as a possible addition to **ECG**, in the next section.

Lemma 8 *(In* **ECG***) Let ABC be any triangle. Then we can determine the handedness of the turn ABC, in the following sense. Let α, β, and γ be the three fixed non-collinear points mentioned in the axioms of* **ECG***, so that $\alpha\beta\gamma$ is a left turn by definition. Let $L = \mathit{Line}\ (\alpha, \beta)$. Then we can construct a point R such that ABC is a left turn if and only if R is on the same side of L as γ, and a right turn if and only if R is on the opposite side of L from γ.*

Proof. The idea of the proof is this: by a series of "moves" (applications of the axioms for *Left* and *Right*, which correspond to translations, dilations, and rotations), the triangle ABC is "reduced" to either triangle $\alpha\beta\gamma$ or triangle $\alpha\gamma\beta$, preserving the handedness of the turns; but Axioms 61 and 62 directly specify the handedness of the turns $\alpha\beta\gamma$ and $\alpha\gamma\beta$. We now give the details.

By a "move", applied to a triple of non-collinear points PQR, we mean a construction of a new triple UVW such that, according to the axioms for *Left* and *Right*, if PQR is a left turn then so is UVW, and if PQR is a right turn then so is UVW. This latter condition we describe for short by saying that the move "preserves handedness". The axioms describe several types of moves that preserve handedness, specifically, moving P along the ray QP, moving R along the ray QR, rotating PR or QR in such a way that the points P,Q, and R never become collinear, and translating the whole triple. We first show that the number of moves (applications of these axioms) required to perform a given rotation is bounded by a fixed constant. Consider the following procedure: First move P to decrease the angle to less than a right angle. Then we can rotate PQR by any angle up to a right angle, using two moves (one moves PQ and one moves QR). An arbitrary rotation can be performed by performing at most four rotations of less than ninety degrees. Thus in four or fewer rotations (requiring eight or fewer moves) we can bring one of its sides onto the desired ("target") ray. Then we can move P by the same amount as in the first step resulting in the desired rotation. Hence ten or fewer applications of the above axioms suffice to perform any rotation.

Now consider three non-collinear points P, Q, and R. We give a procedure for determining whether PQR is a left turn or a right turn. First translate PQR so that Q coincides with the point β (given by a constant of **ECG**). Then rotate it so that P lies on *Ray* (β, α). Then move P to α. By the axioms above, all these steps preserve the handedness of PQR. Now, if R is on the same side of *Line* (α, β) as γ, then by Axioms 65 and 69, PQR is a left turn, since $\alpha\beta\gamma$ is a left turn by definition. And if R (after the moves described) is on the opposite side of *Line* (α, β), then we claim that PQR is a right turn. To see this, let γ' be a point on the same side of *Line* (α, β) as γ, and on the same side of *Line* (Q, R) as α. (Such a point can be constructed by bisecting the angle formed by *Ray* (β, α) and the opposite ray to *Ray* (Q, R).) Then we can move P to γ' without changing the handedness of PQR, and then we can move R to α without changing the handedness of PQR. But now PQR coincides with $\gamma\beta\alpha$, which by definition is a right turn. That completes the proof of the lemma.

Remark. With classical logic, we could prove that ABC is either a right turn or a left turn. To reach that conclusion, we would need to know that if point R is not on $L = $ *Line* (α, β), then either R is on the same side of L as γ or on the opposite side. The constructive status of this statement is discussed below.

Lemma 9 *The predicates Right* (A, B, C) *and Left* (A, B, C) *are definable in Greenberg's theory G, relative to an arbitrary choice of Left* (α, β, γ) *and Right* (α, γ, β) *for some triple of non-collinear points* α*,* β*, and* γ*. This can even be done with intuitionistic logic.*

Proof. It will suffice to define the relation $T(A, B, C, P, Q, R)$ with the meaning "ABC and PQR have the same handedness." First we note that it is possible to define the notion of one triangle being a translation of another; namely, ABC is a translation of PQR if the two triangles are congruent and $AP = BQ = CR$. It is also possible to define the notion of ABC being a rotation of PBQ (when the two angles share vertex B). This requires twenty variables to express, so it is too complex to write down intelligibly, but the definition in question says there exist twenty points representing ten "moves" according to the **ECG** axioms for rotations given above. Then ABC and PQR have the same handedness if there exist P' and Q' such that $P'BR'$ is a translation of PQR and there is a rotation $P''BR''$ of PQR with R'' on *Ray* (B, C) and P'' is on the same side of *Line* (B, C) as A.

Theorem 2 **ECG** *with classical logic is equivalent to Greenberg's system G with only line-circle and circle-circle continuity.*

Remark. Greenberg's system is not completely formal. But we understand here a weak second-order theory, with the first-order variables ranging over points, and circles, lines, segments, rays, and angles treated as sets of points.[8] Incidence means membership. There is a weak form of the comprehension axiom, only for formulas with no set quantifiers.

Proof. It suffices to deal with the point, line, and circle fragment of **ECG**, since the whole theory is conservative over this fragment (as follows either by model theory or cut-elimination from the fact that the axioms of both theories can expressed in this fragment). First we show that set variables are irrelevant in Greenberg's theory G. There are no set variables in the axioms, except those ranging over lines, circles, segments, rays, and angles. Hence, in the axioms, those variables can be replaced with variables of the corresponding sorts of **ECG**, and set membership by the appropriate incidence relations. By Gentzen's cut-elimination theorem, if a formula A without set variables has a proof in **ECG**, then there is a cut-free proof of a sequent $\Gamma \Rightarrow A$, where Γ is a conjunction of axioms. This entire proof contains no set variables, by the cut-elimination theorem; hence it is a proof in the language of **ECG**. We showed above that the axioms of G (interpreted in **ECG** in this way) are provable in **ECG**. That is one direction of the proof.

For the other direction, we define a translation A' of **ECG** into G as follows. In A', variables over lines and circles are replaced by set variables, restricted to appropriate predicates defining

[8] Technically, in Greenberg's book, lines are primitive objects and the incidence relation $on(Point, Line)$ is undefined, but rays, segments, and circles are sets of points and the incidence relations are set membership. See p. 144 for his explanation.

those concepts. Incidence relation symbols are replaced by 'ϵ'. For each term $t(x_1, \ldots x_n)$ of **ECG**, there is a formula $G_t(y, x_1, \ldots, x_n)$ that expresses $y = t(x_1, \ldots, x_n)$. The definition is given inductively on the complexity of t. The most difficult part of this is to interpret the function symbols for the two intersection points of two circles. However, the difficulty, which is defining the concepts *Left* (A, B, C) and *Right* (A, B, C) required to distinguish the two intersection points of two circles, has been taken care of in Lemma 9.

To interpret the function symbols for the intersection points of a line and a circle we need to define the concept "P and Q occur in the same order on *Line* (A, B) as A and B do." That can be done without case distinctions, simply by saying that for some R, RAB is a left turn and RPQ is a left turn. That completes the proof.

9 Euclid's Reasoning

Euclid's proofs have been analyzed in detail by Avigad *et. al.* in [1], and they conclude:

> Euclidean proofs do little more than introduce objects satisfying 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 departures 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.

It is our purpose in this section to argue that Euclid's reasoning can be supported in **ECG**, including the two types of apparently non-constructive reasoning just discussed. The type of *reductio* argument mentioned corresponds to Markov's principle $\neg\neg x = y \supset x = y$, which we have shown follows from the betweenness axioms of **ECG**. The first type (based on the idea that if *ab* and *cd* are unequal then one of them is longer) will be studied metamathematically. But first, we give two examples.

A typical example of such an argument in Euclid is Prop. I.6, whose proof begins

> Let ABC be a triangle having the angle ABC equal to the angle ACB. I say that the side AB is also equal to the side AC. For, if AB is unequal to AC, one of them is greater. Let AB be greater, ...

The same proof also uses an argument by contradiction in the form $\neg\neg x = y \supset x = y$. This principle, the "stability of equality", is included in **ECG**, and is universally regarded as constructively acceptable. The conclusion of I.6, however, is negative (has no $\exists$ or $\vee$), so we can simply put double negations in front of every step, and apply the stability of equality once at the end.

Prop. I.26 is another example of the use of the stability of equality: "... DE is not unequal to AB, and is therefore equal to it." While we have proved that classical arguments can be eliminated from proofs of Euclid's theorems, in fact it seems that the only classical arguments that *occur* in Euclid are applications of the principle "if *ab* and *cd* are unequal then one of them is longer."

In the examples above, this principle is not really needed to reach Euclid's desired conclusion. Since the conclusion concerns the equality of certain points, we can simply double-negate each step of the argument, and then add one application of the stability of equality at the end. The double negation of "if *ab* and *cd* are unequal then one is longer" is provable, since (intuitively speaking) $\neg(p < q)$ is $q \leq p$ and $p \leq q \wedge q \leq p$ implies $p = q$. In fact, this *had* to happen: we prove metatheorems below explaining why the principle in question, and indeed, any uses of classical logic whatsoever, are in principle eliminable from proofs of theorems of the form found in Euclid.

In order to arrive at such metatheorems, we first formulate the principle in question in the language of **ECG**, which does not contain $<$ as a primitive. Our formulation is as follows: If two unequal points B and C are both between A and D, then either B is between A and C or C is between A and B. Formally that is

$$B \neq C \land \mathbf{B}(A,B,D) \land \mathbf{B}(A,C,D) \supset \mathbf{B}(A,B,C) \lor \mathbf{B}(A,C,B) \qquad \text{(Axiom 75)}$$

The point D is a matter of convenience; the axiom is really about A, B, and C and their positions on a ray emanating from A, but **ECG** does not have rays as primitive, so we need point D to express this in **ECG**.

We defined $AB < CD$ in Definition 48. We write $AB > CD$ to mean $CD < AB$.

Lemma 10 *(in* **ECG***) Axiom 75 implies that if $AB \neq CD$, then either $AB > CD$ or $AB < CD$.*

Proof. Given AB and CD, let Q be the intersection point mentioned in the definition of $AB > CD$. Then according to Axiom 75, if $Q \neq B$ then either Q is between A and B or B is between Q and A. In the first case we have $AB > CD$. In the second case, B is inside the circle of radius CD about A. It follows that D is outside the circle of radius AB and center C. That completes the proof of the lemma.

The theory **ECG** has quantifier-free, disjunction-free axioms. It follows (as we will prove in Theorem 8) that no non-trivial disjunction can be proved in **ECG**. That is, if P is negative and **ECG** proves $\forall x\, P(x) \supset A(x) \lor B(x)$, then **ECG** proves $\forall x\, P(x) \supset A(x)$ or **ECG** proves $\forall x\, P(x) \supset B(x)$. Hence, Axiom 75 is not provable in **ECG**.

When we add Axiom 75, we will also introduce a new construction term, which we write $\mathit{if}(AB > CD, P, Q)$. This abbreviates the official version, which is $\mathit{if}(A,B,C,D,P,Q)$. Provided $AB \neq CD$, this term constructs a point which is equal either to P or to Q, depending on whether $AB > CD$ or $AB < CD$. The axiom expressing this is

$$\begin{aligned}&(AB > CD \supset \mathit{if}(AB > CD, P, Q) = P) &\text{(Axiom 76)}\\ &\land(AB < CD \supset \mathit{if}(AB > CD, P, Q) = Q)\end{aligned}$$

Note that Axiom 76 does not contain disjunction, but that by Axiom 75, we have

$$AB \neq CD \supset \mathit{if}(AB > CD, P, Q)\downarrow .$$

Definition 1 *The theory* **ECGD** *is* **ECG** *plus the new function symbol "if", with Axioms 75 and 76.*

Remark. The "D" in **ECGD** is for "disjunction".

The following lemmas give two appealing theorems of **ECGD** that cannot be proved in **ECG** (because they are non-trivial disjunctions).

Lemma 11 *(in* **ECGD***) Let P be point not on line L. Then any point Q is either on the same side of L as P, or on the opposite side.*

Proof. Drop a perpendicular K from P to L, meeting L at point R. Projecting Q to point Q' on K. Extend segment PR past R by the amount RQ' twice, arriving at point D on K. Then both R and Q' are between P and D, so by Axiom 75, either R is between P and Q' or Q' is between P and R. In the first case, Q' and P are on the same side of L, and in the second case, they are on opposite sides. But Q and Q' are on the same side of L. Hence Q and P are on the same side, or on opposite sides, of L. That completes the proof of the lemma.

Lemma 12 *(in* **ECGD***) In any triangle ABC, either ABC is a left turn or ABC is a right turn.*

Proof. In Lemma 8, we have already shown how to construct a point Q not on line $L = \mathit{Line}\,(\alpha, \beta)$, such that ABC is a left turn or a right turn according as Q lies on the same side of L as γ, or on the opposite side. In **ECGD** by Lemma 11, Q must lie on one side or the other of L. Hence ABC is either a left turn or a right turn. That completes the proof.

Terms of **ECGD** that involve the new symbol *if* represent geometrical constructions that can proceed by cases, with comparisons between constructed (unequal) lengths determining the next construction steps. Given Euclid's cavalier approach to case splits, the fact that such constructions are not explicitly mentioned in Euclid does not necessarily mean that they are not required to give a correct and complete version of Euclid. The question thus arises, whether Axiom 75 (or more generally, disjunctive axioms of any kind) are required to formalize Euclid. But because the *theorems* of Euclid do not mention disjunction in any essential way, we can simply take the double-negation interpretation, and eliminate Axioms 75 and 76, as will be shown below. Thus what happened in the example of Proposition I.6 happens necessarily in all examples of similar logical form.

10 Euclid's parallel postulate proved in ECG

Let P be a point not on line L. We consider lines through P that do not meet L (i.e., are parallel to L). Playfair's version of the parallel postulate says that two parallels to L through P are equal. Our Axiom 58 says that if K is a parallel to L through P and M is another line through P, with $M \neq K$, then M meets L. Recall that 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. Suppose P is a point not on line L, and K is parallel to P through L, M is another line through P, and Q any point on L. Let A be any point on M not equal to P and not on L. Euclid's condition that M "make the interior angles less than two right angles" on the side of PQ where A lies can be conveniently expressed by saying that A is between Q and the intersection of QA with K. Thus a formal version of Euclid's parallel axiom is

$$\begin{aligned}&\mathit{Parallel}(L,K) \wedge \mathit{on}(P,K) \wedge \mathit{on}(P,M) \wedge \mathit{on}(A,M) \wedge \qquad \text{(Euclid's Postulate 5)}\\ &\neg \mathit{on}(A,L) \wedge A \neq P \wedge \mathbf{B}(Q, A, \mathit{IntersectLines}(\mathit{Line}\,(Q,A),K))\\ &\quad \supset \mathbf{B}(P, A, \mathit{IntersectLines}(L,M))\end{aligned}$$

Note that the logical axioms of LPT make it superfluous to state in the conclusion that $\mathit{IntersectLines}(L, M)$ is defined. That follows automatically.

Axiom 58 has a weaker conclusion than Euclid's Postulate 5, because it does not specify on which side of P the intersection point will lie. On the other hand, Axiom 58 also has a weaker hypothesis than Euclid's Postulate 5, so its exact relationship to Postulate 5 is not immediately clear. One direction is settled by the following theorem:

Theorem 3 **ECG** *proves Euclid's Postulate 5.*

Proof. Suppose Axiom 58, and let L be a line, P a point not on L, K parallel to L through P, M another line through P, Q be a point on L, A be a point on M not on PQ, and B the intersection of QA with K. Suppose that the interior angles made by L, M, and PQ make less than two right angles, which formally means that A is between Q and and B. Then by Axiom 58, since $M \neq K$, M does meet L at some point R. It remains to show that A is between P and R. By Markov's principle (Axiom 34) it suffices to prove that P is not between R and A, and R is not between P and A. Suppose first that P is between R and A. Then R is on the opposite side of line K from A. But R is not not on the same side of line K as Q, since if R is not on the same side of K as Q, then RQ (which is L) would meet K (which it does not). By

Lemma 7, R and Q are on the same side of line K. Since R is on the opposite side of K from A, it follows that A and Q are on opposite sides of line K. Hence point B, the intersection of AQ with K, must be between A and C. But that contradicts the fact that A is between Q and B. Hence the assumption that P is between R and A has led to a contradiction. Now suppose instead that R is between P and A. Then A and P are on opposite sides of L.

But B and P are on the same side of line L, since $BP = K$ does not meet L. Hence A and B are on opposite sides of L. But then the intersection point of AB and L, which is Q, lies between A and B, contradicting the fact that A lies between Q and B. Hence the assumption that R is between P and A has also led to a contradiction. Hence, as noted already, by Axiom 34, A is between P and R. That completes the proof of Euclid's Postulate 5 from Axiom 58.

Now we consider the converse problem, of deriving Axiom 58 from Euclid's Postulate 5. The obvious proof attempt works only if we assume Axiom 75.

Lemma 13 *Let T be the theory* **ECG** *without any parallel postulate. Then in T, Axiom 75 ("of two unequal segments one is longer") plus Euclid's Postulate 5 implies Axiom 58.*

Proof. Assume Euclid's postulate and let p be a point not on line L, and K parallel to L through p, and line M another line through p as in Axiom 58. Let q be the point on L at the foot of the perpendicular to L from p. Let A on line M and B on line L be on the same side of pq. Then K makes the interior angles Apq and pqB equal to two right angles. Hence M does not, or it would coincide with K. If angle Apq is less than a right angle, then by Euclid's Postulate 5, M meets L. Similarly, if angle Apq is more than a right angle, we can show by Euclid's Postulate that M meets L (on the other side of q from B). By Axiom 75, one of these alternatives must hold. More formally, let S be the intersection point of *Line* (A, B) with K. Then angle Apq is less than a right angle if A is between S and B, and more than a right angle if S is between A and B. By Axiom 75, one of these alternatives holds. Hence by Euclid's Proposition 5, M meets L. That completes the proof of the lemma.

11 Constructive Geometry and Euclidean Fields

We know (from Theorem 2) that **ECG** plus classical logic is not only "reasonable" but equivalent to textbook theories. But when we use only intuitionistic logic, is it still a "reasonable" theory? There are two possible ways to answer that question:

- Can Euclid be formalized in **ECG**?
- Are the models of **ECG** characterizable in some elegant way?

We have argued in the previous section that probably **ECG** suffices to formalize Euclid, and certainly **ECGD** does suffice. In this section we take up the second approach. Classically, the models of G are all planes over a Euclidean field, that is, an ordered field in which every positive element has a square root. Is that same thing true constructively for **ECG**? The main point of this section is to answer that question in the affirmative. The detailed answer is somewhat surprising, though. There turns out to be more than one natural set of axioms for constructive Euclidean fields, and these different versions correspond directly to **ECG**, to **ECGD**, and to **ECG**with our parallel postulate (Axiom 58) replaced by the (apparently) weaker version, Euclid's Postulate 5.

Before turning to the proofs of those correspondences, we must prove some elementary lemmas about intuitionistic geometry. We begin with a two-dimensional version of Markov's principle.

Lemma 14 **ECG** *proves that, if point P does not lie on line L, then some circle with nonzero radius and center P lies entirely on the same side of L as P.*

Proof. Let point P not lie on line L; we will construct a circle with center P lying on the same side of L as P. Let point K on L be the foot of the perpendicular from P to L. Then $K \neq P$. Hence the two circles $C_1 = \mathit{Circle}\,(K, P)$ and $C_2 = \mathit{Circle}\,(P, K)$ that are used to bisect PK have different centers and each contains a point inside the other. Hence the points $X = \mathit{IntersectCircles1}(C_1, C_2)$ and $Y = \mathit{IntersectCircles2}(C_1, C_2)$ are defined. If $X = Y$ then X is between P and K, contradicting $PX = PK$. Hence the midpoint of PK is given by $M = \mathit{IntersectLines}(\mathit{Line}\,(X, Y), \mathit{Line}\,(P, K))$. Hence $\mathit{Circle}\,(P, M)$ lies on the same side of L as P, because K is nearer to P than any other point of L. That completes the proof of the lemma.

Next one may wonder whether the axioms of betweenness of **ECG** are actually sufficient to establish a reasonable theory of order of points on a line. Instead of basing our definition of order on betweenness, we base it instead on *Right* and *Left* . Fix a line L and a point 0 on L. Let K be the perpendicular to L at K and let I be a point on K that is not on L. Define $x < y$ for points x and y on L to mean that xyI is a left turn, and $x < y$ to mean that xyI is a right turn. Then we can establish the fundamental facts about intuitionistic order in **ECG**.

Lemma 15 *With notation as above,* **ECGD** *proves* $x \neq 0 \supset x > 0 \vee x < 0$.

Proof. Since x, 0, and I are distinct and not collinear, xyI is either a left turn or a right turn, by Lemma 12. That completes the proof.

Lemma 16 *(in* **ECG***) With notation as above,* **ECG** *proves* $0 \not< x \wedge x \not< 0 \supset x = 0$.

Proof. Suppose $0 \not< x$ and $x \not< 0$. Then $0xI$ is not a left turn and $x0I$ is not a left turn. Suppose $x \neq 0$. Then $x0I$ is a triangle and by Lemma 8, not not (one of $x0I$ and $0xI$ is a right turn and the other is a left turn). But neither is a left turn; hence $\neg\neg x = 0$. Since in **ECG** we have the axiom $\neg\neg x = y \supset x = y$, we can conclude $x = 0$. That completes the proof.

But note that **ECG** cannot prove $a < b \supset a < x \vee x < b$, since the decision as to which alternative holds cannot be made continuously in x. This theorem is essentially the axiom of apartness.

For points on L, we say that x is on the same side of y as z if y does not lie on $\mathit{Segment}\,(x, z)$. Note that this is equivalent to saying that x and y are on the same side of the line perpendicular to L at y. Then we have

Lemma 17 *(in* **ECG***) With notation as above, fix a point* $1 > 0$ *on* L*. Then* $x > 0$ *if and only if* x *is on the same side of 0 as 1.*

Proof. Since $0 < 1$, $01I$ is a left turn. Assume x is on the same side of 0 as 1. Then, by the axioms for *Left* , $0xI$ is a left turn. Hence $0 < x$. Conversely, assume $0 < x$. Then $0xI$ is a left turn. We must show x is on the same side of 0 as 1; that is, we must show that 0 does not lie on $\mathit{Segment}\,(x, 1)$. Suppose 0 does lie on $\mathit{Segment}\,(x, 1)$. Since $0xI$ is a left turn, $0 \neq x$, since $00I$ is not a left turn. Since $0 \neq 1$, we have $\mathbf{B}(x, 0, 1)$. But $01I$ is a left turn; hence both $0xI$ and $01I$ are left turns, contradicting Lemma 8. That completes the proof.

With the basic properties of order established, we are ready to turn to the characterization of models of **ECG** and **ECGD** as planes over Euclidean fields. We first discuss the axiomatization of Euclidean fields with intuitionistic logic. We use a language with symbols + for addition and $\cdot$ for multiplication, and a unary predicate $P(x)$ for "x is positive". We take the usual axioms for fields, except the axiom for multiplicative inverse, which says that positive elements have multiplicative inverses. If positive elements have inverses, it is an easy exercise to show that negative elements do too. We define a *Euclidean field* to be a commutative ring satisfying the following additional axioms:

$$x \neq 0 \supset \exists y\,(x \cdot y = 1) \qquad \text{EF1}$$
$$P(x) \wedge P(y) \supset P(x + y) \wedge P(x \cdot y) \qquad \text{EF2}$$
$$x + y = 0 \supset \neg(P(x) \wedge P(y)) \qquad \text{EF3}$$

$$x + y = 0 \wedge \neg P(x) \wedge \neg P(y) \supset x = 0 \qquad \text{EF4}$$
$$x + y = 0 \wedge \neg P(y) \supset \exists\, z(z \cdot z = x) \qquad \text{EF5}$$
$$\neg\neg P(x) \supset P(x) \qquad \text{EF6, or Markov's principle}$$

Axiom EF5 says that non-negative elements have square roots. This is a stronger axiom, intuitionistically, than simply specifying that positive elements have square roots. We could conservatively extend field theory by a unary function symbol for negation, $-x$, in which case the last axiom could be more readably written $\neg P(-x) \supset \exists z\, (z^2 = x)$.

As usual, we define $x < y$ to mean $\exists z(P(z) \wedge x + z = y$, or informally, $y - x$ is positive; and $x \leq y$ means $\neg(y < x$. Then Markov's principle is equivalent to $\neg(x \leq 0) \supset 0 < x$.

We also consider *weakly Euclidean fields*, which instead of EF1, are required only to satisfy the weaker axiom

$$P(x) \supset \exists y\, (x \cdot y = 1) \qquad \text{EF0}$$

EF1 implies EF0, because $P(x) \supset x \neq 0$, but to derive EF1 from EF0, we would have to know that each nonzero element is either positive or negative. In the language without a symbol for additive inverse, this can expressed as

$$x + y = 0 \wedge x \neq 0 \supset P(x) \vee P(y) \qquad \text{EF7}$$

Fields that satisfy EF0 and EF2 through EF7 then automatically satisfy EF1 as well. We call these fields *strongly Euclidean.* To recap: In a Euclidean field, all nonzero elements have multiplicative inverses, while in a weakly Euclidean field, it is only required that elements known to be positive or negative have multiplicative inverses. In a strongly Euclidean field, nonzero elements are either positive or negative, so the distinction doesn't matter.

Remark. Since the double negation of EF7 follows from EF4, every Euclidean field is not not strongly Euclidean; in other words, one cannot give an example of a Euclidean field that is not strongly Euclidean. The standard plane is strongly Euclidean if and only if Markov's principle (in the formulation $\neg\neg x > 0 \supset x > 0$) holds in the reals.

Theorem 4 *The models of* **ECG** *are all isomorphic to planes F^2, where F is a Euclidean field. and the relations and function symbols of* **ECG** *are interpreted as usual in reducing geometry to field theory. The models of* **ECGD** *are all isomorphic to planes F^2, where F is a strongly Euclidean field. The models of* **ECG** *with the parallel Axiom 58 replaced by Euclid's Postulate 5 are all isomorphic to planes F^2, where F is a weakly Euclidean field. In fact, each of these geometrical theories interprets, and is interpretable in, the corresponding field theory.*

Remark. Theory A is interpretable in theory B if there is a map from the syntax of A (variables, terms, formulas) to that of B that preserves provability. Thus the theorem implies, for example, that Euclid's Postulate 5 implies Axiom 58 if and only if axiom EF0 implies EF1 with the help of EF2 through EF6.

Proof. We show how to interpret field theory in our geometrical theories. We have three fixed points α, β, and γ, pairwise unequal. Let 0 be another name for α and let 1 be another name for β. Let L be *Line* (α, β), and let K be the perpendicular to L at 0. Let I be a point on K but not on L such that $01I$ is a left turn, and define $P(x)$ to mean that $0xI$ is a left turn. The function symbols $+$ and $\cdot$ of field theory are interpreted by terms *Add* and *Multiply* of **ECG**, defined as shown earlier.

Technically, we ought to exhibit formal proofs in **ECG** of the (interpretations of) the ring axioms, corresponding to the informal proofs in the section on arithmetization, and it may well be possible to exhibit such proofs using a theorem-prover or proof-checker, but here we rely on the reader to be convinced that such proofs exist based on an examination of the proofs and the axioms of **ECG**.

Now we will check axiom EF3. Suppose $x + y = 0$ holds in this model. That means that $Add(x, y) = 0$. Suppose also that $\neg P(x)$ and $\neg P(y)$, i.e. neither $0xI$ nor $0yI$ is a left turn. Then x is not between 0 and 1, since $01I$ is a left turn, and if x is between 0 and 1, $0xI$ would be a left turn too. Similarly y is not between 0 and 1. Now assume $x < 0$. then $Add(x, y) = 0$ implies that x and y are on opposite sides of 0. Since $x0I$ is a left turn, by Lemma 8 $0xI$ is a right turn. Hence $0yI$ is a left turn, so $0 < y$, contradiction. Hence $\neg x < 0$. Since we have proved x is neither negative nor positive, we have $x = 0$ by Lemma 16. That completes the verification of axiom EF3.

Now we turn to axiom EF5. Suppose $x + y = 0$. Suppose also $\neg P(y)$; that is, it is not that case that $y \neq 0$ and $0yI$ is a left turn. Verifying this axiom amounts to checking that in **ECG**, Descartes' square root construction can be extended to a function defined for $x \geq 0$, without requiring a case distinction as to whether or not the argument is zero. We have given just such an extension of Descartes' construction in the section on arithmetization, and now it only remains to remark that the argument given there can be carried out in **ECG**.

Next we check axiom EF7 in **ECGD**. Let $x + y = 0$ and $x \neq 0$. Then x and y are on line L, and we can find a point D on the same side of 0 as 1 and a point A on the opposite side of 0 from 1, x between A and D. Then either 0 is between A and x, in which case x is positive (that is $0xI$ is a left turn), or x is between A and 0, in which case y is on the opposite side of 0 from x, i.e. the same side as 1, and hence y is positive.

Next we verify Markov's principle $\neg\neg P(x) \supset P(x)$. Suppose $\neg\neg P(x)$, i.e. $\neg\neg x > 0$. Let $y = Add(x, 1)$, so that $y > x$. Then $x > 0$ is equivalent to $\mathbf{B}(0, x, y)$. Hence we have $\neg\neg B(0, x, y)$. Then by Markov's principle in **ECG**, we have $B(0, x, y)$, i.e. $x > 0$, i.e. $P(x)$, as desired.

We now turn to the verifications of the parallel Axiom 58 and Euclid's Postulate 5. In order to define the reciprocal $1/x$ in geometry, we use Descartes's method. That is, we fix a line L to serve as the x-axis, and a point 0 on L and a point $1 \neq 0$ on L. Then let X be a point on L, and suppose $X \neq 0$. We wish to define a point $1/X$. Erect the perpendicular K to L at X, and find a point on K at distance 1 from X, for example $Q = \mathit{IntersectLineCircle1}(K, \mathit{Circle}\ (X, 0, 1))$. Then $Q \neq 0$, since Q is not on L, so we can form $M = \mathit{Line}\ (0, Q)$. Erect the perpendicular H to L at 1. Then K and H are parallel, since they are both perpendicular to L. Line M does not coincide with K, since 0 lies on M but not on K (since $X \neq 0$). Hence M is a line through Q that is not parallel to H. Then, by Axiom 58, M meets H in a point R. The segment $R1$ has the desired length. The desired point $1/X$ on line L is one of the intersection points of $\mathit{Circle}\ (0, 1, R)$ with L. Which one it is depends on the sign of X, which we do not know; but the selection is made automatically by the definition of *IntersectLineCircle2*:

$$1/X = \mathit{IntersectLineCircle2}(\mathit{Line}\ (0, X), \mathit{Circle}\ (0, 1, R))$$

since the two intersection points are numbered in the same order as 0 and X occur on L. Hence arithmetic on L satisfies the axioms of a Euclidean field. Similarly, if we only have Euclid's Postulate 5, we can still construct $1/X$ if we know that $X > 0$, as follows. Consider the interior angles made by M and H with the perpendicular dropped from Q to H. They will make less than two right angles if $X > 0$, and more than two right angles if $X < 0$. Hence by Euclid's Postulate 5, if $X > 0$, M meets H as shown in the figure, while if $X < 0$, M meets H at a point south of L. One can then verify that arithmetic on L satisfies the weak Euclidean field axioms. Finally, Axiom 75 erases the distinction in question, and enables us to verify axiom EF7 as well.

Conversely, assume F is a Euclidean field. We will show how to turn F^2 into a model of **ECG** (or, to describe the construction more formally, we will show how to interpret geometry in field theory). As usual in the corresponding classical theories, we take the points to be elements of F^2, and let lines, circles, arcs, and segments have their usual analytic definitions; in particular we define circles so that circles of zero radius are allowed. Hence *Circle3* can be interpreted. Markov's principle in F allows us to verify that axiom of **ECG**. The intersection points of circles and lines, and the intersection points of circles and circles, can be defined by the solution of quadratic equations; and which one is which (i.e. the concept "ABC is a left turn") can be defined as usual in computer graphics, by the cross product, which can be defined in Euclidean field theory (note that division is not required). Then one has to verify that the axioms of **ECG** about handedness are valid. The details follow the sketch given in an earlier section, and are omitted. The verification of the parallel Axiom 58 requires that the reciprocal $1/x$ be defined when $x \neq 0$; the verification of Euclid's Postulate 5 only requires that the reciprocal be defined when $x > 0$. If F is strongly Euclidean (satisfies EF7) then we can also verify Axiom 75, and in that case there is on obvious interpretation of the function symbol *if* of **ECGD** validating Axiom 76. That completes the proof of the theorem.

In [3] it is shown that the different constructive version of Euclidean field theory are not only apparently different, but really are not equivalent with constructive logic; and as a corollary, the different versions of the parallel postulate are also not equivalent.

12 Classical logic not needed for negative theorems

Our plan in this section is to investigate the double-negation interpretation for geometric theories. Since the double-negation interpretation applies a double-negation to atomic formulas, we need to have $\neg\neg A \supset A$ for each atomic formula A. We first consider the case when A has the form $t \downarrow$.

The following schema seems initially to have the character of Markov's principle, since $t \downarrow$, in number theory or even in the recursive model $\mathbf{R}^2$, involves an existential quantifier.

$$\neg\neg t \downarrow \supset t \downarrow \qquad \text{for all terms } t$$

As it turns out, however, this schema is provable in **ECG**; however, we require Axiom 58 for the proof, and Euclid's Postulate 5 apparently does not suffice.

Lemma 18 *Let t be any term of* **ECG**. *Then* **ECG** *proves $\neg\neg t \downarrow \supset t \downarrow$. Moreover, Axiom 34 (Markov's principle for betweenness) is not needed in the proof.*

Proof. We proceed by induction on the complexity of the term t. If t is a variable or constant then $t \downarrow$ is an axiom of **LPT**, so we need consider only compound terms. Consider the induction step when t is a term $f(q, r)$, for terms q and r. By the induction hypothesis, **ECG** proves $\neg\neg q \downarrow \supset q \downarrow$. Argue in **ECG** as follows: Suppose $\neg\neg f(q, r) \downarrow$ We claim $\neg\neg q \downarrow$. For suppose $\neg q \downarrow$. Then by the strictness axioms of **LPT**, we have $\neg f(q, r) \downarrow$, contradicting $\neg\neg f(q, r) \downarrow$. That contradiction completes the proof that $\neg\neg q \downarrow$. Using the proof in **ECG** that $\neg\neg q \downarrow \supset q \downarrow$, we conclude $q \downarrow$. Similarly, we have $r \downarrow$. Hence, to complete the proof, it suffices to prove for each function symbol f of **ECG** that

$$s_1 \downarrow \wedge \ldots \wedge s_n \downarrow \wedge \neg\neg f(s_1, \ldots, s_n) \downarrow \supset f(s_1, \ldots, s_n) \downarrow$$

We will prove this by an exhaustive consideration of each function symbol of **ECG**. First consider the case when f is *IntersectLines*. Suppose that *IntersectLines*(K, L) is not undefined. By Axiom 12, if $K = L$ then *IntersectLines*(K, L) is undefined; hence $K \neq L$. Then by Axiom 58 (the parallel axiom), K meets L. Note that neither Playfair's version of the parallel axiom nor Euclid's would seem to suffice here: Playfair's does not guarantee the existence of the intersection point, and Euclid's guarantees it only when we know on which side of some transversal K makes interior angles less than two right angles.

Consider the case when f is *IntersectLineCircle1* or *IntersectLineCircle2*. Then we have a line L and a circle C that do not fail to meet, and we must show that they do indeed meet. Let P be the center of C, and let K be the line through P perpendicular to L (which we have shown how to construct without knowing whether P is on L or not). Let F be the foot of this perpendicular, i.e. the intersection point of L and K. Then L meets C if and only if PF is less than or equal to the radius r of C. So we have $\neg\neg PF \leq r$. Then (even without Markov's principle) we have $PF \leq r$, so C does meet L as desired.

Next consider the case when f is *IntersectCircles1* or *IntersectCircles2*. Then we have two circles C and K that do not fail to intersect, and we must show that they do intersect. The relevant geometrical fact is that two circles intersect if and only if the distance d between their centers is less than or equal to the sum of their radii $r_1 + r_2$. So if C and K do not fail to intersect implies $\neg\neg d \leq r_1 + r_2$. Even without Markov's principle we then have $d \leq r_1 + r_2$, so C and K do intersect.

Next we consider the constructors. *Circle* (P, Q) is always defined, since we allow zero-radius circles; similarly for *Circle3*(P, Q, R). *Line* (P, Q) is defined if and only if $P \neq Q$. Hence we need $\neg\neg P \neq Q \supset P \neq Q$. This follows from the general intuitionistic logical principal that triple negation is equivalent to single negation. The constructors for segments and arcs can be treated similarly.

Finally we consider the accessors, such as *center*, *pointOn1*, etc. These are all total, so there is nothing to prove. That completes the proof of the lemma.

Let A^- be the Gödel double-negation interpretation of A, obtained by replacing $\exists$ by $\neg\forall\neg$ and $A \vee B$ by $\neg(\neg A \wedge \neg B)$. We do not replace A by $\neg\neg A$ for atomic A since these are equivalent in intuitionistic **ECG**.

Theorem 5 (Double negation interpretation) *Suppose* **ECG** *with classical logic proves* A. *Then* **ECG** *with intuitionistic logic proves* A^-.

Proof. First we observe that $\neg\neg A$ is equivalent to A for atomic A. This is an axiom for all atomic formulas not of the form $t\downarrow$, and also for those of that form when t has the form *IntersectLines*(u, v). For other formulas of the form $t\downarrow$, we have proved it in Lemma 18. Since the axioms of **ECG** are quantifier-free and disjunction-free, it follows that so A^- is equivalent to A for axioms A of the **ECG**. Now the theorem follows as soon as we check the soundness of the double-negation interpretation in a multi-sorted logic with partial terms. But that is straightforward; sorts and partial terms offer no complications over the usual first-order case. That completes the proof.

Corollary 3 **ECG** *with classical logic is conservative over* **ECG** *with intuitionistic logic for negative formulae.*

Proof. For negative A, A^- is identical to A.

A typical theorem of Euclid has the form $H \supset A$, where A will be quantifier-free when formulated in **ECG**, and H is a collection of hypotheses that certain points are distinct, or certain incidence relations hold or do not hold. As Proclus pointed out, sometimes this implies a theorem formulated by cases. For example, Euclid I.2 has one proof if $A = C$ and another proof if $A \neq C$. Using the double-negation interpretation, we find a proof that $A = C \vee A \neq C$ implies the conclusion of Euclid 1.2, but without the law of the excluded middle we cannot conclude the "uniform version" of Euclid I.2.

Recall the example given above of Euclid's Prop. I.6, where Axiom 75 is used, but the same conclusion can be reached without it. Since Axiom 75 is classically tautological, the double-negation interpretation shows that it is *always* eliminable from proofs of negative theorems. But all the theorems in Euclid are either already negative, or assert the existence of some objects that can be constructed using the terms of **ECG**; when formulated more explicitly, they are negative in the sense that they say that the result of a certain construction has certain (quantifier-free) properties.

13 From proofs to geometric algorithms

In this section we take up our plan of doing for **ECG** what cut-elimination and recursive realizability did for intuitionistic arithmetic and analysis, namely, to show that existence proofs lead to programs (or terms) producing the object whose existence is proved. In the case of **ECG** we want to produce geometrical constructions, not just recursive constructions (which could already be produced by known techniques, since **ECG** is interpretable in Heyting's arithmetic of finite types). Terms of **ECG** correspond in a natural way to straightedge and compass constructions.

Theorem 6 (Geometric constructions extracted from intuitionistic proofs) *(i)Suppose* **ECG** *proves* $P(x) \supset \exists y\, \phi(x,y)$ *where* P *is negative (does not contain* $\exists$ *or* $\vee$*). Then there is a term* $t(x)$ *of* **ECG** *(representing a geometric construction) such that* $P(x) \supset (\phi(x,t(x))$ *is also provable in* **ECG**.

(ii) Same as (i) but with **ECGD** *in place of* **ECG**.

(iii) Let **ECG** + **DE** *be* **ECG**, *augmented with a constant* **D** *and the axiom saying* **D** *is a test-for-equality function. Then the analogue of (i) holds for* **ECG** + **DE**.

Proof. We use cut-elimination.[9] Since our axiomatization is quantifier-free, if $\psi \rightarrow \exists y\, \phi$ is provable in constructive **ECG**, then there is a list Γ of quantifier-free axioms such that $\Gamma, \psi \Rightarrow \exists y\, \phi$ is provable by a cut-free (hence quantifier-free) proof. Since our axiomatization is disjunction-free, by [12] (or rather, by its adaptation to multi-sorted logic with the logic of partial terms) we can permute the inferences so that the existential quantifier is introduced at the last step. Then we obtain the desired proof just by omitting the last step of the proof. That completes the proof of part (i). All the work was in arranging the axiom system to be quantifier-free and disjunction-free. Part (iii) is proved in the same way, noting that the axioms for **D** are also disjunction-free.

Part (ii) requires a bit more work. Since Axiom 75 contains a disjunction, there is an issue about permuting the inferences in a cut-free proof. Suppose that we have a cut-proof of $\Gamma \Rightarrow \psi \rightarrow \exists y\, \phi$, where Γ is a list of axioms (or subformulas of axioms) of **ECGD**. Among Γ there may be occurrences of Axiom 75, which contains a disjunction. We prove by induction on the number of disjunctions in Γ that there exists a term t of **ECGD** and a list Δ of axioms of **ECGD** such that $\Gamma, \Delta \Rightarrow \phi(t)$ is provable. The basis case, when there are no disjunctions, is part (i) of the theorem. Now for the induction step. If the $\exists$ on the right is introduced at a lower level (nearer the end-sequent) than the lowest introduction of disjunction on the left, then we can complete the proof as above, since the line just before the $\exists$ is introduced will contain the desired term, and the $\exists$-introduction can just be postponed until the end. Otherwise there is a part of the proof that looks like this:

$$\frac{AB > AC, \Gamma_1 \Rightarrow \exists y\, \phi \qquad AB < AC, \Gamma_2 \Rightarrow \exists y\, \phi}{AB > AC \vee AB < AC, \Gamma_1, \Gamma_2 \Rightarrow \exists y\, \phi}$$

By induction hypothesis, there are terms t_1 and t_2 such that $AB > AC, \Gamma_1 \Rightarrow \phi(t_1)$ is provable and $AB < AC, \Gamma_2 \Rightarrow \phi(t_2)$ is provable. Let t be the term $\mathit{if}\,(AB > AC, t_1, t_2)$. Then **ECGD** proves $AB > AC \supset t = t_1$ (by Axiom 76), and **ECGD** proves $AB < AC \supset t = t_2$. Hence, for some list Δ of axioms of **ECGD**, there is a cut-free proof of $AB > AC, \Delta, \Gamma_1 \Rightarrow \phi(t)$, and a cut-free proof of $AB < AC, \Delta, \Gamma_2 \Rightarrow \phi(t)$. These two proofs can then be combined as follows:

$$\frac{AB > AC, \Delta, \Gamma_1 \Rightarrow \phi(t) \qquad AB < AC, \Delta, \Gamma_2 \Rightarrow \phi(t)}{AB > AC \vee AB < AC, \Delta, \Gamma_1, \Gamma_2 \Rightarrow \phi(t)}$$

That completes the induction step, and with it, the proof of the lemma.

Example 1. The "other intersection point". Many Euclidean constructions involve constructing one intersection point P of a line $L = \mathit{Line}\,(A,B)$ and a circle C, and then we say "Let

[9] In fact, we use cut-elimination for many-sorted logic with the logic of partial terms. The details of the cut-elimination theorem for such logics have not been published, but they are not significantly different from Gentzen's formulation for first-order logic.

Q be the other intersection point of L and C". Of course we can prove "if P lies on L and C and the two intersection points of L and C are not equal, then there exists an x such that $x \neq P \wedge on(x, L) \wedge On(x, C)$." Then by the theorem, there must be a term $t(A, B, C, P)$ such that, under the stated conditions, $t(A, B, C, P)$ is an intersection point of L and C that is not equal to P. It is not immediately obvious what this term t might be, and it would be interesting to extract it by computer from a proof. But we should be able to see directly how such a term could be constructed.

Here is a sketch of such a construction. First, project the center of C onto line L, obtaining point R on L. Since the two intersection points are distinct, $R \neq P$. Then we ask whether (A, B) is in the same order on L as (R, P) or not. Now Definition 71 shows how to construct a certain point E (given by a term involving A and B) such that (A, B) is in the same order as (R, P) if and only if ERP is a left turn. And the proof of Lemma 8 shows how to construct a (complicated) term $\ell(A, B, P, C)$ that will be equal to γ if and only if ERP is a left turn. If ERP is a right turn, then $\ell(A, B, P, C)$ can be arranged to be another specific point α' on the other side of $\alpha\beta$ from γ, the same distance from line $\alpha\beta$ along line $\beta\gamma$ as α, but on the other side. Combining this term ℓ with terms representing an appropriate dilation, translation, and rotations, we can construct a term $f(A, B, P, C)$ such that if (A, B) has the same order as (R, P) then $(f(A, B, P, C) = A$, and if (A, B) has the opposite order as (R, P), it is another point A' with $\mathbf{B}(A, B, A'$. Hence $(f(A, B, P, C), B)$ has the same or opposite order as (A, B), depending on whether (A, B) has the same or opposite order as (R, P). Note that if (A, B) has the same order as (R, P), then $P = \mathit{IntersectLineCircle2}(L, C)$, while if not, $P = \mathit{IntersectLineCircle1}(L, C)$. Hence we can take

$$t(A, B, C, P) = \mathit{IntersectLineCircle1}(\mathit{Line}\ (f(A, B, P, C), B), C)$$

as the "other intersection point" constructor.

In case one thinks, "this is not what Euclid had in mind!", that is of course true; but Euclid never tried to give uniform constructions of this type. Perhaps the complexity of this construction is one reason why not. This example shows that one could conservatively add a function symbol to **ECG** for "the other intersection point".

Theorem 7 (Geometric constructions extracted from classical proofs) *Suppose* **ECG** *with classical logic proves* $P(x) \supset \exists y\, \phi(x, y)$ *where P is quantifier-free and disjunction-free. Then there are terms* $t_1(x), \ldots, t_n(x)$ *of* **ECG** *such that* $P(x) \supset \phi(x, t_1(x)) \vee \ldots \vee \phi(x, t_n(x))$ *is also provable in* **ECG** *with classical logic.*

Proof. This is a special case of Herbrand's theorem.

Example 2. Euclid's proof of Book I, Proposition 2 provides us with two such constructions, $t_1(A, B, C) = C$ and $t_2(A, B, C)$ the result of Euclid's construction of a point D with $AD = BC$, valid if $A \neq B$. Classically we have $\forall A, B, C\, \exists D(AD = BC)$, but we need two terms t_1 and t_2 to cover all cases.

Example 3. Let P and Q be distinct points and L a given line, and A, B, and C points on L, with A and B on the same side of C. Then there exists a point D which is equal to P if B is between A and C and equal to Q if A is between B and C. The two terms t_1 and t_2 for this example can be taken to be the variables P and Q. One term will not suffice, since D cannot depend continuously on A and B, but all constructed points do depend continuously on their parameters. This classical theorem is therefore not constructively provable.

We mentioned above that **ECG** cannot prove any non-trivial disjunctive theorem. That is a simple consequence of the fact that its axioms contain no disjunction. We now spell this out:

Theorem 8 (ECG cannot prove a nontrivial disjunctive theorem) *Suppose* **ECG** *proves* $H(x) \supset P(x) \vee Q(x)$*, where H is negative. Then either* **ECG** *proves* $H(x) \supset P(x)$ *or* **ECG** *proves* $H(x) \supset Q(x)$*. (This result depends only on the lack of disjunction in the axioms of* **ECG***.)*

Proof. Consider a cut-free proof of $\Gamma, H(x) \supset P(x) \vee Q(x)$, where Γ is a list of some axioms of **ECG**. Tracing the disjunction upwards in the proof, if we reach a place where the disjunction was introduced on the right before reaching a leaf of the proof tree, then we can erase the other disjunct below that introduction, obtaining a proof of one disjunct as required. If we reach a leaf of the proof tree with $P(x) \vee Q(x)$ still present on the right, then it occurs on the left, where it appears positively. Its descendants will also be positive, so it cannot participate in in application of the rule for proof by cases (which introduces $\vee$ in the left side of a sequent); and it cannot reach left side of the bottom sequent, namely $\Gamma, H(x)$, as these formulas contain no disjunction. But a glance at the rules of cut-free proof, e.g. on p. 442 of [11], will show that these are the only possibilities. That completes the proof.

Theorem 9 (Disjunction Properties for ECG and ECGD) *Suppose* **ECGD** *proves*

$$H(x) \supset P(x) \vee Q(x)$$

where H is negative. Then there is a term $t(x)$ of **ECGD** *such that* **ECG** *proves*

$$H(x) \supset t(x) = \alpha \vee t(x) = \beta$$

and **ECGD** *proves*

$$t(x) = \alpha \supset P(x) \wedge t(x) = \beta \supset Q(x).$$

(Here α and β are two constants of **ECGD***, with $\alpha \neq \beta$ an axiom.)*

Proof. Suppose **ECGD** proves $H(x) \supset P(x) \vee Q(x)$. Then also **ECGD** proves

$$H(x) \supset \exists y\, ((y = \alpha \supset P(x)) \wedge (y = \beta \supset Q(x))).$$

The formula on the right is disjunction-free, so by Theorem 6, there is a term t as required.

14 Conclusions

We have given a quantifier-free, disjunction-free axiomatization **ECG** of Euclidean Constructive Geometry, making use of the Logic of Partial Terms (**LPT**). We have verified that this theory is a reasonable version of intuitionistic geometry, by checking that its models are planes over Euclidean fields. Past versions of intuitionistic geometry have included either apartness or decidable equality. Both of these destroy the property of continuity that the terms of **ECG** possess. The terms of this theory correspond in a natural way to Euclidean straightedge-and compass constructions. Making use of more-or-less standard proof-theoretical tools, we have shown that proofs of existential theorems contain Euclidean constructions of the objects proved to exist, and that these constructions can be automatically extracted from such proofs. As a corollary, objects proved to exist in **ECG** depend continuously on parameters.

We set out to pursue the analogy

$$\frac{\text{formal number theory}}{\text{Turing computable functions}} = \frac{\text{intuitionistic geometry}}{\text{geometric constructions}}$$

and we think that we have found the correct theory **ECG** to place on the right side of this equation.

We have shown that Euclid is essentially constructive, in the process exposing these interesting facts:

- in constructive geometry, we need a rigid compass (for the uniform version of Prop. I.2).
- We need Markov's principle to prove the fundamental properties of same-side and opposite-side of a line. We therefore conclude that Markov's principle is fundamental to geometry; theories without it do not correspond to our geometrical intuition.

- Different versions of the parallel postulate correspond to different axiom systems for Euclidean fields: Euclid's postulate 5 amounts to assuming $1/x$ is defined when $x > 0 \vee x < 0$, while Axiom 58 amounts to assuming that $x \neq 0$ implies $1/x$ is defined. It may be simpler to determine the logical relations between these field theories than to work directly in geometrical theories.
- Axiom 58 implies Euclid's Postulate 5, and conversely with the aid of Axiom 75 (corresponding to $x \neq 0 \supset 0 < x \vee x < 0$). That the reverse implications are not constructively valid is shown in [3].
- Since **ECG** and **ECGD** correspond well to Euclid and to constructive field theories, we conclude that apartness is not necessary for constructive geometry.

15 Appendix: List of axioms of ECG

In this section we list the axioms of **ECG** given above without comment, for reference. The underlying logic is three-sorted intuitionistic logic (sorts for points, lines, and circles) with the logic of partial terms **LPT** [2], p. 97.

In the following list, the symbol := is used for a definition (macro). The symbols on the left side are to be replaced (with argument substitution) in their subsequent uses by the right hand side of the definition. α, β, and γ are the only constant symbols in **ECG**.

The underlying logic of LPT provides the following:

$$
\begin{array}{ll}
\forall x(t\downarrow \wedge A(x) \supset A(t)) & \\
A(t) \wedge t\downarrow \supset \exists x\, A(x) & \\
f(t_1,\ldots,t_n)\downarrow \supset t_1\downarrow \wedge \ldots \wedge t_n\downarrow & \\
R(t_1,\ldots,t_n) \supset t_1\downarrow \wedge \ldots \wedge t_n\downarrow & \\
t \cong s := (t\downarrow \supset s\downarrow \wedge t = s) \wedge (s\downarrow \supset t\downarrow \wedge t = s) & \\
x\downarrow & \text{for every variable } x \\
c\downarrow & \text{for every constant } c
\end{array}
$$

In listing the axioms of **ECG**, it is not necessary to explicitly indicate the types (or sorts) of the variables, as this can be mechanically deduced from the signature of the relation and function symbols. (That is one reason for distinguishing *on* and *On*.) The signatures of the function symbols have also not been explicitly given here, as the names chosen for them convey that information. The following are the axioms of **ECG**, including the definitions used in stating the axioms.

$$\neg\neg x = y \supset x = y \tag{1}$$
$$\neg\neg\delta(A,B,C,D) \supset \delta(A,B,C,D) \tag{2}$$
$$\neg\neg on(P,L) \supset on(P,L) \tag{3}$$
$$\neg\neg On(P,C) \supset On(P,C) \tag{4}$$
$$P = IntersectLines(L,K) \supset on(P,L) \wedge on(P,K) \tag{5}$$
$$IntersectLines(L,K) \cong IntersectLines(K,L) \tag{6}$$
$$P = IntersectLineCircle1(L,C) \supset on(P,L) \wedge On(P,C) \tag{7}$$
$$P = IntersectLineCircle2(L,C) \supset on(P,L) \wedge On(P,C) \tag{8}$$
$$P = IntersectCircles1(C,K) \supset On(P,C) \wedge On(P,K) \tag{9}$$
$$P = IntersectCircles2(C,K) \supset On(P,C) \wedge On(P,K) \tag{10}$$
$$on(P,L) \wedge \neg on(P,K) \supset IntersectLines(L,K)\downarrow \tag{11}$$

$$\mathit{IntersectLines}(L,K)\downarrow \wedge on(P,L) \wedge on(P,K) \supset P = \mathit{IntersectLines}(L,K) \tag{12}$$
$$on(P,L) \wedge On(P,C) \supset \mathit{IntersectLineCircle1}(L,C)\downarrow \tag{13}$$
$$on(P,L) \wedge On(P,C) \supset \mathit{IntersectLineCircle2}(L,C)\downarrow \tag{14}$$
$$On(P,C) \wedge On(P,K) \supset \mathit{IntersectCircles1}(C,K)\downarrow \tag{15}$$
$$On(P,C) \wedge On(P,K) \supset \mathit{IntersectCircles2}(C,K)\downarrow \tag{16}$$
$$\mathit{Line}\ (A,B)\downarrow \leftrightarrow A \neq B \tag{17}$$
$$\mathit{Circle}\ (A,B)\downarrow \tag{18}$$
$$\mathit{Line}\ (\mathit{pointOn1}(L), \mathit{pointOn2}(L)) = L \tag{19}$$
$$\mathit{pointOn1}(\mathit{Line}\ (A,B)) = A \tag{20}$$
$$\mathit{pointOn2}(\mathit{Line}\ (A,B)) = B \tag{21}$$
$$\mathit{pointOn1}(L) \neq \mathit{pointOn2}(L) \tag{22}$$
$$\mathit{Circle}\ (\mathit{center}(C), \mathit{pointOnCircle}(C)) = C \tag{23}$$
$$\mathit{center}(\mathit{Circle}\ (A,B)) = A \tag{24}$$
$$\mathit{pointOnCircle}(\mathit{Circle}\ (A,B)) = B \tag{25}$$
$$\mathit{center}(C) \neq \mathit{pointOnCircle}(C) \tag{26}$$
$$\neg on(\alpha, \mathit{Line}\ (\beta,\gamma) \tag{27}$$
$$\neg on(\beta, \mathit{Line}\ (\alpha,\gamma) \tag{28}$$
$$\neg on(\gamma, \mathit{Line}\ (\alpha,\beta) \tag{29}$$
$$on(A, \mathit{Line}\ (A,B)) \tag{30}$$
$$on(B, \mathit{Line}\ (A,B)) \tag{31}$$
$$on(P,L) \wedge on(Q,L) \wedge on(R, \mathit{Line}\ (P,Q)) \supset on(R,L) \tag{32}$$
$$\mathbf{B}(a,b,c) \supset \mathbf{B}(c,b,a) \tag{33}$$
$$\neg\neg\mathbf{B}(a,b,c) \supset \mathbf{B}(a,b,c) \tag{34}$$
$$\begin{aligned} & a \neq b \wedge a \neq c \wedge b \neq c \supset \\ & \quad (\neg\mathbf{B}(a,b,c) \wedge \neg\mathbf{B}(b,c,a) \supset \neg\neg\mathbf{B}(c,a,b)) \\ & \quad (\neg\mathbf{B}(b,c,a) \wedge \neg\mathbf{B}(c,a,b) \supset \neg\neg\mathbf{B}(a,b,c)) \\ & \quad (\neg\mathbf{B}(c,a,b) \wedge \neg\mathbf{B}(a,b,c) \supset \neg\neg\mathbf{B}(b,c,a)) \\ & \quad \neg(\mathbf{B}(a,b,c) \wedge \mathbf{B}(b,c,a)) \wedge \neg(\mathbf{B}(a,b,c) \wedge \mathbf{B}(b,a,c)) \\ & \quad \neg(\mathbf{B}(b,c,a) \wedge \mathbf{B}(b,a,c)) \end{aligned} \tag{35}$$
$$\begin{aligned} & \mathbf{B}(P, \mathit{IntersectLines}(\mathit{Line}(P,Q), \mathit{Line}(\\ & \qquad \mathit{IntersectCircles1}(\mathit{Circle}(P,Q), \mathit{Circle}(Q,P)), \\ & \qquad \mathit{IntersectCircles2}(\mathit{Circle}(P,Q), \mathit{Circle}(Q,P))), Q) \end{aligned} \tag{36}$$
$$\mathbf{B}(\mathit{IntersectLineCircle1}(\mathit{Line}(P,Q), \mathit{Circle}(P,Q)), P, Q) \tag{37}$$
$$\mathbf{B}(P, Q, \mathit{IntersectLineCircle2}(\mathit{Line}(P,Q), \mathit{Circle}(Q,P))) \tag{38}$$
$$\mathit{OppositeSide}(P,Q,L) := \mathbf{B}(P, Q, \mathit{IntersectLines}(\mathit{Line}\ (P,Q), L) \tag{39}$$
$$\mathit{SameSide}(P,Q,L) := \neg\mathbf{B}(P, Q, \mathit{IntersectLines}(\mathit{Line}\ (P,Q), L)) \tag{40}$$
$$\mathit{SameSide}(A,B,L) \wedge \mathit{SameSide}(B,C,L) \supset \mathit{SameSide}(A,C,L) \tag{41}$$
$$\mathit{OppositeSide}(A,B,L) \wedge \mathit{OppositeSide}(B,C,L) \supset \mathit{SameSide}(A,C,L) \tag{42}$$
$$on(P, \mathit{Ray}\ (O,B)) := on(P, \mathit{Line}\ (O,B)) \wedge \neg\mathbf{B}(P,O,B) \tag{43}$$
$$On(P, \mathit{Circle}\ (A,Q)) \leftrightarrow \delta(A,P,A,Q) \wedge A \neq Q \tag{44}$$
$$A \neq B \wedge C \neq D \supset \tag{45}$$

$(on(\mathit{IntersectLineCircle1}(\mathit{Line}\ (A,B), \mathit{Circle3}(A,C,D)), \mathit{Ray}\ (A,B)) \wedge$
$\delta(A, \mathit{IntersectLineCircle1}(\mathit{Line}\ (A,B), \mathit{Circle3}(A,C,D)), C, D)$

$$\delta(A,B,C,D) \wedge \delta(A,B,E,F) \supset \delta(C,D,E,F) \tag{46}$$

$$\mathbf{B}(A,B,C) \wedge \mathbf{B}(P,Q,R) \wedge \delta(A,B,P,Q) \wedge \delta(B,C,Q,R) \supset \delta(A,C,P,R) \tag{47}$$

$$AB < CD := \mathbf{B}(C, \mathit{IntersectLineCircle2}(\mathit{Line}\ (C,D), \mathit{Circle3}(C,A,B)), D) \tag{48}$$

$$CD \leq AB := \neg\mathbf{B}(C, \mathit{IntersectLineCircle2}(\mathit{Line}\ (C,D), \mathit{Circle3}(C,A,B)), D) \tag{49}$$

$$AB = CD := \delta(A,B,C,D) \tag{50}$$

$$AB = PQ \wedge BC = QR \wedge AC = PR \wedge PQ = UV \wedge QR = VW \wedge PR = UV \supset \\ AB = UV \wedge BC = VW \wedge AC = UW \tag{51}$$

$$A \neq B \wedge A \neq C \wedge B \neq C \wedge A' \neq B' \wedge \\ B'' = \mathit{IntersectLineCircle1}(\mathit{Line}\ (A',B'), C) \wedge \\ K_1 = \mathit{Circle3}(B'',B,C) \wedge K_2 = \mathit{Circle3}(A'',A,C) \supset \\ \mathit{IntersectCircles1}(K_1,K_2) \downarrow \wedge \mathit{IntersectCircles2}(K_1,K_2) \downarrow \wedge \\ \mathit{OppositeSide}(\mathit{IntersectCircles1}(K_1,K_2), \\ \mathit{IntersectCircles2}(K_1,K_2), \mathit{Line}\ (A',B')) \tag{52}$$

$$AP \leq AB \wedge on(P,L) \supset \mathit{IntersectLineCircle1}(L, \mathit{Circle}\ (A,B)) \downarrow \tag{53}$$

$$AP \leq AB \wedge on(P,L) \supset \mathit{IntersectLineCircle2}(L, \mathit{Circle}\ (A,B)) \downarrow \tag{54}$$

$$A = center(C) \wedge A = center(K) \wedge On(P,C) \wedge On(Q,K) \wedge AP = AQ \supset \\ \mathit{IntersectLineCircle1}(L,C) \cong \mathit{IntersectLineCircle1}(K,C) \wedge \\ \mathit{IntersectLineCircle2}(L,C) \cong \mathit{IntersectLineCircle2}(K,C) \tag{55}$$

$$On(P,C) \wedge AP \leq AB \supset \\ \mathit{IntersectCircles1}(C, \mathit{Circle}\ (A,B)) \downarrow \wedge\ \mathit{IntersectCircles2}(C, \mathit{Circle}\ (A,B)) \downarrow \tag{56}$$

$$A = center(C_1) \wedge A = center(C_1) \wedge On(P,C_1) \wedge On(Q,C_1) \wedge AP = AQ \supset \\ \mathit{IntersectCircles1}(C_1,K) \cong \mathit{IntersectCircles1}(C_2,K) \wedge \\ \mathit{IntersectCircles2}(C_1,K) \cong \mathit{IntersectCircles2}(C_2,K) \wedge \\ \mathit{IntersectCircles1}(K,C_1) \cong \mathit{IntersectCircles1}(K,C_2) \wedge \\ \mathit{IntersectCircles2}(K,C_1) \cong \mathit{IntersectCircles2}(K,C_2) \wedge \tag{57}$$

$$\neg\mathit{IntersectLines}(K,L) \downarrow \wedge on(p,K) \wedge on(p,M) \wedge M \neq K \supset \mathit{IntersectLines}(L,M) \tag{58}$$

$$\mathit{Left}\ (A,B,C) := C = \mathit{IntersectCircles1}(\mathit{Circle}\ (A,C), \mathit{Circle}\ (B,C)) \\ \wedge \neg on(C, \mathit{Line}\ (A,B)) \tag{59}$$

$$\mathit{Right}\ (A,B,C) := C = \mathit{IntersectCircles2}(\mathit{Circle}\ (A,C), \mathit{Circle}\ (B,C)) \\ \wedge \neg on(C, \mathit{Line}\ (A,B)) \tag{60}$$

$$\mathit{Left}\ (\alpha,\beta,\gamma) \tag{61}$$

$$\mathit{Right}\ (\alpha,\gamma,\beta) \tag{62}$$

$$\mathit{Left}\ (P,Q,R) \wedge P \neq P' \wedge R \neq R' \wedge on(P', \mathit{Ray}\ (Q,P)) \wedge on(R', \mathit{Ray}\ (Q,R)) \\ \supset \mathit{Left}\ (P',Q',R') \tag{63}$$

$$\mathit{Left}\ (P,Q,R) \wedge \neg\mathbf{B}(P, \mathit{IntersectLines}(\mathit{Line}\ (Q,R), \mathit{Line}\ (P,P')), P') \\ \supset \mathit{Left}\ (P',Q,R) \tag{64}$$

$$\mathit{Left}\ (P,Q,R) \wedge \neg\mathbf{B}(R, \mathit{IntersectLines}(\mathit{Line}\ (Q,P), \mathit{Line}\ (R,R')), P') \\ \supset \mathit{Left}\ (P,Q,R') \tag{65}$$

$$\mathit{Left}\ (A,B,C) \wedge AB = PQ \wedge BC = QR \wedge \tag{66}$$

$$AC = PR \wedge AP = BQ \wedge AP = CR \supset \mathit{Left}\ (P, Q, R)$$

$$P \neq P' \wedge R \neq R' \wedge on(P', \mathit{Ray}\ (Q, P)) \wedge on(R', \mathit{Ray}\ (Q, R)) \wedge \mathit{Right}\ (P, Q, R) \supset \mathit{Right}\ (P', Q', R') \tag{67}$$

$$\mathit{Right}\ (P, Q, R) \wedge \neg \mathbf{B}(P, \mathit{IntersectLines}(\mathit{Line}\ (Q, R), \mathit{Line}\ (P, P')), P') \supset \mathit{Right}\ (P', Q, R) \tag{68}$$

$$\mathit{Right}\ (P, Q, R) \wedge \neg \mathbf{B}(R, \mathit{IntersectLines}(\mathit{Line}\ (Q, P), \mathit{Line}\ (R, R')), P') \supset \mathit{Right}\ (P, Q, R') \tag{69}$$

$$\mathit{Right}\ (A, B, C) \wedge AB = PQ \wedge BC = QR \wedge AC = PR \wedge AP = BQ \wedge AP = CR \supset \mathit{Right}\ (P, Q, R) \tag{70}$$

$$\mathit{SameOrder}\ (A, B, P, Q) := A \neq B \wedge P \neq Q \wedge on(P, \mathit{Line}\ (A, B)) \wedge on(Q, \mathit{Line}\ (A, B)) \wedge \mathit{Left}\ (P, Q, \mathit{IntersectCircles1}(\mathit{Circle}\ (A, B), \mathit{Circle}\ (B, A)) \tag{71}$$

$$P = \mathit{IntersectLineCircle1}(\mathit{Line}\ (A, B), C) \wedge Q = \mathit{IntersectLineCircle2}(\mathit{Line}\ (A, B), C) \wedge P \neq Q \supset \mathit{SameOrder}\ (A, B, P, Q) \tag{72}$$

$$R = \mathit{IntersectCircles1}(\mathit{Circle}\ (A, P), \mathit{Circle}\ (B, Q)) \supset \neg \mathit{Right}\ (A, B, R) \tag{73}$$

$$R = \mathit{IntersectCircles2}(\mathit{Circle}\ (A, P), \mathit{Circle}\ (B, Q)) \supset \neg \mathit{Left}\ (A, B, R) \tag{74}$$

The following two axioms are not part of **ECG**, but are used in **ECGD**:

$$B \neq C \wedge \mathbf{B}(A, B, D) \wedge \mathbf{B}(A, C, D) \supset \mathbf{B}(A, B, C) \vee \mathbf{B}(A, C, B) \tag{75}$$

$$(AB > CD \supset if(AB > CD, P, Q) = P) \wedge (AB < CD \supset if(AB > CD, P, Q) = Q) \tag{76}$$

References

[1] Avigad, J., Dean, Edward, and Mumma, John, A formal system for Euclid's *Elements*, *Review of Symbolic Logic* (to appear in 2009).

[2] Beeson, M., *Foundations of Constructive Mathematics*, Springer-Verlag, Berlin/ Heidelberg/ New York (1985).

[3] Beeson, M., Constructive geometry and the parallel postulate, to appear. Preprint available on the author's website.

[4] Bishop, E., *Foundations of Constructive Analysis*, McGraw-Hill, New York (1967).

[5] Borsuk, K. and Szmielew, W., *Foundations of Geometry*, Amsterdam, North Holland (1960).

[6] Descartes, R., *La Geometrie*. Facsimile with English translation published as *The Geometry of Rene Descartes*, Dover, New York (1954).

[7] Euclid, *The Thirteen Books of the Elements*, second edition of the Heath translation, Dover, New York (1956).

[8] Greenberg, M. J., *Euclidean and Non-Euclidean Geometries*, fourth edition, Freeman, New York (2008).

[9] Heyting, A., *Intuitionism, An Introduction*, North-Holland, Amsterdam (1956).

[10] Hilbert, D., *Foundations of Geometry*, translated from the tenth German edition by Leo Unger, Open Court, La Salle, Illinois (1971).

[11] Kleene, S. C., *Introduction to Metamathematics*, van Nostrand, Princeton (1952).

[12] Kleene, S. C., Permutability of inferences in Gentzen's calculi LK and LJ, in: *Two Papers on the Predicate Calculus*, A.M.S. Memoirs **10** (1952), A. M. S., Providence, R. I.

[13] Kijne, D., *Plane Construction Field Theory*, Ph. D. thesis, Rijksuniversiteit te Utrecht (1956).

[14] von Plato, Jan, The axioms of constructive geometry, *Annals of Pure and Applied Logic* **76**, pp. 169–200 (1995) .

[15] Pambuccian, V., Axiomatizing geometric constructions, *J. Applied Logic* **6**, pp. 24–46, 2008.

[16] Pambuccian, V., Axiomatizations of hyperbolic and absolute geometries, in *Non-Euclidean geometries, János Bolyai memorial volume. Papers from the International Conference on Hyperbolic Geometry held in Budapest, July 6–12, 2002*, pp. 119–153. Edited by András Prékopa and Emil Molnár. Mathematics and Its Applications (New York), **581**. Springer, New York (2006).