Old and New Results in the Foundations of Elementary Plane Euclidean and Non-Euclidean Geometries

Marvin Jay Greenberg

By "elementary" plane geometry I mean the geometry of lines and circles—straight-edge and compass constructions—in both Euclidean and non-Euclidean planes. An axiomatic description of it is in Sections 1.1, 1.2, and 1.6. This survey highlights some foundational history and some interesting recent discoveries that deserve to be better known, such as the hierarchies of axiom systems, Aristotle's axiom as a "missing link," Bolyai's discovery—proved and generalized by William Jagy—of the relationship of "circle-squaring" in a hyperbolic plane to Fermat primes, the undecidability, incompleteness, and consistency of elementary Euclidean geometry, and much more. A main theme is what Hilbert called "the purity of methods of proof," exemplified in his and his early twentieth century successors' works on foundations of geometry.

1. AXIOMATIC DEVELOPMENT

1.0. Viewpoint. Euclid's *Elements* was the first axiomatic presentation of mathematics, based on his five postulates plus his "common notions." It wasn't until the end of the nineteenth century that rigorous revisions of Euclid's axiomatics were presented, filling in the many gaps in his definitions and proofs. The revision with the greatest influence was that by David Hilbert starting in 1899, which will be discussed below. Hilbert not only made Euclid's geometry rigorous, he investigated the minimal assumptions needed to prove Euclid's results, he showed the independence of some of his own axioms from the others, he presented unusual models to show certain statements unprovable from others, and in subsequent editions he explored in his appendices many other interesting topics, including his foundation for plane hyperbolic geometry without bringing in real numbers. Thus his work was mainly *metamathematical*, not geometry for its own sake.

The disengagement of elementary geometry from the system of real numbers was an important accomplishment by Hilbert and the researchers who succeeded him [**20**, Appendix B]. The view here is that elementary Euclidean geometry is a much more ancient and simpler subject than the axiomatic theory of real numbers, that the discovery of the independence of the continuum hypothesis and the different versions of real numbers in the literature (e.g., Herman Weyl's predicative version, Errett Bishop's constructive version) make real numbers somewhat controversial, so we should not base foundations of elementary geometry on them. Also, it is unaesthetic in mathematics to use tools in proofs that are not really needed. In his eloquent historical essay [**24**], Robin Hartshorne explains how "the true essence of geometry can develop most naturally and economically" without real numbers.[1]

Plane Euclidean geometry without bringing in real numbers is in the spirit of the first four volumes of Euclid. Euclid's Book V, attributed to Eudoxus, establishes a

doi:10.4169/000298910X480063

[1]Hartshorne's essay [**24**] elaborates on our viewpoint and is particularly recommended to those who were taught that real numbers precede elementary geometry, as in the ruler and protractor postulates of [**32**].

theory of proportions that can handle any quantities, whether rational or irrational, that may occur in Euclid's geometry. Some authors assert that Eudoxus' treatment led to Dedekind's definition of real numbers via cuts (see Moise [**32**, §20.7], who claimed they should be called "Eudoxian cuts"). Eudoxus' theory is applied by Euclid in Book VI to develop his theory of similar triangles. However, Hilbert showed that the theory of similar triangles can actually be fully developed in the plane without introducing real numbers and without even introducing Archimedes' axiom [**28**, §14–16 and Supplement II]. His method was simplified by B. Levi and G. Vailati [**10**, Artikel 7, §19, p. 240], cleverly using an elementary result about cyclic quadrilaterals (quadrilaterals which have a circumscribed circle), thereby avoiding Hilbert's long excursion into the ramifications of the Pappus and Desargues theorems (of course that excursion is of interest in its own right). See Hartshorne [**23**, Proposition 5.8 and §20] for that elegant development.

Why did Hilbert bother to circumvent the use of real numbers? The answer can be gleaned from the concluding sentences of his *Grundlagen der Geometrie* (Foundations of Geometry, [**28**, p. 107]), where he emphasized the *purity of methods of proof*.[2] He wrote that "the present geometric investigation seeks to uncover which axioms, hypotheses or aids are necessary for the proof of a fact in elementary geometry ... " In this survey, we further pursue that investigation of what is necessary in elementary geometry.

We next review Hilbert-type axioms for elementary plane Euclidean geometry—because they are of great interest in themselves, but also because we want to exhibit a standard set of axioms for geometry that we can use as a reference point when investigating other axioms. Our succinct summaries of results are intended to whet readers' interest in exploring the references provided.

1.1. Hilbert-type Axioms for Elementary Plane Geometry Without Real Numbers. The first edition of David Hilbert's *Grundlagen der Geometrie*, published in 1899, is referred to as his *Festschrift* because it was written for a celebration in memory of C. F. Gauss and W. Weber. It had six more German editions during his lifetime and seven more after his death (the fourteenth being the centenary in 1999), with many changes, appendices, supplements, and footnotes added by him, Paul Bernays, and others (see the Unger translation [**28**] of the tenth German edition for the best rendition in English, and see [**22**] for the genesis of Hilbert's work in foundations of geometry). Hilbert provided axioms for three-dimensional Euclidean geometry, repairing the many gaps in Euclid, particularly the missing axioms for betweenness, which were first presented in 1882 by Moritz Pasch. Appendix III in later editions was Hilbert's 1903 axiomatization of plane hyperbolic (Bolyai-Lobachevskian) geometry. Hilbert's plane hyperbolic geometry will be discussed in Section 1.6.

Hilbert divided his axioms into five groups entitled Incidence, Betweenness (or Order), Congruence, Continuity, and a Parallelism axiom. In the current formulation, for the first three groups and only for the plane, there are three incidence axioms, four betweenness axioms, and six congruence axioms—thirteen in all (see [**20**, pp. 597–601] for the statements of all of them, slightly modified from Hilbert's original).

The primitive (undefined) terms are *point*, *line*, *incidence* (point lying on a line), *betweenness* (relation for three points), and *congruence*. From these, the other standard geometric terms are then defined (such as circle, segment, triangle, angle, right angle,

[2]For an extended discussion of purity of methods of proof in the *Grundlagen der Geometrie*, as well as elsewhere in mathematics, see [**21**] and [**7**]. For the history of the *Grundlagen* and its influence on subsequent mathematics up to 1987, see [**2**].

perpendicular lines, etc.). Most important is the definition of two lines being *parallel*: by definition, l is parallel to m if no point lies on both of them (i.e., they do not intersect).

We briefly describe the axioms in the first three groups:

The first incidence axiom states that two points lie on a unique line; this is Euclid's first postulate (Euclid said to draw the line). The other two incidence axioms assert that every line has at least two points lying on it and that there exist three points that do not all lie on one line (i.e., that are not *collinear*).

The first three betweenness axioms state obvious conditions we expect from this relation, writing $A * B * C$ to denote "B is between A and C": if $A * B * C$, then A, B, and C are distinct collinear points, and $C * B * A$. Conversely, if A, B, and C are distinct and collinear, then exactly one of them is between the other two. For any two points B and D on a line l, there exist three other points A, C, and E on l such that $A * B * D$, $B * C * D$, and $B * D * E$. The fourth betweenness axiom—the Plane Separation axiom—asserts that every line l bounds two disjoint half-planes (by definition, the *half-plane* containing a point A not on l consists of A and all other points B not on l such that segment AB does not intersect l). This axiom helps fill the gap in Euclid's proof of I.16, the Exterior Angle theorem [**20**, p. 165]. It is equivalent to Pasch's axiom that a line which intersects a side of a triangle between two of its vertices and which is not incident with the third vertex must intersect exactly one of the other two sides of the triangle.

There are two primitive relations of congruence—congruence of segments and congruence of angles. Two axioms assert that they are equivalence relations. Two axioms assert the possibility of laying off segments and angles uniquely. One axiom asserts the additivity of segment congruence; the additivity of angle congruence can be proved and need not be assumed as an axiom, once the next and last congruence axiom is assumed.

Congruence axiom six is the side-angle-side (SAS) criterion for congruence of triangles; it provides the connection between segment congruence and angle congruence. Euclid pretended to prove SAS by "superposition." Hilbert gave a model to show that SAS cannot be proved from the first twelve axioms [**28**, §11]; see [**20**, Ch. 3, Exercise 35 and Major Exercise 6] for other models. The other familiar triangle congruence criteria (ASA, AAS, and SSS) are provable. If there is a correspondence between the vertices of two triangles such that corresponding angles are congruent (AAA), those triangles are by definition *similar*. (The usual definition—that corresponding sides are proportional—becomes a theorem once proportionality has been defined and its theory developed.)

A model of those thirteen axioms is now called a *Hilbert plane* ([**23**, p. 97] or [**20**, p. 129]). For the purposes of this survey, we take *elementary plane geometry* to mean the study of Hilbert planes.

The axioms for a Hilbert plane eliminate the possibility that there are no parallels at all—they eliminate spherical and elliptic geometry. Namely, a parallel m to a line l through a point P not on l is proved to exist by "the standard construction" of dropping a perpendicular t from P to l and then erecting the perpendicular m to t through P [**20**, Corollary 2 to the Alternate Interior Angle theorem]. The proof that this constructs a parallel breaks down in an elliptic plane, because there a line does not bound two disjoint half-planes [**20**, Note, p. 166].

The axioms for a Hilbert plane can be considered one version of what J. Bolyai called *absolute plane geometry*—a geometry common to both Euclidean and hyperbolic plane geometries; we will modify this a bit in Section 1.6. (F. Bachmann's axioms based on reflections furnish an axiomatic presentation of geometry "absolute" enough to also include elliptic geometry and more—see Ewald [**12**] for a presentation in English.)

For the foundation of Euclidean plane geometry, Hilbert included the following axiom of parallels (John Playfair's axiom from 1795, usually misstated to include existence of the parallel and stated many centuries earlier by Proclus [**39**, p. 291]):

Hilbert's Euclidean Axiom of Parallels. *For every line l and every point P not on l, there does not exist more than one line through P parallel to l.*

It is easily proved that for Hilbert planes, this axiom, the fourteenth on our list, is equivalent to Euclid's fifth postulate [**20**, Theorem 4.4]. I propose to call models of those fourteen axioms *Pythagorean planes*, for the following reasons: it has been proved that those models are isomorphic to Cartesian planes F^2 coordinatized by arbitrary *ordered Pythagorean fields*—ordered fields F such that $\sqrt{a^2+b^2} \in F$ for all $a, b \in F$ [**23**, Theorem 21.1]. In particular, $\sqrt{2} \in F$, and by induction, $\sqrt{n} \in F$ for all positive integers n. The field F associated to a given model was constructed from the geometry by Hilbert: it is the *field of segment arithmetic* ([**23**, §19] or [**28**, §15]). Segment arithmetic was first discovered by Descartes, who used the theory of similar triangles to define it; Hilbert worked in the opposite direction, first defining segment arithmetic and then using it to develop the theory of similar triangles.

Another reason for the name "Pythagorean" is that the Pythagorean equation holds for all right triangles in a Pythagorean plane; the proof of this equation is the usual proof using similar triangles [**23**, Proposition 20.6], and the theory of similar triangles does hold in such planes [**23**, §20]. The smallest ordered Pythagorean field is a countable field called the *Hilbert field* (he first introduced it in [**28**, §9]); it coordinatizes a countable Pythagorean plane. The existence theorems in Pythagorean planes can be considered constructions with a straightedge and a transporter of segments, called *Hilbert's tools* by Hartshorne [**23**, p. 102, Exercise 20.21 and p. 515].

These models are not called "Euclidean planes" because one more axiom is needed in order to be able to prove all of Euclid's plane geometry propositions.

1.2. Continuity Axioms. Most of the plane geometry in Euclid's *Elements* can be carried out rigorously for Pythagorean planes, but there remain several results in Euclid which may fail in Pythagorean planes, such as Euclid I.22 (the Triangle Existence theorem in [**32**, §16.5]), which asserts that given three segments such that the sum of any two is greater than the third, a triangle can be constructed having its sides congruent to those segments—[**23**, Exercise 16.11] gives an example of I.22 failure. Hilbert recognized that I.22 could not be proved from his *Festschrift* axioms—see [**22**, p. 202].

Consider this fifteenth axiom, which was not one of Hilbert's:

Line-Circle Axiom. *If a line passes through a point inside a circle, then it intersects the circle (in two distinct points).*

This is an example of an *elementary* continuity axiom—it only refers to lines and/or circles. For a Pythagorean plane coordinatized by an ordered field F, this axiom holds if and only if F is a *Euclidean field*—an ordered field in which every positive element has a square root [**23**, Proposition 16.2]. Another elementary continuity axiom is:

Circle-Circle Axiom. *If one circle passes through a point inside and a point outside another circle, then the two circles intersect (in two distinct points).*

Euclid's proof of his very first proposition, I.1—the construction of an equilateral triangle on any base—tacitly uses circle-circle continuity in order to know that the circles he draws do intersect [**20**, p. 130]. Hartshorne and Pambuccian have independently shown [**23**, p. 493] that Euclid I.1 does not hold in all Hilbert planes. It does hold in all Pythagorean planes, as can be shown by first constructing the altitude standing on the midpoint of that base, using the fact that the field of segment arithmetic contains $\sqrt{3}$.

For arbitrary Hilbert planes, the Circle-Circle axiom is equivalent to the Triangle Existence theorem [**20**, Corollary, p. 173] and implies the Line-Circle axiom [**20**, Major Exercise 1, p. 200]. Conversely, Strommer [**43**] proved that Line-Circle implies Circle-Circle in all Hilbert planes, first proving the Three-Chord theorem (illustrated on the cover of [**23**] and explained in Section 1.2.1 below), then introducing a third circle and the radical axis it determines in order to invoke Line-Circle. Alternatively, Moise [**32**, §16.5] gave a proof of this implication for Pythagorean planes, and by using Pejas' classification of Hilbert planes ([**37**] or [**27**]), I was able to reduce the general case to that one.

Hartshorne [**23**, p. 112] calls a *Euclidean plane* any Pythagorean plane satisfying the Circle-Circle axiom. Euclidean planes are, up to isomorphism, just Cartesian planes F^2 coordinatized by arbitrary Euclidean fields F [**23**, Corollary 21.2]. The diligent reader can check that **every plane geometry proposition in Euclid's *Elements* can be proved from those fifteen axioms**. (For Euclid's propositions about "equal area," see [**23**, Chapter 5], where Hartshorne carefully specifies which results are valid in all Pythagorean planes.) Thus Euclid's plane geometry has been made completely rigorous without bringing in real numbers and Hartshorne's broader definition is justified.

Consider the Euclidean plane $\mathcal{E}$ coordinatized by the *constructible field* $\mathbf{K}$ (called the *surd field* by Moise in [**32**]); $\mathbf{K}$ is the closure (in $\mathbb{R}$, say) of the rational numbers under the field operations and the operation of taking square roots of positive numbers. (The Hilbert field consists of all the *totally real* numbers in $\mathbf{K}$—see [**23**, Exercises 16.10–16.14].) The countable model $\mathcal{E}$ is used to prove the impossibility of the three classical construction problems (trisecting every angle, squaring every circle, and duplicating a cube) using straightedge and compass alone ([**23**, §28] or [**32**, Ch. 19]). This application shows the importance of studying Euclidean planes other than $\mathbb{R}^2$. In the language of mathematical logic, it also shows that the theory of Euclidean planes is *incomplete*, meaning that there are statements in the theory that can be neither proved nor disproved. An example of such a statement is "every angle has a trisector"; this is true in $\mathbb{R}^2$ but false in $\mathcal{E}$. Moise said that the plane $\mathcal{E}$ is "all full of holes" [**32**, p. 294].

Notes. Hartshorne studied constructions with *marked ruler* and compass, such as trisecting any angle and constructing a regular heptagon; Viète formulated a new axiom to justify using a marked ruler [**20**, p. 33]. It is an open problem to determine the models of the theory with Viète's axiom added. Hartshorne in [**23**, §30–31] solved this problem in the special case where the mark is only used between two lines (not between a line and a circle). In an Archimedean Euclidean plane, the models are the Cartesian planes coordinatized by those Euclidean subfields of $\mathbb{R}$ in which one can find real roots of cubic and quartic equations—another lovely application of algebra to geometry.

In [**36**] Victor Pambuccian surveys the many works in which geometric constructions became part of the axiomatizations of various geometries (starting only in 1968). Michael Beeson [**1**] has written about geometric constructions using intuitionist logic.

1.2.1. Continuity Axioms in Hilbert's Work and Elsewhere. The Line-Circle and Circle-Circle axioms do not appear in Hilbert's *Grundlagen*. The treatise [**22**] edited by M. Hallett and U. Majer presents Hilbert's notes in German for geometry courses he taught from 1891 to 1902, as well as his 1899 *Festschrift*. They provide extensive discussion and explication in English of those materials, in which Hilbert covers many topics not included in his *Grundlagen*. One can see his ideas about the foundations evolving over time and see him and others solving most problems that arose along the way (such as problems related to the theorems of Desargues and Pappus).

In Hilbert's lectures of 1898–1899 on Euclidean geometry, he discussed the Three Chord theorem: if three circles whose centers are not collinear intersect each other pairwise, then the three chord-lines determined by those pairs of intersection points are concurrent. Hilbert gave a proof of this theorem which he noticed depends on the Triangle Existence theorem mentioned above. He also noticed the related line-circle and circle-circle properties and said that assuming those properties amounts to assuming that a circle is "a closed figure," which he did not define. Hilbert gave an example of a Pythagorean plane in which those three properties do not hold—he constructed a Pythagorean ordered field F which is not a Euclidean field and used the Cartesian plane coordinatized by F. Hallett discusses this in detail on pp. 200–206 of [**22**] and wonders why Hilbert did not add the circle-circle property as an axiom.

In his ingenious article [**25**], Robin Hartshorne proved the Three Chord theorem for any Hilbert plane in which the line-circle property holds. His proof uses the classification of Hilbert planes by W. Pejas ([**37**] or [**27**]).

So what continuity axioms did Hilbert assume? In his *Festschrift*, he only assumed Archimedes' axiom, which will be discussed in the next section; it is unclear what that axiom has to do with continuity, except that it allows measurement of segments and angles by real numbers. The models of his planar *Festschrift* axioms are all the Cartesian planes coordinatized by Pythagorean subfields of the field $\mathbb{R}$ of real numbers. In the second edition of his *Grundlagen*, Hilbert added a "completeness" axiom, as his second continuity axiom, stating that it is impossible to enlarge the sets of points and lines, and extend the relations of incidence, betweenness, and congruence to these larger sets, in such a way that the Pythagorean axioms and Archimedes' axiom are still satisfied. This is obviously not a geometric statement and not a statement formalizable in the language used previously, so what does it accomplish? The addition of those two "continuity" axioms to the fourteen axioms for a Pythagorean plane allowed him to prove that all models of those sixteen axioms are isomorphic to the Cartesian plane coordinatized by the entire field $\mathbb{R}$ [**28**, p. 31]. It is essential to notice that Hilbert did not use his completeness axiom for any ordinary geometric results in his development. Hallett gives a very thorough explication of the purpose of that axiom on pp. 426–435 of [**22**].[3]

Alternatively, Szmielew and Borsuk in [**3**] assumed only one continuity axiom:

Dedekind's Axiom. *Suppose the set of points on a line l is the disjoint union of two nonempty subsets such that no point of either subset is between two points of the other subset (such a pair of subsets is called a Dedekind cut of the line). Then there exists a unique point O on l such that one of the subsets is a ray of l emanating from vertex O and the other subset is its complement on l.*

They too prove that all models of their planar axioms are isomorphic to the Cartesian plane coordinatized by $\mathbb{R}$ (so their theory is also *categorical*). Unlike Hilbert's

[3]Hilbert used the term "complete" here in a different sense than the usual meaning of a theory being deductively complete. His "completeness" is a maximality condition.

completeness axiom, Dedekind's axiom implies Archimedes' axiom [**20**, pp. 135–136]. Hilbert refused to assume Dedekind's axiom because he studied in depth the role Archimedes' axiom plays by itself.[4] Dedekind's axiom also implies the Line-Circle axiom [**20**, p. 136] and the Circle-Circle axiom [**11**, vol. 1, p. 238]. It is in fact the mother of all continuity properties in geometry.

1.3. Archimedes' Axiom. Hilbert called Archimedes' axiom the "the axiom of measure" [**28**, p. 26], because it allows measurement of segments and angles by real numbers [**20**, pp. 169–172]. It is not strictly speaking a geometric axiom, for it includes a natural number variable n—not just geometric variables—in its statement.[5]

Archimedes' Axiom. *If CD is any segment, A any point, and r any ray with vertex A, then for every point $B \neq A$ on r, there exists a natural number n such that when CD is laid off n times on r starting at A, a point E is reached such that $n \cdot CD \cong AE$ and either $B = E$ or B is between A and E.*

If CD is taken as the unit segment, the length of AB is $\leq n$; or, if AB is taken as the unit segment, the length of CD is $\geq 1/n$. Thus, Archimedes' axiom states that no segment is infinitely large or infinitesimal with respect to any other segment as unit. Assuming this axiom, the analogous result can be proved for measurement of angles [**23**, Lemma 35.1].

Although Euclid did not list it as one of his postulates, Archimedes' axiom does occur surreptitiously in Euclid's *Elements* in his Definition V.4, where it is used to develop the theory of proportions. It appears in an equivalent form for arbitrary "magnitudes" in Proposition X.1. Archimedes assumed it, crediting it to Eudoxus, in his treatise *On the Sphere and Cylinder*.

Note. Non-Archimedean geometries were first considered by Giuseppe Veronese in his 1891 treatise *Fondamenti di Geometria*, a work that Hilbert called "profound" [**28**, p. 41, footnote]; for Hilbert's example of a non-Archimedean geometry, see [**28**, §12]. In 1907, H. Hahn discovered non-Archimedean completeness using his "Hahn field" instead of $\mathbb{R}$. Other eminent mathematicians (R. Baer, W. Krull, F. Bachmann, A. Prestel, M. Ziegler) developed the subject. See [**8**] for the categoricity theorem generalizing Hilbert's, plus other very interesting results. Ehrlich [**9**] is a fascinating history of non-Archimedean mathematics (I was astonished to learn that Cantor, whose infinite cardinal and ordinal numbers initially generated so much controversy, rejected infinitesimals). In [**6**], Branko Dragovich applied non-Archimedean geometry to theoretical physics.

1.4. Aristotle's Axiom. Closely related to Archimedes' axiom is the following axiom due to Aristotle:

Aristotle's Angle Unboundedness Axiom. *Given any acute angle, any side of that angle, and any challenge segment AB, there exists a point Y on the given side of the angle such that if X is the foot of the perpendicular from Y to the other side of the angle, then $YX > AB$.*

---

[4] E.g., in [**28**, §32] Hilbert proved that if an ordered division ring is Archimedean, then it is commutative, and hence a field; that commutativity in turn implies Pappus' theorem in the Desarguesian geometry coordinatized by such a ring.

[5] It becomes a geometric statement if one enlarges the logic to *infinitary logic*, where the statement becomes the infinite disjunction $cd \geq ab$ or $2cd \geq ab$ or $3cd \geq ab$ or $\ldots$.

In other words, *the perpendicular segments from one side of an acute angle to the other are unbounded*—no segment AB can be a bound.

Aristotle made essentially this statement in Book I of his treatise *De Caelo* ("On the heavens"). I will refer to it simply as "Aristotle's axiom." The mathematical importance of his axiom was highlighted by Proclus in the fifth century C.E., when Proclus used it in his attempted proof of Euclid's fifth postulate ([**39**, p. 291] and [**20**, p. 210]), which we will refer to from now on as "Euclid V." It is easy to prove that Euclid V implies Aristotle's axiom [**20**, Corollary 2, p. 180].

As examples of geometries where Aristotle's axiom does not hold, consider spherical geometry, where "lines" are interpreted as great circles, and plane elliptic geometry, obtained by identifying antipodal points on a sphere. On a sphere, as you move away from the angle vertex, the rays of an acute angle grow farther apart until they reach a maximum width, and then they converge and meet at the antipodal point of the vertex.

In Section 1.6 it will be shown that Aristotle's axiom is a consequence of Archimedes' axiom for Hilbert planes, but not conversely—it is a weaker axiom. Moreover, it is a *purely geometric axiom*, not referring to natural numbers. It is important as a "missing link" when the Euclidean parallel postulate is replaced with the statement that the angle sum of every triangle is 180°. It is also important as a "missing link" in the foundations of hyperbolic geometry, as will be discussed as well in Section 1.6.

1.5. Angle Sums of Triangles. The second part of Euclid's Proposition I.32 states that the angle sum of any triangle equals two right angles (we will say "is 180°" though no measurement is implied—there is no measurement in Euclid). The standard proof (not the one in Euclid's *Elements*) in most texts [**20**, proof of Proposition 4.11] was called by Proclus "the Pythagorean proof" [**39**, p. 298] because it was known earlier to the Pythagorean school. It uses the converse to the Alternate Interior Angle theorem, which states that if parallel lines are cut by a transversal, then alternate interior angles with respect to that transversal are congruent to each other. That converse is equivalent, for Hilbert planes, to Hilbert's Euclidean Axiom of Parallels [**20**, Proposition 4.8]. Thus the Pythagorean proof is valid in Pythagorean planes.

How about going in the opposite direction: given a Hilbert plane in which the angle sum of any triangle is 180°, can one prove Hilbert's Euclidean Axiom of Parallels? Using only the Hilbert plane axioms, the answer is no. We know that from the following counterexample displayed by Hilbert's student Max Dehn in 1900:

Let F be a Pythagorean ordered field that is non-Archimedean in the sense that it contains *infinitesimal* elements (nonzero elements x such that $|x| < 1/n$ for all natural numbers n). Such fields exist (see Exercise 18.9 or Proposition 18.2 of [**23**]). Within the Cartesian plane F^2, let Π be the full subplane whose points are those points in F^2 both of whose coordinates are infinitesimal, whose lines are the nonempty intersections with that point-set of lines in F^2, and whose betweenness and congruence relations are induced from F^2. Then Π is a Hilbert plane in which the angle sum of every triangle is 180°. But given point P not on line l, there are infinitely many parallels to l through P—the standard parallel plus all the lines through P whose extension in F^2 meets the extension of l in a point whose coordinates are not both infinitesimal, i.e., in a point which is not part of our model Π.

Definition. A Hilbert plane in which the angle sum of every triangle is 180° is called *semi-Euclidean*.[6]

[6]Voltaire's geometer said "Je vous conseille de douter de tout, excepté que les trios angles d'un triangle sont égaux à deux droits." ("I advise you to doubt everything, except that the three angles of a triangle are equal to two right angles." In Section 1.6 we will doubt even that.)

Thus a Pythagorean plane must be semi-Euclidean, but a semi-Euclidean plane need not be Pythagorean. You might think, from Dehn's example, that the obstruction to proving the Euclidean parallel postulate in a semi-Euclidean plane is the failure of Archimedes' axiom, and you would be semi-correct.

Proposition 1. *If a semi-Euclidean plane is Archimedean, then it is Pythagorean, i.e., Hilbert's Euclidean parallel postulate holds.*

Observe also that while the proposition states that Archimedes' axiom is sufficient, it is certainly not necessary (thus "semi-correct"): the plane F^2, when F is a non-Archimedean Pythagorean ordered field, is Pythagorean and non-Archimedean. So an interesting problem is to find a *purely geometric axiom* $\mathcal{A}$ to replace the axiom of Archimedes such that $\mathcal{A}$ is sufficient *and necessary* for a semi-Euclidean plane to be Pythagorean. Such an axiom $\mathcal{A}$ is what I call "a missing link."

Historically, many attempts to prove Euclid V from his other postulates focused on proving that the angle sum of every triangle equals 180°—assuming Archimedes' axiom in a proof was then considered acceptable. Saccheri, Legendre, and others took that approach in some of their attempted proofs ([**20**, Chapter 5] or [**40**, Chapter 2]).

Proposition 1 can be proved by reductio ad absurdum. Assume on the contrary that there is a line l and a point P not on l such that there is more than one parallel to l through P. We always have one parallel m by the standard construction. Let n be a second parallel to l through P, making an acute angle θ with m, let s be the ray of n with vertex P on the same side of m as l, let Q be the foot of the perpendicular from P to l, and let r be the ray of l with vertex Q on the same side of PQ as s. Using Archimedes' axiom, one can prove that there exists a point R on r such that $\measuredangle PRQ < \theta$ as follows (see Figure 1):

Start with a random point R on r. If $\measuredangle PRQ \geq \theta$, lay off on r segment $RR' \cong PR$, with R between Q and R', so as to form isosceles triangle PRR'. First, by hypothesis, every triangle has angle sum 180°. So $\measuredangle PR'Q = \frac{1}{2}\measuredangle PRQ$ by the congruence of base angles. Repeating this process of halving the angle at each step, one eventually obtains an angle $< \theta$ (by the Archimedean property of angles).

Since ray s does not intersect ray r, ray PR is between ray PQ and ray s—otherwise the Crossbar theorem [**20**, p. 116] would be violated. Then

$$\measuredangle PRQ + \measuredangle RPQ < \theta + \measuredangle RPQ < 90°.$$

Hence right triangle PRQ has angle sum less than 180°, contradicting our hypothesis that the plane is semi-Euclidean. ■

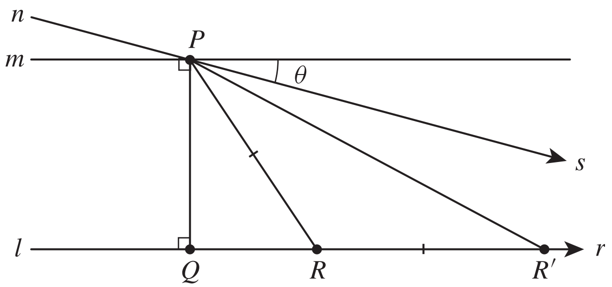

Figure 1. Proof of Proposition 1.

The crucial geometric statement $\mathcal{C}$ used in this argument is the following:

$\mathcal{C}$. *Given any segment PQ, line l through Q perpendicular to PQ, and ray r of l with vertex Q, if θ is any acute angle, then there exists a point R on r such that $\measuredangle PRQ < \theta$.*

Note that if S is any point further out on l—i.e., such that R lies between Q and S—then we also have $\measuredangle PSQ < \measuredangle PRQ < \theta$, by the Exterior Angle theorem [**20**, Theorem 4.2]. So statement $\mathcal{C}$ says that *as R recedes from Q along a ray of l with vertex Q, $\measuredangle PRQ$ goes to zero*. Note also that we have proved that if $\mathcal{C}$ holds in a Hilbert plane which is *non-Euclidean* in the sense that the negation of Hilbert's Euclidean axiom of parallels holds, then that plane has a triangle whose angle sum is $< 180°$.

Is statement $\mathcal{C}$ a missing link? Yes!

Theorem 1. *A Hilbert plane is Pythagorean if and only if it is semi-Euclidean and statement $\mathcal{C}$ holds.*

Proof. It only remains to show that the Euclidean parallel postulate implies $\mathcal{C}$. Assume the negation of $\mathcal{C}$. Then there exist P, Q, and l as above and an acute angle θ such that $\measuredangle PRQ > \theta$ for all $R \neq Q$ on a ray r of l emanating from Q. (We can write $>$ and not just $\geq$ because if we had equality for some R, then any point R' with R between Q and R' would satisfy $\measuredangle PR'Q < \theta$ by the Exterior Angle theorem.) Emanating from vertex P there is a unique ray s making an angle θ with the standard line m through P perpendicular to PQ and such that s is on the same side of m as l and of PQ as r. If s does not meet r, then the line containing s is a second parallel to l through P. If s does meet r in a point R, then the angle sum of right triangle PQR is greater than $180°$. In either case we have a contradiction of the Euclidean parallel postulate. ■

Now **statement $\mathcal{C}$ is a consequence of Aristotle's axiom**, which can be seen as follows: Let PQ, l, r, and θ be given as in statement $\mathcal{C}$. Apply Aristotle's axiom with challenge segment PQ and angle θ to produce Y on a side of angle θ and X the foot of the perpendicular from Y to the other side of θ such that $XY > QP$. Say O is the vertex of angle θ. Lay off segment XY on ray QP starting at Q and ending at some point S, and lay off segment XO on ray r of l starting at Q and ending at some point R. By the SAS axiom, $\measuredangle QRS$ is congruent to θ. Since $XY > QP$, P is between Q and S, so ray RP is between rays RQ and RS. Hence $\measuredangle QRP < \theta$, which is the conclusion of statement $\mathcal{C}$. ■

Thus Aristotle's axiom $\mathcal{A}$ is also a missing link:

Theorem of Proclus. *A semi-Euclidean plane is Pythagorean if and only if it satisfies Aristotle's axiom.*

For the easy proof that Euclid V implies $\mathcal{A}$, see [**20**, Corollary 2, p. 180].

I have named this theorem after Proclus because he was the first to recognize the importance of Aristotle's axiom in the foundations of geometry when he used it in his failed attempt to prove Euclid V ([**39**, p. 291] or [**20**, p. 210]). His method provides the proof of sufficiency I found [**20**, p. 220], in which Aristotle's axiom is used directly, without the intervention of $\mathcal{C}$. Also, the proof above that $\mathcal{A}$ implies $\mathcal{C}$ used the fact that any two right angles are congruent, which was Euclid's fourth postulate; that postulate became a theorem in Hilbert's axiom system, and Hilbert's proof of it is based on the idea Proclus proposed ([**39**, p. 147] and [**20**, Proposition 3.23]).

Summary. We have studied in these sections the following strictly decreasing chain of collections of models: Hilbert planes ⊃ semi-Euclidean planes ⊃ Pythagorean planes ⊃ Euclidean planes ⊃ Archimedean Euclidean planes ⊃ {real Euclidean plane}.

1.6. Aristotle's Axiom in Non-Euclidean Geometry. Warning: Readers unfamiliar with the axiomatic approach to hyperbolic geometry may find this section heavy going. They could prepare by reading [**23**, Chapter 7] and/or [**20**, Chapters 4, 6, and 7].

Call a Hilbert plane *non-Euclidean* if the negation of Hilbert's Euclidean parallel postulate holds; it can be shown that this implies that for every line l and every point P not on l, there exist infinitely many lines through P that do not intersect l.

A surprise is that without real number coordinatization (i.e., without Dedekind's axiom) there are many strange non-Euclidean models besides the classical hyperbolic one (but all of them have been determined by W. Pejas—see [**37**], [**27**], or [**20**, Appendix B, Part II]).

In a general Hilbert plane, quadrilaterals with at least three right angles are convex and are now called *Lambert quadrilaterals* (although Johann Lambert was not the first person to study them—Girolamo Saccheri and medieval geometers such as Omar Khayyam worked with them). Return for a moment to a semi-Euclidean plane. There, the angle sum of any convex quadrilateral is 360°, as can be seen by dissecting the convex quadrilateral into two triangles via a diagonal. In particular, if a quadrilateral in a semi-Euclidean plane has three right angles, then the fourth angle must also be right, so that quadrilateral is by definition a *rectangle*.

Conversely, suppose every Lambert quadrilateral is a rectangle. To prove that the angle sum of every triangle is 180°, it suffices (by dropping an appropriate altitude in a general triangle) to prove this for a right triangle ABC. Let the right angle be at C. Erect a perpendicular t to CB at B, and drop a perpendicular from A to t, with D the foot of that perpendicular. Since $CBDA$ is a Lambert quadrilateral, its fourth angle at A is right, by hypothesis. By [**20**, Corollary 3, p. 180], opposite sides of a rectangle are congruent, so the two triangles obtained from a diagonal are congruent by SSS, and hence each has angle sum 180°. Thus the plane is semi-Euclidean.

This proves the equivalence in part (i) of the following fundamental theorem.

Uniformity Theorem. *A Hilbert plane must be one of three distinct types:*

(i) *The angle sum of every triangle is* 180° *(equivalently, every Lambert quadrilateral is a rectangle).*

(ii) *The angle sum of every triangle is* $< 180°$ *(equivalently, the fourth angle of every Lambert quadrilateral is acute).*

(iii) *The angle sum of every triangle is* $> 180°$ *(equivalently, the fourth angle of every Lambert quadrilateral is obtuse).*

Moreover, all three types of Hilbert plane exist.

For a proof of uniformity, see [**20**, Major Exercises 5–8, pp. 202–205]. The footnote on p. 43 of [**28**] credits Max Dehn with having proved it, and states that later proofs were produced by F. Schur and J. Hjelmslev. Actually, Saccheri had this result back in 1733—except that his proof used continuity; Hartshorne [**23**, p. 491] states that Lambert proved it without using continuity. In planes of types (ii) and (iii), similar triangles must be congruent [**20**, proof of Proposition 6.2], so there is no similarity theory. W. Pejas ([**37**] or [**27**]) showed that in a plane of obtuse type (iii), the *excess* by which the angle sum of a triangle is $> 180°$ is infinitesimal.

Saccheri and Lambert did not know of the existence of planes of obtuse angle type (iii). Again, it was Dehn who gave an example of a non-Archimedean Hilbert plane satisfying obtuse angle hypothesis (iii) [**20**, p.189, footnote 8]. Another example is the infinitesimal neighborhood of a point on a sphere in F^3, where F is a non-Archimedean ordered Pythagorean field [**23**, Exercise 34.14]. These examples contradict the assertion made in some books and articles that the hypothesis of the obtuse angle is inconsistent with the first four postulates of Euclid. In fact, it is consistent with the thirteen axioms for Hilbert planes (which imply those four postulates).

These examples again suggest that Archimedes' axiom would eliminate the obtuse angle type (iii), and indeed that was proved by Saccheri and Legendre ([**28**, Theorem 35] or [**23**, Theorem 35.2]). However, **the weaker axiom of Aristotle suffices to eliminate the obtuse angle type (iii)**. Namely, Aristotle's axiom implies $\mathcal{C}$, and we showed in the proof of Theorem 1, Section 1.5, that in a non-Euclidean Hilbert plane, $\mathcal{C}$ implies that there exists a triangle whose angle sum is $< 180°$. For a direct proof not using $\mathcal{C}$ but using the idea of Proclus, see [**20**, p. 185], the Non-Obtuse Angle theorem.

Turning now to elementary plane hyperbolic geometry, Hilbert axiomatized it without bringing in real numbers by adding just his hyperbolic parallel axiom to the thirteen axioms for a Hilbert plane. His axiom states that given any line l and any point P not on l, *there exists a limiting parallel ray s to l emanating from P making an acute angle* with the perpendicular ray PQ dropped from P to l, where Q again is the foot of the perpendicular from P to l ([**20**, p. 259], [**23**, axiom L, p. 374], or [**28**, Appendix III]). Ray s is by definition "limiting" in that *s does not intersect l and every ray emanating from P between s and perpendicular ray PQ does intersect l*. The Crossbar theorem [**20**, p. 116] tells us that for any ray s' emanating from P such that ray s lies between s' and perpendicular ray PQ, s' does not intersect l. The acute angle between s and ray PQ is called *the angle of parallelism* for segment PQ; its congruence class depends only on the congruence class of segment PQ. Reflecting across line PQ, the other limiting parallel ray from P to l is obtained. A *hyperbolic plane* is, by definition, a Hilbert plane satisfying *Hilbert's hyperbolic parallel axiom* just stated. Hilbert's axiom made explicit the limiting existence assumption tacitly made by Saccheri, Gauss, Bolyai, and Lobachevsky.

Hilbert pointed out that his approach was a breakthrough over the earlier ones by Felix Klein and Bolyai-Lobachevsky in that it did not resort to three dimensions to prove its theorems for the plane; he also did not use Archimedes' axiom.

Hilbert's hyperbolic parallel axiom follows, in non-Euclidean planes, from Dedekind's axiom [**20**, Theorem 6.2], but without Dedekind's axiom, there exist many non-Euclidean Hilbert planes which are not hyperbolic (e.g., Dehn's semi-Euclidean non-Euclidean one or the ones of types (ii) and (iii) described in [**20**, Appendix B, Part II], including Pejas' example of a type (ii) plane which is Archimedean and not hyperbolic).

All hyperbolic planes are known (up to isomorphism): they are the isomorphic Klein and Poincaré models [**20**, Chapter 7] coordinatized by arbitrary Euclidean fields. Again, Hilbert cleverly showed how the field is extracted from the geometry: he constructed the coordinate field F as his *field of ends* ([**20**, Appendix B], [**23**, §41], or [**28**, Appendix III]). The "ends" or "ideal points" are defined essentially as equivalence classes of rays under the relationship of limiting parallelism [**20**, pp. 276–279]; they form a *conic at infinity* in the projective completion of the hyperbolic plane [**20**, p. 286], and F excludes one end denoted ∞. From the way multiplication of ends is defined, the reason every positive has a square root in F is that every segment has a midpoint [**20**, p. 576]. That in turn guarantees that the Circle-Circle axiom holds [**23**, Corollary 43.4], unlike the case of Pythagorean planes.

Hyperbolic planes satisfy the acute angle hypothesis (ii) in the Uniformity theorem [**23**, Corollary 40.3]. The acute angle hypothesis alone implies that each side adjacent to the acute angle in a Lambert quadrilateral is greater than the opposite side [**20**, Corollary 3 to Proposition 4.13]. That result is the key to Saccheri's proof that Archimedes' axiom implies Aristotle's axiom for planes of acute type (Prop. 21 of his *Euclides Vindicatus* or [**23**, Proposition 35.6])—namely, he showed that if an arbitrarily chosen perpendicular segment from the given side of the acute angle to the other side is not greater than challenge segment AB, then by repeatedly doubling the distance from the vertex along the given side, eventually a perpendicular segment $> AB$ will be obtained. Planes of obtuse type (iii) are non-Archimedean and non-Artistotelian, as was mentioned above. As for type (i), if a semi-Euclidean plane is Archimedean, then by Proposition 1 in Section 1.5, Euclid V holds, and Aristotle's axiom follows from that. Thus **Archimedes' axiom implies Aristotle's axiom in all Hilbert planes**. Since Aristotle's axiom holds in all Pythagorean and hyperbolic planes, including the non-Archimedean ones, the converse implication is invalid.

János Bolyai gave a straightedge and compass construction of the limiting parallel ray in a hyperbolic plane ([**20**, Figure 6.11, p. 259] or [**23**, Proposition 41.10]).

Bolyai's construction can be carried out in any Hilbert plane satisfying the acute angle hypothesis and the Line-Circle axiom, but the ray obtained need not be a limiting parallel ray without a further hypothesis on the plane (see [**18**], where is shown the special property of that ray that shocked Saccheri: the "common perpendicular at infinity"). Friedrich Schur gave a counterexample similar to Dehn's: in the infinitesimal neighborhood of the origin in the Klein model coordinatized by a non-Archimedean Euclidean field, limiting parallel rays do not exist. This example suggests that taking Archimedes' axiom as a further hypothesis would make Bolyai's construction work, and indeed that is the case. Again Archimedes' axiom is not necessary and Aristotle's axiom is a missing link.

Advanced Theorem. *A Hilbert plane satisfying the acute angle hypothesis (ii) and the Line-Circle axiom is hyperbolic if and only if it satisfies Aristotle's axiom; in that case, Bolyai's construction provides the limiting parallel rays.*

That Aristotle's axiom holds in hyperbolic planes follows from the fact previously mentioned that a side adjacent to the acute angle in a Lambert quadrilateral is greater than the opposite side—see [**20**, Exercise 13, p. 275] or [**23**, Proposition 40.8]. That Bolyai's construction provides the limiting parallel ray in a hyperbolic plane is part of Engel's theorem [**20**, Theorem 10.9], which provides other important constructions as well.

The sufficiency part of this theorem is much deeper than its Euclidean analogue—Proclus' theorem—because its only proof known so far depends on W. Pejas' classification ([**37**] or [**27**]) of Hilbert planes (hence it is "advanced"). Pejas' work implies that the plane can be embedded as a "full" submodel of a Klein model; then Aristotle's axiom, via its corollary $\mathcal{C}$, implies maximality of the submodel, so that it equals the Klein model. See [**19**] for details of my proof and see [**20**, Appendix B, Part II] for a description of Pejas' great work (built upon earlier work by F. Bachmann, G. Hessenberg, and J. Hjelmslev).

A slightly stronger version of the Advanced theorem is that a non-Euclidean Hilbert plane is hyperbolic if and only if it satisfies the Line-Circle axiom and Aristotle's axiom (that the type of the plane is (ii) is provable). Thus we now have an elegant axiomatization of what Bolyai intended by his study of the common part of plane

Euclidean and hyperbolic geometries (at least the elementary part): the 13 axioms for a Hilbert plane plus the Line-Circle axiom plus Aristotle's axiom.

Bolyai's construction gives the angle of parallelism corresponding to a given segment PQ, which Lobachevsky denoted $\Pi(PQ)$. Conversely, given an acute angle θ, there is a very simple straightedge and compass construction of a segment PQ such that $\theta = \Pi(PQ)$—see [**20**, George Martin's theorem, p. 523].

Note. Hilbert gets the credit for the main ideas in the foundations of plane hyperbolic geometry without real numbers, but he only sketched much of what needed to be done. The details were worked out by others as described in the introduction to Pambuccian [**35**]; the best exposition of those details is Hartshorne [**23**, Chapter 7], where a remarkable new hyperbolic trigonometry without real numbers is presented. Using it, all arguments in other treatises using classical hyperbolic trigonometry can be rephrased so as to avoid their apparent dependence on real numbers [**23**, Exercise 42.15].

Summary. We have the following strictly decreasing chain of collections of planes: Non-Euclidean Hilbert planes $\supset$ Hilbert planes satisfying the acute angle hypothesis $\supset$ hyperbolic planes $\supset$ Archimedean hyperbolic planes $\supset$ {real hyperbolic plane}.

2. BOLYAI'S CIRCLE-ANGLE CONSTRUCTION AND REGULAR POLYGONING A CIRCLE. Rectangles, in particular squares, do not exist in hyperbolic planes. So the classical problem of "squaring" a circle must be reinterpreted there with a regular 4-gon replacing the square. In the *real* hyperbolic plane, the area of a circle can be defined as the limit of the areas of the regular polygons inscribed in it as the number of sides goes to infinity (so here the real numbers definitely are necessary). The answer is $4\pi \sinh^2(r/2)$ [**20**, Theorem 10.7], where r is the length of the radius when lengths are normalized so that *Schweikart's segment*, a segment whose angle of parallelism is $\pi/4$, has length arcsinh 1. From this formula one sees that areas of circles are unbounded.

János Bolyai found a remarkable construction ("construct" henceforth refers to straightedge and compass) of an auxiliary acute angle θ associated to a circle, in terms of which the formula for area of the circle takes the familiar form πR^2, where $R = \tan\theta$. So "par abus de langage" I will call *R the Euclidean radius of the hyperbolic circle*. Conversely, given acute angle θ, a segment of length r can be constructed from it so that $\tan\theta$ is the Euclidean radius of the circle of hyperbolic radius r. See [**20**, Chapter 10, Figure 10.32 and Project 2] or [**16**, pp. 69–75]. The "radii" are related by $R = 2\sinh(r/2)$.

Some popular writers have claimed that Bolyai "squared the circle" in a hyperbolic plane. That is not true if by "the" circle is meant an arbitrary circle. One trivial reason is that normalized areas of regular 4-gons, equal to their defects [**23**, §36], are bounded by 2π (using radian measure, the *defect* of a 4-gon is by definition 2π minus the angle sum), whereas areas of circles are unbounded. But even if the circle has area $< 2\pi$ and has a constructible radius, the regular 4-gon with the same area may not be constructible—see W. Jagy [**30**, Theorem B] for a counterexample due to N. M. Nesterovich.

The question of determining exactly when a circle and a regular 4-gon having the same area are *both* constructible in a hyperbolic plane has a surprising answer. As J. Bolyai argued, using an argument that was completed by Will Jagy [**30**], the answer is in terms of those integers $n > 1$ for which the angle $2\pi/n$ is constructible in the smallest Euclidean plane. Gauss determined those numbers, so let's call them *Gauss numbers*: they are the numbers $n > 1$ such that the only *odd* primes—if any—

occurring in the prime factorization of n occur to the first power and are *Fermat primes* [**23**, Theorem 29.4]. The only Fermat primes known at this time are 3, 5, 17, 257, and 65,537. The prime 2 may occur with any exponent ≥ 0 in the factorization of a Gauss number, so there are infinitely many Gauss numbers.[7]

This problem of constructing both figures comes down to constructing two angles—the auxiliary angle θ for which the circle has area $\pi \tan^2 \theta$ and the acute corner angle σ of the regular 4-gon. The equation for equal areas is then

$$\pi \tan^2 \theta = 2\pi - 4\sigma. \tag{$*$}$$

Bolyai's Construction Theorem (Jagy's Theorem A). *Suppose that a regular* 4*-gon with acute angle σ and a circle in the hyperbolic plane have the same area $\omega < 2\pi$. Then both are constructible if and only if σ is an integer multiple of $2\pi/n$, where n is a Gauss number > 2. In that case, if $R = \tan\theta$ is the Euclidean radius of the circle, then R^2 is a rational number, which, when expressed in lowest terms, has denominator a divisor of n.*

For example, when $\omega = \pi$, $n = 8$: the hyperbolic circle of Euclidean radius 1 and the regular 4-gon with acute angle $\pi/4$ are both constructible [**20**, p. 521].

The proof plays off the ambiguity in equation ($*$), whereby the area of the 4-gon can be considered either an angle—its defect—or a real number—the radian measure of its defect. Jagy's insight was to apply the Gelfond-Schneider theorem about transcendental numbers to prove that R^2 is *rational* if both the circle and the regular 4-gon are constructible.

To dispense with the trivial reason mentioned above, I asked Jagy to consider regular m-gons having the same area as a circle, for arbitrary $m \geq 4$. A regular m-gon in the real hyperbolic plane can have any corner angle σ such that $m\sigma < (m-2)\pi$. *That regular m-gon will be constructible if and only if σ is a constructible angle and π/m is constructible*: by joining the center of the m-gon to the midpoint and an endpoint of one of its sides, a right triangle is formed with one acute angle $\sigma/2$ and the other π/m. Constructing the regular m-gon comes down to constructing that right triangle, which can be done if and only if the acute angles $\sigma/2$ and π/m are constructible [**20**, p. 506, Right Triangle Construction theorem]. Thus, just as in the Euclidean case, *m must be a Gauss number*. The equation of areas to consider then becomes

$$\pi \tan^2 \theta = (m-2)\pi - m\sigma.$$

When the circle of Euclidean radius R and the regular m-gon with corner angle σ are both constructible, Jagy's argument on p. 35 of [**30**] using the Gelfond-Schneider theorem shows that $R = \tan\theta$ is the square root of a rational number k/n in lowest terms and n is a Gauss number or $n = 1$. Solving for σ gives

$$\sigma = \frac{(m-2)}{m}\pi - \frac{k}{mn}\pi$$

so that when both are constructible, $\frac{k}{mn}\pi$ must be constructible. For such n, mn is only a Gauss number when $\gcd(m, n)$ is a power of 2 (including the 0th power). Thus when

[7] An angle is constructible in the Euclidean plane coordinatized by the field $\mathbf{K}$ of constructible numbers if and only if it is constructible in the hyperbolic plane coordinatized by $\mathbf{K}$ [**20**, p. 587]. If a marked ruler instead of a straightedge is used in constructions, then the prime 3 can occur to any power in the factorization of n, and the other odd primes that may occur to the first power must be *Pierpont primes*—see [**23**, Corollary 31.9] and [**4**].

m is a power of 2, the result is the same as before. Jagy expresses the general result as follows:

Jagy's Theorem. *Let a circle and a regular m-gon in the real hyperbolic plane with equal normalized area ω be given. Then both are constructible if and only if the following four conditions hold:*

1. $\omega < (m-2)\pi$,
2. *ω is a rational multiple of π, and if that rational multiple is k/n in lowest terms, n is a Gauss number or $n = 1$,*
3. *m is a Gauss number, and*
4. *m and n have no odd prime factors in common.*

For example, if m is any Gauss number ≥ 5 and we again consider the constructible circle γ of area π, the regular m-gon with angle $\pi(m-3)/m$ is constructible and also has area π. Thus γ can not only be "squared" in the real hyperbolic plane, it can also be "regular pentagoned," "regular hexagoned," "regular octagoned," etc. The trebly asymptotic triangle also has area π, and, since it can be constructed via Hilbert's construction of lines of enclosure, γ can also be considered to be "regular triangled." In fact, Jagy prefers to write $\omega \leq (m-2)\pi$ in condition 1 so as to allow asymptotic m-gons.

After presenting his result for $m = 4$ at the end of his immortal Appendix, Bolyai referred admiringly to "the theory of polygons of the illustrious Gauss (remarkable invention of our, nay of every age)." Tragically, Gauss did not return the compliment by assisting the mathematical career of Bolyai in any way. Gauss feared "the howl of the Boeotians" should he publicly endorse non-Euclidean geometry [**20**, p. 244]. Bolyai years later sharply criticized Gauss for his timidity [**20**, p. 242].

In addition to the results above due to Bolyai and Jagy, constructible numbers are again key to solving problems about constructions with straightedge and compass in a hyperbolic plane. Given two perpendicular axes and a choice of ends on them labeled 0 and ∞ on one axis and 1 and -1 on the other, I conjectured and Hartshorne proved [**20**, p. 587] that a segment in a hyperbolic plane is constructible from that initial data if and only if Hartshorne's *multiplicative length* of that segment [**23**, Proposition 41.7] is a constructible number in the field of ends based on that data. That multiplicative length is a key discovery which Hartshorne and Van Luijk have applied to algebraic geometry and number theory [**26**].

3. UNDECIDABILITY AND CONSISTENCY OF ELEMENTARY GEOMETRY

3.0. Metamathematical Aspects. In this section we will sketch some important metamathematical results about our axiom systems. Readers who are not trained in mathematical logic may refer to any logic textbook or to wikipedia.org for good explanations of concepts that may be unfamiliar, but those readers can still understand the gist of this section.

3.1. Relative Consistency. It has been known since the Euclidean plane models of Beltrami-Klein and Poincaré were exhibited in the late nineteenth century that *if plane Euclidean geometry is a consistent theory, then so is plane hyperbolic geometry* [**20**, Chapter 7]. Those isomorphic models ended the doubts about the validity of this alternative geometry. The Poincaré models have become especially important for applications of hyperbolic geometry to other branches of mathematics and to Escher's art [**20**, pp. 382ff].

The converse result was shown in 1995 by A. Ramsey and R. D. Richtmyer; their model is explained in [**20**, pp. 514–515] (see also [**20**, Project 1, p. 537]). Previously, only a model of plane Euclidean geometry in hyperbolic three-space had been exhibited: the *horosphere*, with its *horocycles* as the interpretation of "lines."

In the Ramsey-Richtmyer model, an origin O is chosen in the hyperbolic plane and then all the hyperbolic lines through O and all the equidistant curves for only those lines are the interpretation of "Euclidean lines" (the *equidistant curve* through a point P not on a line l consists of all points R on the same side of l as P such that P and R are at the same perpendicular distance from l). "Euclidean points" are interpreted as all the points of the hyperbolic plane. "Betweenness" is induced by the betweenness relation in the hyperbolic plane. "Congruence" is more subtle and is explained in the reference above. Curiously, this model shows that ancient geometers such as Clavius, who thought that equidistant curves were Euclidean lines ([**23**, p. 299] and [**20**, pp. 213–214]), were partially correct, provided they were working in a hyperbolic plane!

Once he established the relative consistency of hyperbolic geometry via one of his Euclidean models, Poincaré wrote [**17**, p. 190]:

> "No one doubts that ordinary [Euclidean] geometry is exempt from contradiction. Whence is the certainty derived, and how far is it justified? That is a question upon which I cannot enter here, because it requires further work, but it is a very important question."

The fact that Poincaré considered this question "very important" and said that it "requires further work" indicates to me that he was asking a mathematical question (as well as a philosophical question). Logicians have solved this consistency problem for elementary Euclidean geometry. Section 3.5 has a brief discussion of their work.

3.2. Undecidability. Roughly speaking, a theory is called *undecidable* if there is no algorithm for determining whether or not an arbitrary statement in the theory is provable in that theory.

The elementary theory of Euclidean planes is undecidable. Victor Pambuccian [**34**, p. 67] expressed this undecidability result by saying that in general "there is no way to bulldoze one's way through a proof via analytic geometry," as was previously believed generally possible in an article he cited there. Thus elementary Euclidean geometry is genuinely creative, not mechanical.

Here is a rough indication of why that is so: Each Euclidean plane is isomorphic to the Cartesian plane coordinatized by its field of segment arithmetic, which is a Euclidean field F. As Descartes showed, every geometric statement about that plane translates into an algebraic statement about F. Descartes hoped to be able to solve every geometric problem by applying algebra to the translated problem. So the theory of Euclidean planes is thereby included in the theory of Euclidean fields.

But M. Ziegler [**46**] has proved that the theory of Euclidean fields is undecidable. In fact, Ziegler proved that *any finitely axiomatized first-order theory of fields having the real number field* $\mathbb{R}$ *as a model must be undecidable*. This includes the theory of fields itself, the theory of ordered fields, and the theories of Pythagorean ordered fields and Euclidean fields.

The theory of Cartesian planes coordinatized by any of those types of fields is then undecidable as well, because the field can be recovered from the geometry—see [**41**, Section II.3, pp. 218–263].

It is necessary to provide an equivalent[8] first-order axiomatization **E** of our theory of elementary Euclidean plane geometry in order to be able to apply Ziegler's theorem. First-order systems, first emphasized by Thoralf Skolem, are explained on the first pages of Pambuccian [**35**]. In first-order logic, quantification ("for all" or "there exists") is only allowed over individuals, not over sets of or relations among individuals.

3.3. Tarski-elementary Euclidean Geometry. Alfred Tarski's idea of what was "elementary" in geometry differed from what we've been discussing in previous sections. He regarded as "elementary" that part of Euclidean geometry which can be formulated and established without the help of any set-theoretic devices, hence within first-order logic.

Tarski's various first-order axiom systems for plane Euclidean geometry were based on the single primitive notion of "point" and on two undefined relations among points first introduced by O. Veblen: *betweenness* for three points and *equidistance* (or *congruence*) for two pairs of points (think of each pair as the endpoints of a segment). The relation of *collinearity* of three points is defined in terms of betweenness (A, B, and C are collinear if and only if one of them is between the other two), so he did not need "line" or "incidence" as primitive notions.

But Euclid V or Hilbert's Euclidean parallel postulate or the converse to the Alternate Interior Angle theorem or other familiar equivalent statements all refer to lines, so how did Tarski express an equivalent to a Euclidean parallel postulate in his system? In one version he used the less-familiar equivalent statement that for any three noncollinear points, there exists a fourth point equidistant from all three of them (that point is the *circumcenter* of the triangle formed by the three points—see [**20**, Chapter 5, Exercise 5 and Chapter 6, Exercise 10]).

Tarski worked to develop a first-order replacement for real Euclidean geometry. The axiom that makes Euclidean geometry "real" is Dedekind's cut axiom; cuts are infinite sets of points (so Dedekind's axiom is "second-order"). In order to formulate a first-order replacement for those sets, Tarski had to introduce a countably infinite set of axioms, all of the same form, referred to as the *continuity axiom schema*. It is explained in [**15**], along with his original twenty ordinary axioms (later reduced to fewer than twenty). (See also [**35**] in which Pambuccian reviews some of the history and presents first-order axiomatizations of "absolute" and hyperbolic geometries.)

Geometrically, Tarski-elementary plane geometry certainly seems mysterious, but the representation theorem illuminates the analytic geometry underlying it: its models are all Cartesian planes coordinatized by *real-closed fields*. A real-closed field can be characterized in at least seven different ways; for our purpose, the simplest definition is "a Euclidean field F in which every polynomial of odd degree in one indeterminate, with coefficients in F, has a root." Six other characterizations as well as much more information can be found on wikipedia.org. Of course $\mathbb{R}$ is a real-closed field, but so is its countable subfield of real algebraic numbers. Every Euclidean field has an algebraic extension which is real-closed and is unique up to isomorphism. So we can insert a different link into our strictly decreasing chain of collections of models concluding Section 1.5. It now terminates with

$$\text{Euclidean planes} \supset \text{Tarski-elementary Euclidean planes} \supset \{\text{real Euclidean plane}\}.$$

[8] A first-order theory **E** is "equivalent" to ours if its models are the same as the models we earlier called "Euclidean planes"—models of the 14 Pythagorean plane axioms plus the Circle-Circle axiom. To say the models are "the same" requires providing translation instructions to correctly interpret each theory's language in the language of the other.

The primary significance of Tarski-elementary geometry is its three metamathematical properties, the first two of which are opposite from the elementary theory we have been discussing: it is *deductively complete* and *decidable*, meaning that every statement in its language is either a theorem (i.e., provable) or its negation is a theorem, and there is an algorithm to determine which is the case. See [**15**] and [**5**]. The third property, which both theories share, is *consistency*, to be discussed in Section 3.5.

3.4. Very Brief Historical/Philosophical Introduction to Finitary Consistency Proofs. Hilbert invented *syntactic* proofs of consistency—i.e., proofs that are not based on set-theoretic models of the theory but on "proof theory."

Such proofs are only as worthwhile epistemologically as the principles upon which they are based. Hilbert was well aware of this limitation, so he proposed the doctrine that the most convincing consistency proofs are the ones that are "finitary." Although it was not originally clear how Hilbert's finitism was to be precisely formulated, mathematical logicians now generally agree that we can take it to mean proving within *primitive recursive arithmetic* (PRA). Paul Cohen [**5**] gave a finitary proof, in this sense, of the decidability[9] of RCF, the first-order theory of real-closed fields.

It follows from the pioneering work of Kurt Gödel in the early 1930s that no such finitary proof of the consistency of the full Peano Arithmetic (PA) is possible.

In 1936, Gerhard Gentzen published a nonfinitary consistency proof for PA using transfinite induction along the countable ordinal ε_0, the smallest solution to the ordinal equation

$$\varepsilon = \omega^{\varepsilon}$$

where ω is the ordinal of the natural numbers. Lest the reader get the impression that Gentzen's proof is heavily set-theoretic, according to Martin Davis (personal communication) it is just a matter of allowing induction with respect to a particular computable well ordering of the natural numbers. Gentzen's new ideas led to a flourishing of proof theory.

For an excellent discussion of these matters, see the Stanford Encyclopedia of Philosophy article [**42**].

The next section reports on recent successes in providing *finitary* proofs for the consistency of elementary Euclidean geometry, in stark contrast to the impossibility of such proofs for Peano Arithmetic. In that sense Euclid's geometry is simpler than number theory.

3.5. Consistency of Tarski-elementary Geometry and Our Elementary Geometry. One first-order version **E** of elementary Euclidean geometry is a finitely axiomatized subtheory of Tarski's theory—namely, keep Tarski's twenty ordinary axioms and replace Tarski's continuity axiom schema with the single Segment-Circle axiom asserting that if point a lies inside a given circle and point b lies outside the circle (with center c and radius cp), then the circle intersects segment ab. The circle is not described as a set of points but only indirectly in terms of its center and a radius. The explicit first-order version of this is

$$(\forall\, abcpqr)[B(cqp) \wedge B(cpr) \wedge ca \equiv cq \wedge cb \equiv cr \Rightarrow (\exists\, x)(cx \equiv cp \wedge B(axb))],$$

where $\equiv$ denotes the congruence relation and $B(axb)$ means x is between a and b.

[9]Comparing this decidability result with Ziegler's theorem, we see that RCF cannot be finitely axiomatized.

So **E** is consistent if Tarski's theory is. The consistency of the latter follows from the consistency of the theory RCF of real-closed fields. There are finitary proofs of the consistency of RCF, such as the one Harvey Friedman posted on his website [**14**], using only *exponential function arithmetic*, a sub-theory of PRA. Fernando Ferreira, using work he coauthored with Antonio Fernandes [**13**], has another proof of the consistency of RCF using the stronger *Gentzen arithmetic*, also a sub-theory of PRA.

Much earlier, Hilbert and Bernays [**29**, pp. 38–48] gave a proof that their geometric system—essentially what I called Pythagorean geometry (no continuity axiom)—is consistent, based on the model $\mathcal{E}$ coordinatized by the field **K** of constructible numbers. In [**45**], F. Tomás gave a different proof for a slightly different theory.

Tarski himself proved the consistency of his theory, and his proof may be finitary according to a private communication I received from Steven Givant, who refers to footnote 15 in [**44**].

We remain with the psychological/philosophical puzzle of whether those finitary proofs of consistency provide the certainty Poincaré questioned. Most of us take for granted the consistency of elementary Euclidean geometry because of 2,400 years of experience of not finding any contradictions plus all the successful practical applications of that theory. Some people add that their geometric visualizing ability gives them certainty. It is nevertheless a remarkable technical accomplishment to have developed finitary proofs of its consistency in mathematical logic.

ACKNOWLEDGMENTS. I am very grateful to geometers Robin Hartshorne and William C. Jagy,[10] to mathematical logicians Michael Beeson, Andreas Blass, Solomon Feferman, and Fernando Ferreira, and especially to Victor Pambuccian, who is both a geometer and a logician, for their valuable communications regarding this survey. I also thank the referees for their helpful suggestions.

REFERENCES

1. M. Beeson, Constructive geometry, to appear in *Proceedings of the Tenth Asian Logic Colloquium, Kobe, Japan, 2008*, T. Arai, ed., World Scientific, Singapore, 2009.
2. G. Birkhoff and M. K. Bennett, Hilbert's "Grundlagen der Geometrie," *Rendiconti del Circolo Matematico di Palermo*, Serie II, **XXXVI** (1987) 343–389. doi:10.1007/BF02844894
3. K. Borsuk and W. Szmielew, *Foundations of Geometry: Euclidian and Bolyai-Lobachevskian Geometry* (trans. from Polish by E. Marquit), North-Holland, Amsterdam, 1960.
4. C. Caldwell, The Prime Glossary: Pierpont prime (2009), available at http://primes.utm.edu/glossary/xpage/PierpontPrime.html.
5. P. J. Cohen, Decision procedures for real and p-adic fields, *Comm. Pure Appl. Math.* **22** (1969) 131–151. doi:10.1002/cpa.3160220202
6. B. Dragovich, Non-Archimedean geometry and physics on adelic spaces, in *Contemporary Geometry and Related Topics*, eds. N. Bokan et al., World Scientific, Singapore, 2004, 140–158.
7. M. Detlefsen, Purity as an ideal of proof, in *The Philosophy of Mathematical Practice*, P. Mancosu, ed., Oxford University Press, Oxford, 2008, 179–197.
8. P. Ehrlich, From completeness to Archimedean completeness. An essay in the foundations of Euclidean geometry, *Synthese* **110** (1997) 57–76. doi:10.1023/A:1004971123925
9. ———, The rise of non-Archimedean mathematics and the roots of a misconception. I. The emergence of non-Archimedean systems of magnitudes, *Arch. Hist. Exact Sci.* **60** (2006) 1–121. doi:10.1007/s00407-005-0102-4
10. F. Enriques, *Fragen der Elementargeometrie, I Teil: Die Geometrischen Aufgaben, ihre Lösung und Lösbarkeit*, Teubner, Leipzig, 1911.
11. Euclid, *The Thirteen Books of The Elements*, 3 vols., (tr. and annotated by T. L. Heath), Dover, New York, 1956.
12. G. Ewald, *Geometry: An Introduction*, Wadsworth, Belmont, CA, 1971.

[10]Jagy can be contacted at jagy@msri.org or at P. O. Box 2146 , Berkeley, CA 94702-0146.

13. A. Fernandes and F. Ferreira, Groundwork for weak analysis, *J. Symbolic Logic* **67** (2002) 557–578. doi:10.2178/jsl/1190150098
14. H. Friedman, A consistency proof for elementary algebra and geometry (1999), available at http://www.math.ohio-state.edu/~friedman/pdf/ConsRCF8%5B1%5D.23.99.pdf.
15. S. Givant and A. Tarski, Tarski's system of geometry, *Bull. Symbolic Logic* **5** (1999) 175–214. doi:10.2307/421089
16. J. J. Gray, *János Bolyai: Non-Euclidean Geometry and the Nature of Space*, MIT Press, Cambridge, MA, 2004.
17. ———, *Plato's Ghost: The Modernist Transformation of Mathematics*, Princeton University Press, Princeton, NJ, 2008.
18. M. J. Greenberg, On J. Bolyai's parallel construction, *J. Geom.* **12** (1979) 45–64. doi:10.1007/BF01920232
19. ———, Aristotle's axiom in the foundations of hyperbolic geometry, *J. Geom.* **33** (1988) 53–57. doi:10.1007/BF01230603
20. ———, *Euclidean and Non-Euclidean Geometries: Development and History*, 4th revised and enlarged ed., W. H. Freeman, New York, 2007.
21. M. Hallett, Reflections on the purity of method in Hilbert's *Grundlagen der Geometrie*, in *The Philosophy of Mathematical Practice*, P. Mancosu, ed., Oxford University Press, Oxford, 2008, 198–255.
22. M. Hallett and U. Majer, eds., *David Hilbert's Lectures on the Foundations of Geometry*, Springer, New York, 2004.
23. R. Hartshorne, *Geometry: Euclid and Beyond*, Springer, New York, 2000.
24. ———, Teaching geometry according to Euclid, *Notices Amer. Math. Soc.* **47** (2000) 460–465.
25. ———, Non-Euclidean III.36, *Amer. Math. Monthly* **110** (2003) 495–502. doi:10.2307/3647905
26. R. Hartshorne and R. Van Luijk, Non-Euclidean Pythagorean triples, a problem of Euler, and rational points on K3 surfaces, *Math. Intelligencer* **30**(4) (2008) 4–10. doi:10.1007/BF03038088
27. G. Hessenberg and J. Diller, *Grundlagen der Geometrie*, de Gruyter, Berlin, 1967.
28. D. Hilbert, *The Foundations of Geometry* (trans. L. Unger from 10th German ed.) 2nd English ed., revised and enlarged by P. Bernays, Open Court, Peru, IL, 1988.
29. D. Hilbert and P. Bernays, *Grundlagen der Mathematik II*, Zweite Auflage, Die Grundlehren der mathematischen Wissenschaften, Band 50., Springer-Verlag, Berlin, 1970.
30. W. C. Jagy, Squaring circles in the hyperbolic plane, *Math. Intelligencer* **17**(2) (1995) 31–36. doi:10.1007/BF03024895
31. F. Loeser and J. Sebag, Motivic integration on smooth rigid varieties and invariants of degenerations, *Duke Math. J.* **119** (2003) 315–344. doi:10.1215/S0012-7094-03-11924-9
32. E. E. Moise, *Elementary Geometry from an Advanced Standpoint*, 3rd ed., Addison-Wesley, Reading, MA, 1990.
33. J. Nicaise and J. Sebag, Greenberg approximation and the geometry of arc spaces (2009), available at http://arxiv.org/abs/0901.1806.
34. V. Pambuccian, review of *Geometry: Euclid and Beyond* by R. Hartshorne, *Amer. Math. Monthly* **110** (2003) 66–70. doi:10.2307/3072361
35. ———, Axiomatizations of hyperbolic and absolute geometry, in *Non-Euclidean Geometries: János Bolyai Memorial Volume*, A. Prékopa and E. Molnár, eds., Springer, New York, 2005, pp. 119–153.
36. ———, Axiomatizing geometric constructions, *J. Appl. Log.* **6** (2008) 24–46. doi:10.1016/j.jal.2007.02.001
37. W. Pejas, Die Modelle des Hilbertshen Axiomensystems der absoluten Geometrie, *Math. Ann.* **143** (1961) 212–235. doi:10.1007/BF01342979
38. B. Poonen, Computing torsion points on curves, *Experiment. Math.* **10** (2001) 449–466.
39. Proclus, *A Commentary on the First Book of Euclid's Elements* (trans., notes, and intro. by G. R. Morrow), Princeton University Press, Princeton, NJ, 1992.
40. B. A. Rosenfeld, *A History of Non-Euclidean Geometry: Evolution of the Concept of a Geometric Space*, Springer, New York, 1988.
41. W. Schwabhäuser, W. Szmielew, and A. Tarski, *Metamathematische Methoden in der Geometrie*, Springer-Verlag, Berlin, 1983.
42. Stanford Encyclopedia of Philosophy, Hilbert's Program (2003), available at http://plato.stanford.edu/entries/hilbert-program/.
43. J. Strommer, Über die Kreisaxiome, *Period. Math. Hungar.* **4** (1973) 3–16. doi:10.1007/BF02018030
44. A. Tarski, A decision method for elementary algebra and geometry, PROJECT RAND Report R-109 (1951), available at http://www.rand.org/pubs/reports/2008/R109.pdf.
45. F. Tomás, Sobre la consistencia absoluta de la geometrá de la regla y el transportador de segmentos, *An. Inst. Mat. Univ. Nac. Autónoma México* **19** (1979) 41–107.
46. M. Ziegler, Einige unentscheidbare Körpertheorien, *Enseign. Math.* **28** (1982) 269–280.

MARVIN JAY GREENBERG is Emeritus Professor of Mathematics, University of California. In addition to [**20**], he is also the author of *Lectures on Forms in Many Variables* (Benjamin, 1969, to be updated and reprinted), and coauthor with John Harper of *Algebraic Topology: A First Course* (Perseus, 1981). His Princeton Ph.D. thesis, directed by Serge Lang, and his early journal publications are in the subject of algebraic geometry, where he discovered a functor J.-P. Serre named after him (see p. 450 of [**38**] or Section 2.3 of [**31**]) and an approximation theorem J. Nicaise and J. Sebag named after him (see [**33**]). He is the translator of Serre's *Corps Locaux*.
66 Poppy Lane, Berkeley, CA 94708
mjg0@pacbell.net

A Logical Name for a Roman Bookstore

—Submitted by Jonathan Sondow, New York, NY