1 The Question of Parallels

We sketch a very brief history of certain aspects of geometry as background for the material in this book. Euclid is the central figure but is neither the beginning nor the end of our story.

"Geometry" is originally a Greek word that means "measurement of the earth." The abstract concepts of shape, length, area, and their relationships have been known, at least on a practical level, from the time when human beings first began to construct buildings, mark the boundaries of agricultural fields, and chart the location and movement of the moon and stars. Some nontrivial geometric results were widely distributed as early as the Neolithic period (about 4000 B.C. to 2000 B.C.). For example, the theorem of Pythagoras (that the square of the hypotenuse of a right triangle equals the sum of the squares of the other two sides) appears to have existed then in Babylonia, the British Isles, Greece, India, China, and Egypt. The evidence for this includes the inscription of tables of Pythagorean triples (integers that are the dimensions of right triangles, such as (3, 4, 5) or (5, 12, 13)) on Babylonian baked clay tablets from the time of Hammurabi (around 1800 B.C.); configurations of right triangles of dimensions 12, 35, 37 (measured in half-integer multiples of the megalithic yard of 0.829 m) that occur in rings of standing stones near Inverness in Scotland and at Woodhenge in southern England; and an explicit statement of the theorem in the Baudhayana Sulvasutra, an ancient Hindu manual that gives precise details for the construction of altars of specified form and size.

The desire to understand why geometric facts are true, if not to prove them rigorously, also seems to be of ancient origin. A diagram and brief explanation presents an informal proof of the theorem of

Pythagoras in the earliest known Chinese text on astronomy and mathematics, the *Chou Pei Suan Ching* ("Arithmetical Classic of the Gnomon and the Circular Paths of Heaven"; "gnomon" means both the raised part of a sundial and a figure that equates a rectangle with a difference of squares). It is done explicitly for a triangle of dimensions 3, 4, 5 but is in fact quite general. The text was written during the Han Dynasty, around 200 B.C., but is believed to reproduce knowledge that is much older. The diagram and a hint for the proof are given in Problem 1 at the end of this chapter.

The results found on the Babylonian clay tablets, or on ancient Egyptian papyrii, are examples of primary historical evidence. That is, the artifacts were created at or near the time when (historians believe) the information written on them was discovered. There have been no such primary sources found for the more recent and more extensive contributions to geometry made by Greek civilization. The physical fragility of the Greek manuscripts and their exposure to various political and social upheavals probably accounts for this. But the study of copies and translations made hundreds of years after the fact yields a fairly clear picture of the nature and magnitude of the geometry developed in ancient Greece.

Greek mathematicians assimilated much of the geometric lore of Babylonia and Egypt. Then, from around 600 B.C., they began to apply to the subject the methods of logic developed by the Greek philosophers. In doing so they transformed geometry from a collection of empirical facts and isolated explanations into a unified discipline in which all results are proved by rigorous arguments from a small set of initial assumptions. Thales of Miletus, who lived during the sixth century B.C., is credited as being the first to use logical deduction as a standard procedure for doing geometry. About 50 years later, Pythagoras and the mystical religious brotherhood that he founded in southern Italy extended the method to connect long chains of results. The theorem that is his namesake preceded him by centuries, but Pythagoras and his school contributed many original and fundamental results. They used properties of parallel lines to investigate proportionality in similar figures and to prove that the angle sum of any triangle is equal in measure to a straight angle. One of the Pythagoreans, Hippocrates of Chios, is believed to have made the first attempt at a cohesive logical treatise on the entire field. Several other such works appeared over the next two hundred years. These efforts culminated in Euclid's *Elements*, written around 300 B.C.

Euclid served as the first Professor of Mathematics at the University of Alexandria, in Egypt at the Mediterranean Sea. The conquests of Alexander the Great made Egypt part of the Greek empire. The school at Alexandria became a major center of learning and remained so for centuries. From there, the *Elements* was disseminated through much of the world. Because of its clarity and cohesiveness, it soon became the standard text for geometry and replaced all of the works of Euclid's predecessors. Apart from some references to them in later sources, these earlier efforts no longer exist.

The initial assumptions (postulates, axioms) and definitions on which Euclid based his development are recalled in Appendix I. (Euclid's statements, both in this chapter and in Appendix I, are quoted from T. L. Heath, *The Thirteen Books of Euclid's Elements*, Vol. I, Dover Publications, New York, 1956.) Euclid's use of further tacit assumptions is discussed briefly in Chapter 13. No one saw a need to make these unspoken suppositions explicit until the late nineteenth century, when the discovery of hyperbolic geometry generated new questions about the nature of geometric truth and its verification. But the mathematical world was focused for more than 2000 years on decreasing rather than increasing the number of Euclid's assumptions. The goal was to eliminate the need for the Fifth Postulate (see Figure 1.1):

> That, if a straight line falling on two straight lines makes the interior angles on the same side less than two right angles, the two straight lines, if produced indefinitely, meet on that side on which are the angles less than the two right angles.

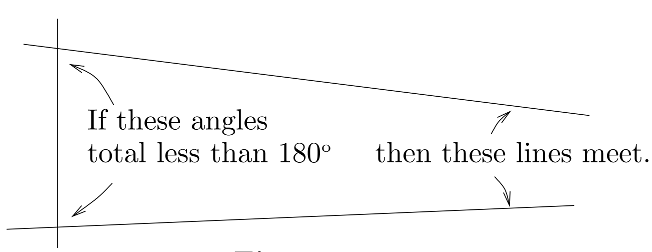

Figure 1.1

This statement is about parallel (that is, nonintersecting) lines in the plane. It is equivalent, in the context of Euclid's other assumptions, to Playfair's Postulate (after the Scottish physicist and mathematician John Playfair [1748–1819]):

Through a given point not lying on a given line there can be drawn only one line parallel to the given line.

A discussion of the equivalence of these two statements, in the more general setting of our text, is given in Chapter 17. (See Theorem 17.6; see also Problems 3 and 4 at the end of this chapter.)

The properties of parallel lines and their role in plane geometry were big issues for many years before Euclid wrote the *Elements*. There is some evidence that in work from the fourth century B.C. but now lost, the philosopher Aristotle articulated as a basic assumption a principle very similar to the Fifth Postulate. In his treatise *Prior Analytics*, which still exists, he cites the implicit supposition of properties of parallel lines as an example of the logical error of begging the question.

The Fifth Postulate is both more complicated and less obvious than Euclid's other assumptions. So it was inevitable that the urge to modify it, either by proving it as a consequence of the other postulates or by replacing it with a more self-evident statement, would be reborn many times over two millennia. Only a few decades after Euclid, the great Greek mathematician Archimedes wrote a treatise (now lost) called *On Parallel Lines*. The astronomer Ptolemy of Alexandria attempted a proof of the Fifth Postulate in the second century A.D. A number of Byzantine scholars tried to construct a proof, in particular Simplicius in the sixth century. There is reason to believe that Archimedes may have redefined parallel lines as lines that are everywhere equidistant. If so, then he tacitly made an assumption that is equivalent to the Fifth Postulate. Simplicius made another such tacit assumption when he based his proof by contradiction on the "fact" that for any point in the interior of an angle, there exists a line through the point that meets both sides of the angle. This again is a property of the Euclidean plane that is equivalent to the Fifth Postulate. (See Problem 7.) Ptolemy's proof foundered on simply a gross misuse of language and logic.

All of the numerous other attempts at proving the Fifth Postulate inevitably failed. This is because there really exists another type of plane (now called hyperbolic) that is as valid as the standard Euclidean plane, and in which Euclid's other postulates hold but the fifth one fails. Both a specific model for the hyperbolic plane and its general properties are developed throughout our text, beginning in Chapter 2. It took until well into the nineteenth century before the hyperbolic plane was discovered and its existence definitively proved. We mention here only a few of the many who tried to prove the Fifth Postulate. Most of them committed an error of a type similar to one of those given in the preceding paragraph, even though the various geometers usually were rather perceptive in identifying the mistakes of their predecessors. Some were explicit in noting that they assumed an additional premise.

Mathematical leadership resided in the Arabic-speaking world from the sixth through the thirteenth centuries. Notable among those from Arabic cultures who tackled the Fifth Postulate were the Baghdad mathematician and astronomer Thabit Ibn Qurra in the ninth century and the Egyptian physicist, mathematician and astronomer Abu Ali Ibn al-Haytham in the tenth century. Both Ibn Qurra and Ibn al-Haytham introduced a quadrilateral with three right angles and presented flawed proofs that such a figure must be a rectangle. This quadrilateral was later considered in the eighteenth century by Lambert and is discussed later.

The celebrated Persian mathematician, astronomer, philosopher and poet Omar Khayyam also tried to establish the Fifth Postulate. Khayyam explicitly assumed Aristotle's principle. He also introduced the quadrilateral studied more extensively 600 years later by Saccheri and also discussed later (see as well the problems for Chapter 19).

A diverse set of mathematicians in medieval Europe attempted to prove Euclid's parallel postulate. Among them were the Polish scholar Vitello in the thirteenth century, and the Jewish religious philosopher Levi ben Gerson, who lived in southern France in the fourteenth century. John Wallis, in seventeenth-century England, adopted the postulate (equivalent to the Fifth) that there exist similar plane figures of arbitrary size. He regarded this as a natural generalization of Euclid's Third Postulate, that there are circles of arbitrary radius.

The eighteenth century Italian Jesuit priest Girolamo Saccheri was not happy with Wallis's proof. Saccheri considered quadrilaterals with two sides of equal length both perpendicular to the base, as did Khayyam. He attempted to show, as did Khayyam and others, that the two (necessarily equal) angles at the top of the quadrilateral could not be acute. He assumed the angles were acute and tried to argue toward a contradiction. This acute angle hypothesis is exactly what is true in the hyperbolic plane, but of course Saccheri was not aware of this. He was careful not to make any further assumptions, and he derived many consequences of his basic hypothesis. Thus he unknowingly developed many properties of the hyperbolic plane. In particular, he showed that any two nonintersecting straight lines either have a common perpendicular or come asymptotically close to each other. His efforts to derive a contradiction from this or consequent properties led either to an error in calculating or to the apparently frustrated declaration that the "hypothesis of acute angle is absolutely false; because (it is) repugnant to the nature of the straight line." Saccheri clearly was aware that he had not produced a satisfactory argument.

Johann Heinrich Lambert, working in Munich and Berlin in the latter half of the eighteenth century, studied the same quadrilateral with three right angles as did Ibn Qurra and Ibn al-Haytham. Lambert assumed toward a contradiction that the fourth angle was acute. Ibn al-Haytham had dismissed this situation by the same sort of redefinition of parallel lines as being everywhere equidistant that was attributed to Archimedes. But Lambert argued more carefully and derived even more properties of the hyperbolic plane than had Saccheri. In particular, he showed that the angle sum of a triangle is less than two right angles (that is, less than 180 in degree measure) and noted the resulting connection between angle sum and area. He also observed the analogy between the consequences he found of the acute angle hypothesis and the results of spherical geometry, where the fourth angle of a quadrilateral with three right angles must be obtuse. He even had the glimmer of an idea that perhaps some geometric surface, an imaginary sphere, exists where the Fifth Postulate fails and Euclid's other postulates hold. Nevertheless, he felt obliged to finish the quest for a contradiction. His argument stumbled over an unwarranted assumption about the set of points all the same distance from a given line, an assumption that resembles the one made by Ibn al-Haytham.

The Italian-French mathematician Adrien-Marie Legendre made an extended series of attempts at proving the Fifth Postulate in the 12 editions of his textbook, *Elements de Geometrie*, which appeared between 1794 and 1823. He was one of many who replicated the error made by Simplicius some 1300 years before.

The truth about the nature of the Fifth Postulate finally was grasped in the nineteenth century, by three men working independently: Nicolai Ivanovich Lobachevsky (1793–1856) of Russia, Janos Bolyai (1802–1860) of Hungary, and the great mathematician Karl Friedrich Gauss (1777–1855) of Germany. In 1829, Lobachevsky published a paper entitled "On the Principles of Geometry" where he asserted that there is a consistent system of plane geometry in which there exist multiple lines through a point that do not meet a given line and angle sums of triangles are less than the measure of a straight angle. He did not have a proof of consistency. But he did derive the trigonometric formulas that hold in his alternative plane and he showed that these formulas are internally consistent. He observed that they can be obtained from the analogous formulas of spherical trigonometry by multiplying the sides of a triangle by an imaginary unit. Thus, he termed his plane "imaginary." He also used its properties to compute definite integrals, some known and some previously unknown. His effort was received by his Russian colleagues with disbelief and ridicule. But Lobachevsky remained convinced of the value of his discovery and published further articles about it over the next 25 years.

Bolyai's remarkably similar results constituted an article called "Supplement Containing the Absolutely True Science of Space, Independent of the Truth or Falsity of Euclid's [Fifth Postulate] (That Can Never Be Decided A Priori)." It appeared in 1832 as an appendix to a book by his father, Farkas, which surveyed previous attempts to prove the parallel postulate. J. Bolyai termed "absolute" a plane that could be either Euclidean or the alternative one with respect to properties of parallel lines.

Gauss made many attempts to prove the Fifth Postulate early in his mathematical life and then became convinced that the alternative plane must exist. But he never published any of this work, most likely because he felt it would have a cool reception. He was delighted to read J. Bolyai's and Lobachevsky's articles, and he noted in private correspondence that they had replicated some of his own results. In a letter to Farkas Bolyai he wrote of Janos's paper,

> Indeed the whole contents of the work, the path taken by your son, the results to which he is led, coincide almost entirely with my own meditations, which have occupied my mind partly for the last thirty or thirty-five years. So I remained quite stupefied. So far as my own work is concerned, of which up till now I have put little on paper, my intention was not to let it be published during my lifetime. Indeed the majority of people have not clear ideas upon the questions of which we are speaking, and I have found very few people who could regard with any special interest what I communicated to them on this subject....I am very glad that it is just the son of my old friend who takes the precedence of me in such a remarkable manner. [R. Bonola, *Non-Euclidean Geometry*, Dover Publications, New York, 1955, p. 100]

Gauss was also interested in Lobachevsky's work and promoted it through private letters to other European mathematicians. The publication of one of these letters, a few years after Gauss's death, made an especially strong impression and generated a surge of interest in the new geometry. Several important papers about it were written in the late 1860s. In particular, the Italian Eugenio Beltrami published two articles in 1868 that contain the first concrete models of this alternative plane. One of them is the example studied in this book (see Example 4 of Chapter 2, and Chapter 11) and is known as the Beltrami-Klein model. That is, Beltrami found a subset of the Euclidean plane on which he defined lines, angles, distance, and angle measure in such a way that the properties of the alternative plane hold. So he proved that the "imaginary" plane of Lobachevsky is as consistent as the Euclidean plane.

Beltrami's results firmly established the existence of this non-Euclidean plane and finally settled the issue of the Fifth Postulate. It turns out that Beltrami's model was implicit in the more general geometric structures introduced in 1859 by the Englishman Arthur Cayley. But Cayley did not see the link between his work and the new planes of Lobachevsky and Bolyai. It remained for another eminent German mathematician, Felix Klein, to make this connection in 1871. Klein also coined the term “hyperbolic” for the new geometry. Other models for the hyperbolic plane were discovered by the French mathematician Henri Poincaré in the 1880s. A great deal of work has been done on this plane and its higher-dimensional analogs up to the present day, and many applications have been found throughout broad areas of mathematics and mathematical physics.

Why did over 2100 years elapse before the true status of the Fifth Postulate and the existence of the hyperbolic plane was revealed? Certainly the nature of the Euclidean plane as an abstraction from everyday human experience lends it an aura of uniqueness. It seemed imperative to almost everyone who considered the question that no other planar structure is possible. It is not surprising that some of the would-be provers of the Fifth Postulate resorted in the end to an appeal to real-world objects. Alexis Claude Clairaut in the eighteenth century justified the existence of a rectangle (as opposed to a Saccheri quadrilateral with acute angles) by citing the “form of houses, gardens, rooms, walls.” The various mathematicians described previously did not suffer such a blatant lapse of methodology, but they all worked under an imposing psychological burden. In fact, the eighteenth-century German philosopher Immanuel Kant believed that Euclid’s assumptions existed as part of the structure of the human mind and that no correct reasoning about geometry was even possible without them. Lobachevsky, Bolyai, Gauss, and Beltrami proved him wrong.

Why didn’t the existence of spherical geometry provide a clue that other surfaces should be investigated as possible sources for an alternative plane? The study of the sphere, because of its applications to astronomy and the calendar, existed in ancient Babylon and Egypt. The earliest mathematical treatise that has come down to us is “On the Rotating Sphere” by Autolycus in the fourth century B.C. The geometry of the spherical surface was developed to a level comparable to that of the Euclidean plane by Menelaus of Alexandria in his work, “On the Sphere,” written in the first century A.D. The results of our text that pertain to the sphere are all essentially to be found in Menelaus’s book. Later contributions, especially to spherical trigonometry, were made by many others, until the great Swiss mathematician Leonhard Euler put the subject in its modern form in the eighteenth century. So the sphere was not unknown to those pursuing the Fifth Postulate. But lines on the sphere, the "great circles" that divide a sphere into two equal-sized hemispheres (see Chapter 2), are finite in extent, whereas lines in the plane, by a (tacit) assumption of Euclid, are infinite. So the sphere seemed irrelevant to the issue of parallel lines in the plane. And of course, there are no parallel lines on the sphere.

Consideration of standard problems of Euclidean geometry was never likely to lead to the discovery of a model of the hyperbolic plane. In every model, either distance or angle measure must be modified from Euclidean distance or measure by some strange sort of twist, as we will begin to see in Chapter 2. It would take the development of an extensive theory of geometric surfaces, a knowledge of this theory as possessed by Beltrami and Klein, and the motivation prompted by the message of Lobachevsky, Gauss, and Bolyai in order to find such a model.

So the hyperbolic plane took a long time to arrive. Once it did, the scientific world's perception of geometry, and indeed of all of mathematics, was fundamentally changed. No longer could mathematics be regarded as simply the investigation of unique structures that had a prior existence as ideal objects in some philosophical sense. Now, alternative assumptions could be made from which different yet equally valid consequences could be derived. Thus mathematics became a more arbitrary and at the same time a more powerful enterprise. The broader perspective has led to deeper insights.

The discovery of the hyperbolic plane generated new concerns about the foundations of Euclidean geometry. The intuition and visualization that for over 2000 years had quietly resided in Euclid's postulates and arguments became suspect. New axiomatic treatments were written, with the goal of bringing Euclid's tacit assumptions to the surface and making all strands of the arguments explicit. The first of these foundational works was completed in 1882 by the German mathematician Moritz Pasch. One of his contributions was to clarify the distinction between undefined and defined terms. He recognized that he must present some terms as undefined, or primitive, as otherwise either an infinite or a circular sequence of definitions must follow. He saw that Euclid's definition of a point as "that which has no part" (see Appendix I) was not really a definition in any rigorous sense, as "part" has no meaning. Presumably, Euclid's purpose here was to provide some intuition; he never used this definition, or several others, in any of his arguments. Pasch let the word "point" be an undefined term, and his postulates stated all that he allowed to be assumed about it. Our text also leaves "point" as undefined. However, the notion of "betweenness" among points was also an undefined notion for Pasch, whereas we define it in terms of distance (and collinearity) based on an assumed preexistence of the real number system (see Chapters 5 and 6).

The most influential axiomatic study of Euclidean geometry was published in 1899 by the famous German mathematician David Hilbert. His treatment presented 21 axioms or postulates (and for us these terms mean the same thing; namely, assumptions) and showed that the full theory of Euclidean plane and solid geometry, including the development of the real numbers, follows. The American mathematician George David Birkhoff, in an article he published in 1932, assumed the prior existence of the real numbers and formulated just four postulates from which he showed that the theory of the Euclidean plane can be derived. Two of these postulates are essentially the Ruler and Protractor Axioms that we discuss in Chapter 13. A number of authors have considered the simultaneous development of the Euclidean and hyperbolic planes. The first axiomatic setting of which we are aware for the simultaneous study of the Euclidean, hyperbolic, and spherical planes occurs in the 1969 book *College Geometry* by David C. Kay (Holt, Rinehart and Winston, New York). Much but by no means all of the axiom system developed in our text borrows from Kay's system.

Problem Set 1

The solutions to these problems will have to be informal arguments based on informal recollections of concepts and results from Euclidean geometry. These should include the area of a rectangle and a (right) triangle, congruence of triangles, the interior of an angle, the angle sum of (the three angles of) a triangle, and the property (tacit in Euclid) that any line in a plane separates the plane into disjoint halfplanes. But note that the Fifth Postulate (or Playfair's Postulate) is not to be used except where indicated. Most of these concepts will be introduced and carefully developed in the chapters to follow, where rigorous proofs then will be feasible and expected.

1. Figure 1.2 appears in the ancient Chinese text *Chou Pei Suan Ching* with $a = 3$ and $b = 4$ and presents an implicit proof of the Pythagorean Theorem. Make this proof explicit, for general a and b, through the following steps:

   (a) Note that each of the four $a \times b$ rectangles in the picture is split into two congruent right triangles by a diagonal whose length is denoted d.

   (b) Show that the quadrilateral $PQRS$ is a square (that is, each vertex angle is a right angle). You may assume that the three angles of a triangle add to $180°$.

   (c) Use the decomposition of $PQRS$ into four triangles and a square to find the area of $PQRS$ in terms of a and b.

   (d) Conclude that $d^2 = a^2 + b^2$.

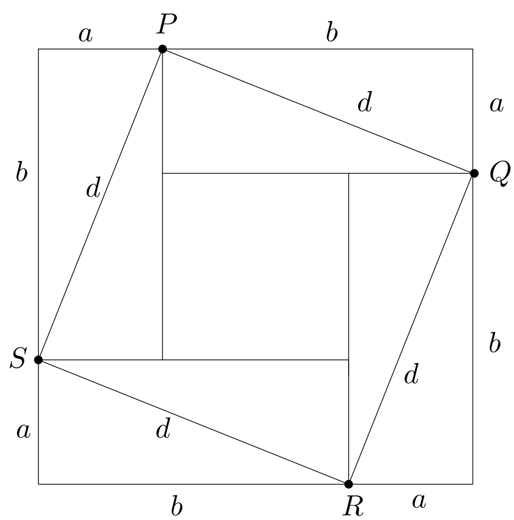

Figure 1.2

2. Euclid proved the Exterior Angle Inequality, which says that an exterior angle of a triangle is larger than either remote interior (that is, nonadjacent) angle, without using the Fifth Postulate. (See Chapters 2 and 15.) Use the Exterior Angle Inequality to

show that if line l crosses lines m and n so that the interior angles on one side add to two right angles (see Figure 1.3), then m and n are parallel. (Hint: Suppose that m and n meet and find a contradiction. Do *not* assume that the three angles of a triangle add to 180.)

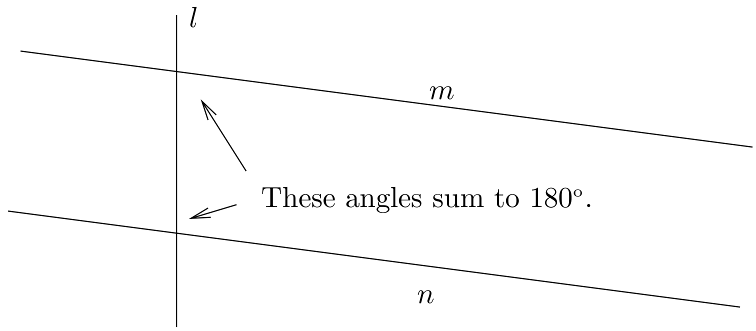

Figure 1.3

3. Show that Playfair's Postulate implies the Fifth Postulate. (Hint: Use Problem 2.)

4. Show that the Fifth Postulate implies Playfair's Postulate.

5. Show that if the Fifth Postulate holds, then the angle sum of any triangle equals two right angles (180°). (Hint: Consider the line through a vertex of the triangle that is parallel to the opposite side, as in Figure 1.4. This pair of parallel lines is crossed by each of the other two sidelines of the triangle. What can you say about the interior angles in these configurations?)

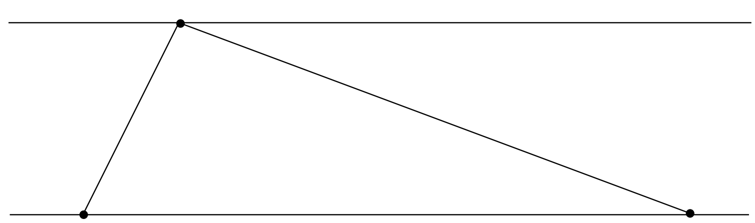

Figure 1.4

6. Assume a plane where Playfair's Postulate is false. That is, there are lines $\overleftrightarrow{QP}$ and $\overleftrightarrow{RP}$ both parallel to line l, as in Figure

1.5. Can you find more lines through P that are parallel to l? Explain.

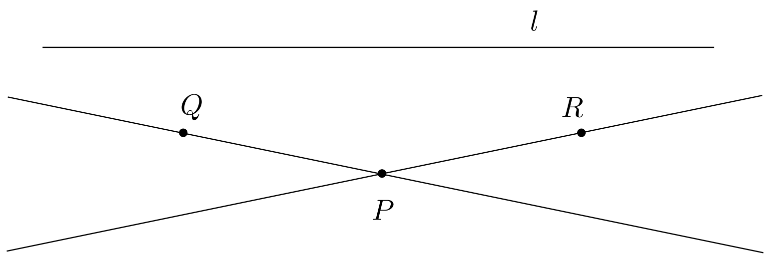

Figure 1.5

7. Consider the following claim (as in Figure 1.6):

   (*) If X is any point in the interior of a proper angle $\underline{\angle QPR}$, then there is a line through X that meets both sides of the angle.

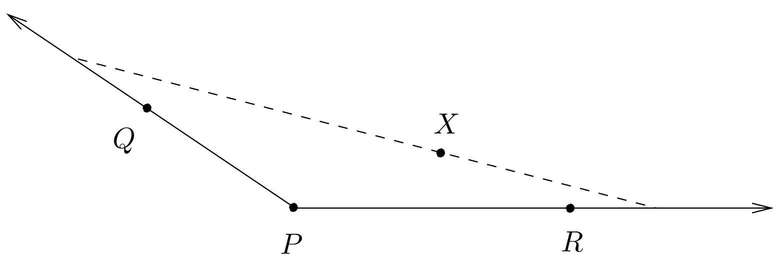

Figure 1.6

   (a) Assume a plane where Playfair's Postulate is true. Show that (*) must also be true.

   (b) Assume a plane where Playfair's Postulate is false. That is, there are lines $\overleftrightarrow{QP}$ and $\overleftrightarrow{RP}$ both parallel to line l, as in Figure 1.5. You may assume that l is contained in the interior of angle $\underline{\angle QPR}$. Explain why (*) must be false for $\underline{\angle QPR}$ and any point X on l. (That is, no line through X can meet both sides of $\underline{\angle QPR}$.)

8. Suppose a plane where there is a (Saccheri) quadrilateral $ABCD$ with angles $\underline{\angle C}$ and $\underline{\angle D}$ right angles and $\underline{\angle A}$ and $\underline{\angle B}$ acute (as in Figure 1.7).

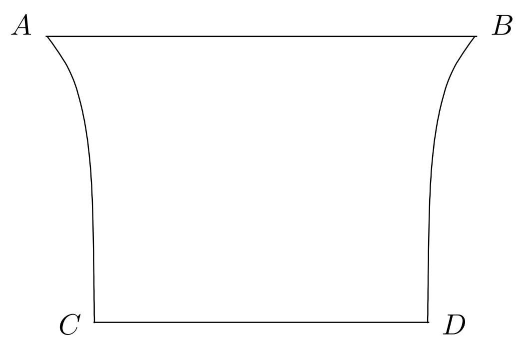

Figure 1.7

Show that the Fifth Postulate must then be false. (Hint: What does Problem 2 say about lines $\overleftrightarrow{AC}$ and $\overleftrightarrow{BD}$?)

9. Explain how a Saccheri quadrilateral can be obtained from a Lambert quadrilateral (with three right angles and an acute angle) by taking a mirror image across one side.

10. Assume that the Fifth Postulate is false; that is, angles $\underline{\angle SPQ}$ and $\underline{\angle PQO}$ sum to less than 180 in Figure 1.8, but lines $\overleftrightarrow{PS}$ and $\overleftrightarrow{QO}$ are parallel. Show that there exists a triangle with angle sum less than 180. (You may assume that points R exist on the ray $\overrightarrow{QO}$ with angle $\underline{\angle PRQ}$ arbitrarily small; see Propositions 19.2 and 19.3.)

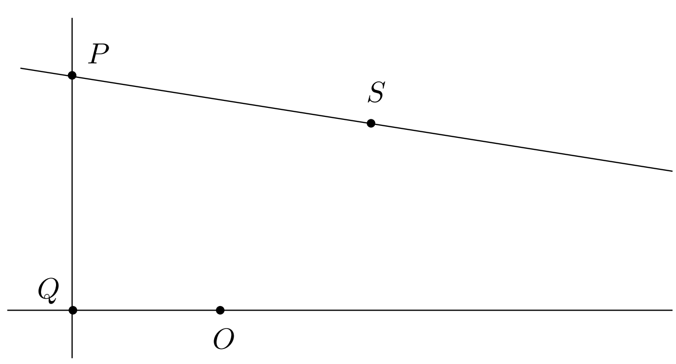

Figure 1.8

2 Five Examples

An important idea in this text is that there is more than one kind of geometry. The axiom system we will build describes not only the ordinary Euclidean plane but other interesting geometric structures as well. So the theorems we prove in the context of our axiom system yield information about a number of different examples all at once. The following five examples satisfy most of the axioms to be introduced, and three of them will satisfy all twenty-one of the axioms. It will be helpful to keep the examples in mind as we develop the general theory.

(1) $\mathbb{E}$: The usual *Euclidean plane*, points, and lines.

Coordinates: The points in $\mathbb{E}$ are in one-to-one correspondence with the ordered pairs of real numbers: Each point A corresponds to a pair of real numbers (x, y), called the *coordinates* of A, where the pair is assigned in the familiar way (Figure 2.1).

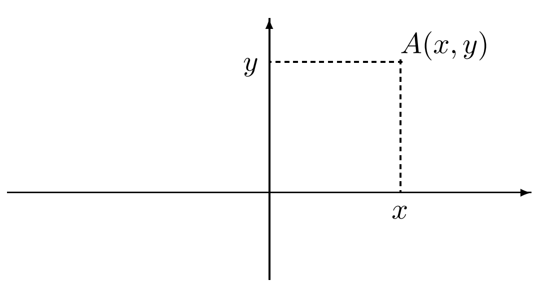

Figure 2.1

We often identify A with its pair of coordinates (x, y).

Equations of Lines. Each *nonvertical line* l in $\mathbb{E}$ consists of all points (x, y), where $y = mx + b$ for some fixed m and b. Each *vertical line* l consists of all (x, y), where $x = a$ for some fixed a.

For any two points $A(x_1, y_1)$ and $B(x_2, y_2)$, the *slope* of the line l through A and B is

$$m = \frac{y_2 - y_1}{x_2 - x_1} \qquad (\text{if } x_2 \neq x_1),$$

and an equation for l is given by

$$y - y_1 = m(x - x_1) \qquad (\text{if } x_2 \neq x_1).$$

The *Euclidean distance* $e(AB)$ between A and B satisfies the formula

$$e(AB) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

Example. Let $A(1, 2)$, $B(-1, -3)$ be given. The line through A and B has slope $m = \frac{-3-2}{-1-1} = \frac{-5}{-2} = \frac{5}{2}$ and equation $y - 2 = \frac{5}{2}(x - 1)$, or $y = \frac{5}{2}x - \frac{1}{2}$. Distance $e(AB) = \sqrt{(-3-2)^2 + (-1-1)^2} = \sqrt{25 + 4} = \sqrt{29}$.

Proposition 2.1. If $A(x_1, y_1)$ and $B(x_2, y_2)$ are on the line $y = mx + b$, then $e(AB) = |x_1 - x_2|\sqrt{m^2 + 1}$.

Proof. Problem 3.

(2) $\mathbb{M}$: The *Minkowski plane*, or *taxicab plane*.

$\mathbb{M}$ has the same points, lines, and coordinates as does $\mathbb{E}$, but *distance is different*: For any $A(x_1, y_1)$ and $B(x_2, y_2)$, define

$$d_{\mathbb{M}}(AB) = |x_2 - x_1| + |y_2 - y_1|.$$

So the Minkowski distance $d_{\mathbb{M}}(AB)$ is defined as the sum of the horizontal and vertical "ordinary distances" (Figure 2.2).

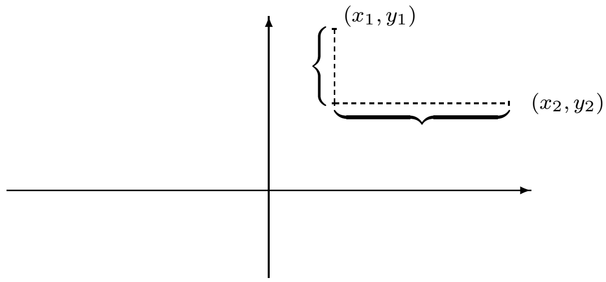

Figure 2.2

Example. Given $A(1,2)$, $B(-1,-3)$, then

$$d_{\mathbb{M}}(AB) = |-1-1| + |-3-2| = 2 + 5 = 7.$$

Proposition 2.2. If $A(x_1, y_1)$ and $B(x_2, y_2)$ are on the line $y = mx + b$, then $d_{\mathbb{M}}(AB) = |x_1 - x_2|(1 + |m|)$.

Proof. Problem 4.

(3) $\mathbb{S}(r)$: The (surface of the) *sphere* of radius r; that is, the *spherical plane.*

Once r is fixed, we shorten the notation to $\mathbb{S}$. We shall assume that our spheres are centered at the origin $(0,0,0)$ in three-dimensional space. Then $\mathbb{S}$ is the set of all (x, y, z) with $x^2 + y^2 + z^2 = r^2$. Points are as usual, and lines on $\mathbb{S}$ are defined to be the *great circles.* (A great circle is the intersection of the sphere with a plane that cuts the sphere in half; the equator is a great circle on the earth's surface.) Then any two points have a unique line joining them, *unless* they are opposite (antipodes), when they have infinitely many lines joining them (as infinitely many longitudinal great circles join the north and south poles).

Distance in $\mathbb{S}$: For points A, B on $\mathbb{S}$, define distance

$d_{\mathbb{S}}(AB)$ = length of the minor (i.e., shorter) arc of the great circle (line) through A and B.

To compute $d_{\mathbb{S}}(AB)$ more easily, we first recall a formula for *arc length in a circle of radius* r (see Figure 2.3).

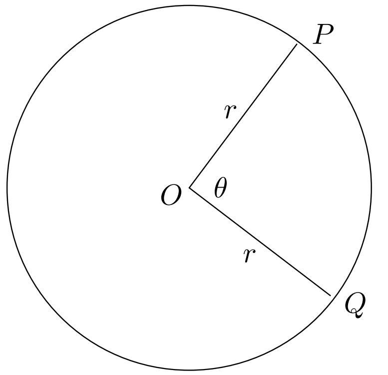

Figure 2.3

Let θ be the radian measure of $\underline{\angle POQ}$. The angle that sweeps out the full circle has measure 2π, and the circumference is $2\pi r$. The sector formed by $\underline{\angle POQ}$ makes up $\frac{\theta}{2\pi}$ of the full circle, so

$$\text{arc length } PQ = \frac{\theta}{2\pi} \cdot 2\pi r = \theta r.$$

For instance, $\theta = \frac{\pi}{3}$ implies that arc length equals $\frac{\pi}{3} r$.

Example. In $\mathbb{S}$ of radius 1, let $A = (1, 0, 0)$, $B = (\frac{\sqrt{3}}{2}, \frac{1}{2}, 0)$. We find $d_{\mathbb{S}}(AB)$: A, B lie on a great circle in the xy-plane (Figure 2.4).

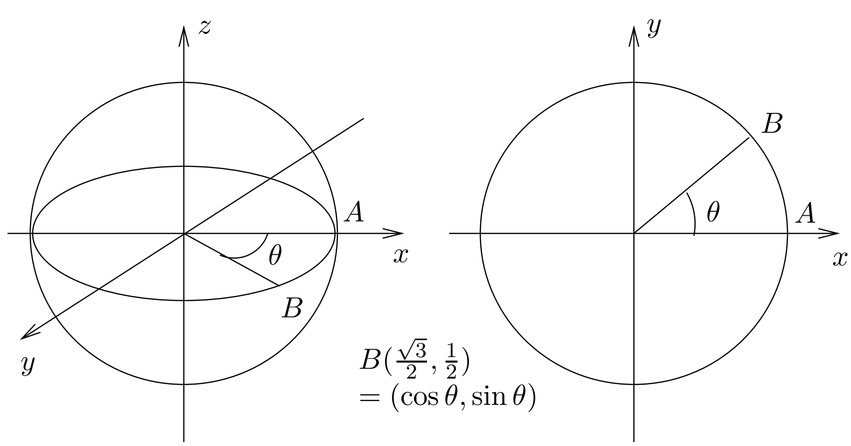

Figure 2.4

$$\frac{\sqrt{3}}{2} = \cos\frac{\pi}{6} \text{ and } \frac{1}{2} = \sin\frac{\pi}{6}, \text{ so } \theta = \frac{\pi}{6}. \text{ So } d_{\mathbb{S}}(AB) = \theta r = \frac{\pi}{6}.$$

An explicit formula for the spherical distance between two points, in terms of their coordinates, is given next. It follows from the distance formula for three-dimensional space and the Law of Cosines and is derived in Problem 10. Here, "$\cos^{-1}$" means the usual inverse cosine or arc cosine function, measured in radians.

If $P(a, b, c)$ and $Q(x, y, z)$ are points on the surface of the sphere of radius r centered at $(0, 0, 0)$ then

$$d_{\mathbb{S}}(PQ) = r\cos^{-1}\left(\frac{ax + by + cz}{r^2}\right).$$

Example. Let $A = (1, 0, 0)$ and $B = \left(\frac{\sqrt{3}}{2}, \frac{1}{2}, 0\right)$ on the sphere of radius 1, as in the previous example. Then by the formula,

$$d_{\mathbb{S}}(AB) = 1 \cdot \cos^{-1}\left(1 \cdot \frac{\sqrt{3}}{2} + 0 \cdot \frac{1}{2} + 0 \cdot 0\right) = \cos^{-1}\frac{\sqrt{3}}{2} = \frac{\pi}{6},$$

which agrees with the previous calculation.

Example. In $\mathbb{S}$ of radius 2, $P = \left(\frac{2}{3}, \frac{4}{3}, \frac{-4}{3}\right)$ and $Q = \left(\frac{6}{5}, 0, \frac{8}{5}\right)$. Then

$$\begin{aligned} d_{\mathbb{S}}(PQ) &= 2 \cdot \cos^{-1}\left(\frac{\frac{2}{3} \cdot \frac{6}{5} + \frac{4}{3} \cdot 0 - \frac{4}{3} \cdot \frac{8}{5}}{2^2}\right) \\ &= 2 \cdot \cos^{-1}\left(-\frac{1}{3}\right) = 3.82\ldots\,. \end{aligned}$$

(4) $\mathbb{H}$: The *hyperbolic plane.*

$\mathbb{H}$ consists of all points *inside* (but not on) the unit circle in $\mathbb{E}$, that is, all (x, y) with $x^2 + y^2 < 1$.

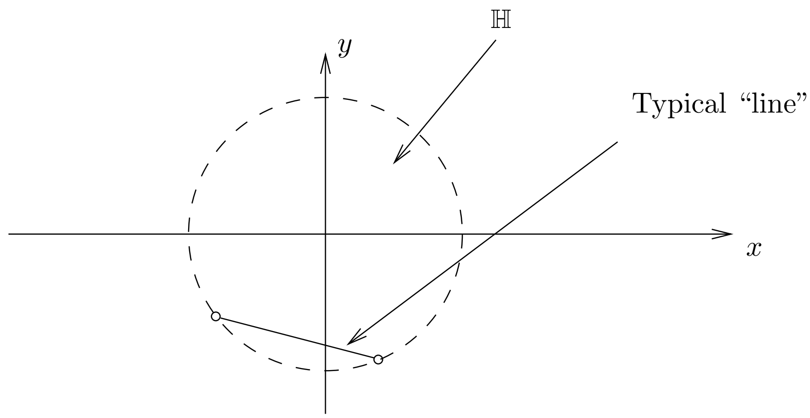

Figure 2.5

Lines in $\mathbb{H}$ are defined to be the chords of the circle (as in Figure 2.5). If A, B are two points in $\mathbb{H}$, define $d_{\mathbb{H}}(AB)$, the *distance* between them in $\mathbb{H}$, as follows: Draw the chord AB, and let M, N be the points where the chord meets the unit circle (M, N are in $\mathbb{E}$ but not in $\mathbb{H}$). Label so that B separates A and N (Figure 2.6).

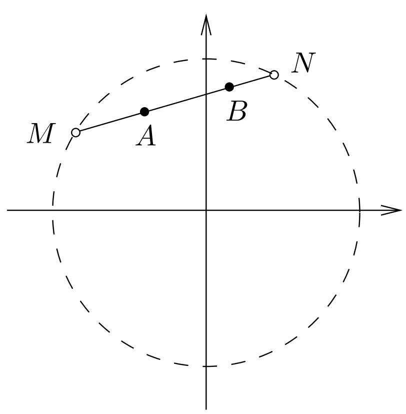

Figure 2.6

Let $e(PQ)$ denote the usual Euclidean distance between points, and *define*

$$d_{\mathbb{H}}(AB) = \ln\left(\frac{e(AN)e(BM)}{e(AM)e(BN)}\right),$$

where "ln" denotes the natural logarithm function. Since $e(AN) > e(BN)$ and $e(BM) > e(AM)$, we have $\frac{e(AN)}{e(BN)} > 1$ and $\frac{e(BM)}{e(AM)} > 1$.

Hence $\frac{e(AN)e(BM)}{e(AM)e(BN)} = \frac{e(AN)}{e(BN)}\frac{e(BM)}{e(AM)} > 1$. It follows from a property of ln that $d_{\mathbb{H}}(AB) > 0$. Note that $d_{\mathbb{H}}(BA) = \ln \frac{e(BM)e(AN)}{e(BN)e(AM)} = d_{\mathbb{H}}(AB)$. Also,

$$d_{\mathbb{H}}(AB) = \left|\ln\left(\frac{e(AN)e(BM)}{e(AM)e(BN)}\right)\right| = \left|\ln\left(\frac{e(AM)e(BN)}{e(AN)e(BM)}\right)\right|.$$

So if absolute value is used in this way, then we need not worry about which point on the unit circle is marked M and which is marked N.

If $A = B$ in $\mathbb{H}$, take any chord through A and let M, N be as previously. Since $\frac{e(AN)e(AM)}{e(AM)e(AN)} = 1$, it is consistent with the preceding definition to set $d_{\mathbb{H}}(AA) = 0$.

Examples. If $A = (.8, 0)$, $B = (.9, 0)$, then $M = (-1, 0)$, $N = (1, 0)$ and

$$d_{\mathbb{H}}(AB) = \ln\left(\frac{(.2)(1.9)}{(1.8)(.1)}\right) = .7472144\ldots\ .$$

If $A = (0, 0)$, $B = (.999, 0)$, then $M = (-1, 0)$, $N = (1, 0)$ and

$$d_{\mathbb{H}}(AB) = \ln\left(\frac{(1)(1.999)}{(1)(.001)}\right) = 7.6004023\ldots\ .$$

If $A = (-\frac{1}{2}, 0)$, $B = (0, \frac{1}{2})$, then the line joining them is $y = x + \frac{1}{2}$, which meets $x^2 + y^2 = 1$ at

$$M = (-\frac{1}{4} - \frac{1}{4}\sqrt{7}, \frac{1}{4} - \frac{1}{4}\sqrt{7}), \quad N = (-\frac{1}{4} + \frac{1}{4}\sqrt{7}, \frac{1}{4} + \frac{1}{4}\sqrt{7}).$$

Then by Proposition 2.1, $e(AM) = |-\frac{1}{4} + \frac{1}{4}\sqrt{7}|\sqrt{2} = \frac{\sqrt{2}}{4}(\sqrt{7} - 1) = e(BN)$, and $e(AN) = \frac{\sqrt{2}}{4}(\sqrt{7} + 1) = e(BM)$. So

$$\begin{aligned} d_{\mathbb{H}}(AB) &= \ln\left(\frac{\frac{\sqrt{2}}{4}(\sqrt{7}+1)\frac{\sqrt{2}}{4}(\sqrt{7}+1)}{\frac{\sqrt{2}}{4}(\sqrt{7}-1)\frac{\sqrt{2}}{4}(\sqrt{7}-1)}\right) \\ &= \ln\frac{(\sqrt{7}+1)^2}{(\sqrt{7}-1)^2} = 2\ln\frac{(\sqrt{7}+1)}{(\sqrt{7}-1)} = 1.59\ldots\ . \end{aligned}$$

(5) $\mathbb{G}$: The *gap*, or *missing strip plane*.

The points of $\mathbb{G}$ are all those of $\mathbb{E}$ *except* those (x, y) with $0 < x \leq 1$ (Figure 2.7).

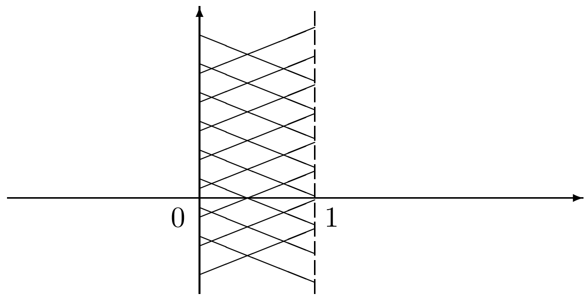

Figure 2.7

So the y-axis is part of $\mathbb{G}$, but the line $x = 1$ is *not* (and neither is any vertical line $x = a$ if $0 < a < 1$).

Lines in $\mathbb{G}$ are defined to be the same as in $\mathbb{E}$, *except* that for any nonvertical line $y = mx + b$, the part in the missing strip is deleted. So a typical nonvertical line l consists of all (x, y) with $y = mx + b$ (m, b fixed) *and* with $x \leq 0$ or $x > 1$ (Figure 2.8).

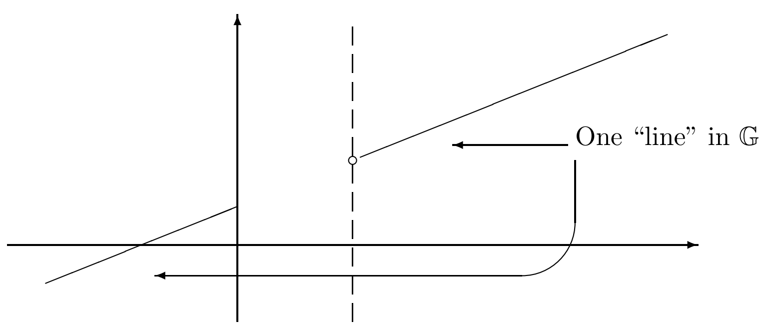

Figure 2.8

Distance: For points A, B in $\mathbb{G}$, we define $d_{\mathbb{G}}(AB)$ as follows. First, if A and B lie on opposite sides of the gap, let C be the point where segment $\overline{AB}$ meets the y-axis, and D the point where $\overline{AB}$ meets the vertical line $x = 1$ (D is not in $\mathbb{G}$; see Figure 2.9).

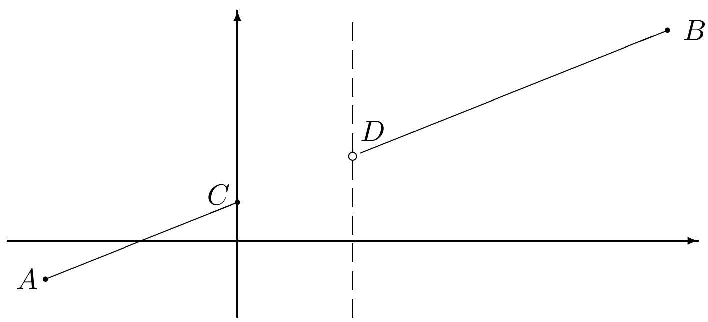

Figure 2.9

Now define

$$d_{\mathbb{G}}(AB) = \begin{cases} e(AB) & \text{for } A, B \text{ on the same side of the gap;} \\ e(AB) - e(CD) & \text{for } A, B \text{ on opposite sides.} \end{cases}$$

Example. Let $A = (-2, 0), B = (2, 0)$ (see Figure 2.10). Then $C = (0, 0), D = (1, 0)$ and

$$d_{\mathbb{G}}(AB) = e(AB) - e(CD) = 4 - 1 = 3.$$

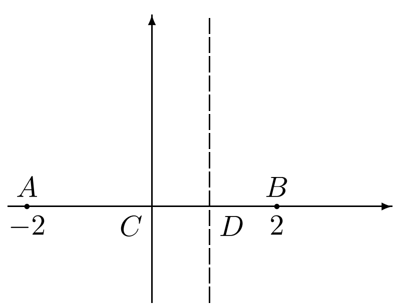

Figure 2.10

Example. $A = (-2, 0), B = (2, 1)$ lie on line l: $y = \dfrac{1}{4}x + \dfrac{1}{2}$, hence $C = (0, \dfrac{1}{2}), D = (1, \dfrac{3}{4})$ (see Figure 2.11).

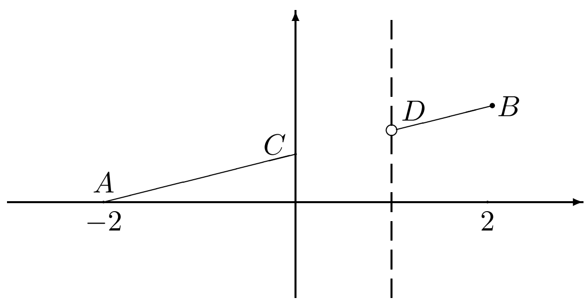

Figure 2.11

Now

$$e(AB) = \sqrt{(2+2)^2 + (1-0)^2} = \sqrt{17}\ ,$$

and

$$e(CD) = \sqrt{(1-0)^2 + (\frac{3}{4} - \frac{1}{2})^2} = \sqrt{1 + \frac{1}{16}} = \sqrt{\frac{17}{16}} = \frac{1}{4}\sqrt{17}\ .$$

So $d_{\mathbb{G}}(AB) = \sqrt{17} - \frac{1}{4}\sqrt{17} = \frac{3}{4}\sqrt{17}$.

Each of the preceding five examples is a specific model of some geometric system. The similarities and differences among these models will help us to understand not only that alternative systems of geometry are just as real as is the Euclidean plane, but also that difficulties (and dangers) may arise when arguments use steps that are not explicitly justified.

For instance, consider the following proof of a form of the Exterior Angle Theorem in the Euclidean plane. This theorem plays a crucial role in the study of parallel lines (see, for example, the proof of Theorem 17.6). We will assume for the present discussion that you recall the relevant concepts from high school geometry. If not, don't worry. All of the ideas will be developed from scratch later in our more general setting.

Exterior Angle Inequality. An exterior angle of a triangle is greater than either remote interior angle. That is, if $\triangle ABC$ is any triangle and D is on the extension of $\overline{BC}$ through C, then $\angle ACD$ is greater than each of $\angle A$ and $\angle B$ (Figure 2.12).

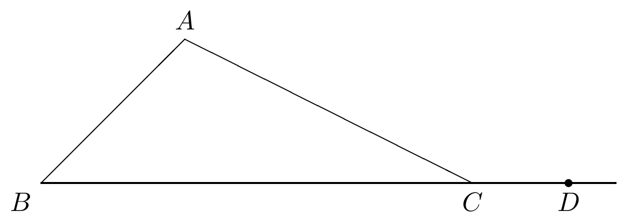

Figure 2.12

Proof. Let M be the midpoint of $\overline{AC}$, and extend segment $\overline{BM}$ its own length through M to E. Then lengths $AM = MC$ and $BM = ME$, while angle measures $\angle AMB = \angle CME$ (vertical angles are equal). Thus $\triangle AMB \cong \triangle CME$ (by the Side-Angle-Side criterion for congruence of triangles). So $\angle ECM = \angle BAM$ (corresponding angles of congruent triangles are congruent) (Figure 2.13).

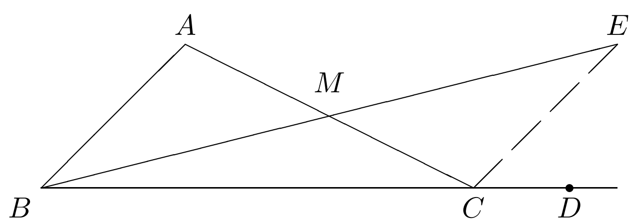

Figure 2.13

Since

$$\begin{aligned}\angle ACD &= \angle ACE + \angle ECD\\ &= \angle ECM + \angle ECD\\ &= \angle BAM + \angle ECD > \angle BAM = \angle A,\end{aligned}$$

we have that $\angle ACD > \angle A$.

To show that $\angle ACD > \angle B$, extend $\overline{AC}$ through C to F, forming $\angle BCF$ (Figure 2.14). Then use the procedure of the first part of the proof to show $\angle BCF > \angle B$. (Let N be the midpoint of $\overline{BC}$; extend $\overline{AN}$ its own length through N, etc.)

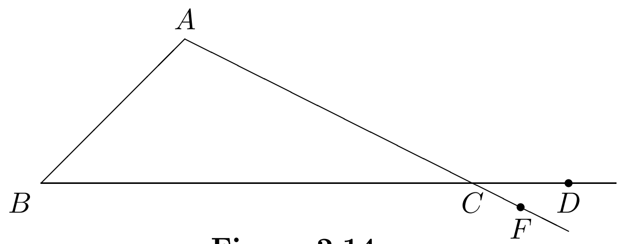

Figure 2.14

Finally, note that $\angle ACD = \angle BCF$ (vertical angles), and so $\angle ACD > \angle B$.

Now consider $\mathbb{S}$, on which "lines" are great circles. At a point where two great circles meet they form (four) angles, which may be given the "usual" angle measure. That is, assign the standard measure of the angles formed by the tangent lines to the great circles at the point of intersection. Assume that the notions of segment, triangle, and congruence all make sense in $\mathbb{S}$, in the most obvious way (they do!). Vertical angles have equal measure, and the Side-Angle-Side criterion for congruence of triangles is valid in $\mathbb{S}$. So it might seem plausible that the Exterior Angle Inequality holds here also. That is, a direct analog of the argument given for $\mathbb{E}$ might work in $\mathbb{S}$. However, consider the triangle $\triangle ABC$ shown in Figure 2.15.

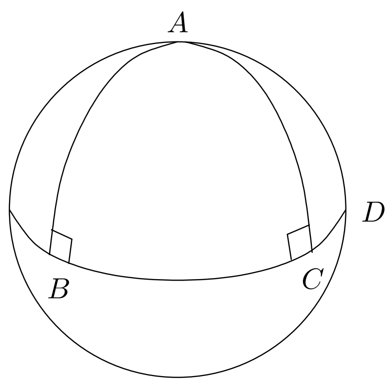

Figure 2.15

Think of A as the north pole, and B and C on the equator with $\angle A = \angle BAC > 90$. Then the exterior angle $\angle ACD = 90$, which is *less* than $\angle A$. So what goes wrong when we try to apply the preceding argument?

The problem lies in the assertion $\angle ACD = \angle ACE + \angle ECD$. There is a tacit assumption that the rays $\overrightarrow{CA}$, $\overrightarrow{CE}$, and $\overrightarrow{CD}$ are positioned so that the angle measures add in this way (i.e., that $\overrightarrow{CE}$ is "between" rays $\overrightarrow{CA}$ and $\overrightarrow{CD}$). There was no justification for this in the previous argument, except for "the picture" (Figure 2.13). When our rays, or segments, are parts of great circles in $\mathbb{S}$, a picture (see Problem 12) indicates that the relative positions may be quite different. The failure of the Exterior Angle Inequality in $\mathbb{S}$ is related to the fact that in the triangle $\triangle ABC$ under discussion, $\angle A + \angle B + \angle C$ *does not equal* 180. The extent to which the Exterior Angle Inequality does hold in $\mathbb{S}$ (and more generally) is discussed in Chapter 15.

It is important to note that assertions such as "ray $\overrightarrow{CE}$ is between $\overrightarrow{CA}$ and $\overrightarrow{CD}$" or "$\angle ACD = \angle ACE + \angle ECD$" *cannot be proved* from the given postulates of classical Euclidean geometry (see Appendix I). They follow only from unspoken assumptions about "betweenness" of rays and points, which are certainly plausible from pictures of configurations in the plane but which are never explicitly mentioned by Euclid. The preceding comparison shows that it pays to be aware of concepts such as "betweenness" and of the assumptions that we make about them. One of the goals of our rigorous development will be to make explicit all of the underlying suppositions about our geometry and thus to gain a deeper understanding of its fundamental concepts.

Problem Set 2

1. (a) In $\mathbb{E}$, find the distance between $A(-\frac{1}{3}, 0)$ and $B(0, \frac{1}{3})$.

   (b) In $\mathbb{M}$, find the distance between $A(-\frac{1}{3}, 0)$ and $B(0, \frac{1}{3})$.

   (c) In $\mathbb{H}$, find the distance between $A(-\frac{1}{3}, 0)$ and $B(0, \frac{1}{3})$.

   (d) In $\mathbb{S}$, (radius $r = 1$), find the distance between $C(0, 0, 1)$ and $D(0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$.

(e) In $\mathbb{S}$, (radius $r = \frac{1}{2}$), find the distance between $P\left(\frac{1}{4}, \frac{\sqrt{2}}{4}, -\frac{1}{4}\right)$ and $Q\left(\frac{1}{6}, -\frac{1}{3}, \frac{1}{3}\right)$.

(f) In $\mathbb{G}$, find the distance between $A(-2, -3)$ and $B(4, 6)$.

2. Find two points A, B in $\mathbb{H}$ such that $d_{\mathbb{H}}(AB) > 13$. Show the calculation that justifies your answer.

3. Prove Proposition 2.1.

4. Prove Proposition 2.2.

5. Let $A(x_1, y_1)$ and $B(x_2, y_2)$ be two points on opposite sides of the gap in $\mathbb{G}$ and on the line $l: \; y = mx + b$. *Derive a formula* for $d_{\mathbb{G}}(AB)$ in terms of x_1, x_2 and m.

6. Find (sketch) the unit circle in $\mathbb{M}$; that is, the set of all points that are (Minkowski) distance 1 from the origin $(0, 0)$.

7. Find all points of the form $P(x, 0)$ in $\mathbb{H}$ whose ($\mathbb{H}$) distance from $O(0, 0)$ is $\ln 2$.

8. True or false: For all points, P, Q, R in $\mathbb{G}$,

$$d_{\mathbb{G}}(PQ) + d_{\mathbb{G}}(QR) \geq d_{\mathbb{G}}(PR)?$$

Justify your answer.

9. True or false: For all noncollinear points P, Q, R in $\mathbb{M}$,

$$d_{\mathbb{M}}(PQ) + d_{\mathbb{M}}(QR) > d_{\mathbb{M}}(PR)?$$

Justify your answer.

10. Let $P(a, b, c)$ and $Q(x, y, z)$ be points on the sphere $\mathbb{S}$ of radius r centered at $O(0, 0, 0)$, as in Figure 2.16.

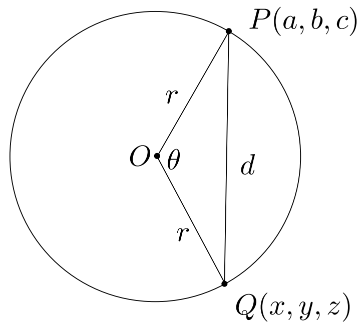

Figure 2.16

Let d be the Euclidean distance PQ and θ be the radian measure of $\angle POQ$.

(a) Recall the Law of Cosines for the triangle POQ and use it to show that $\cos\theta = (2r^2 - d^2)/2r^2$.

(b) Recall the Euclidean distance formula for points in three-dimensional space and use it and part (a) to show that $\cos\theta = \dfrac{ax + by + cz}{r^2}$.

(c) Use (b) to derive that $d_{\mathbb{S}}(PQ) = r\cos^{-1}\left(\dfrac{ax + by + cz}{r^2}\right)$.

11. Let $P(1, 0, 0)$ be on the sphere $\mathbb{S}$ of radius 1. Find the set of points $Q(x, y, z)$ on $\mathbb{S}$ such that $d_{\mathbb{S}}(PQ) = 1$. That is, solve for the coordinates x, y, z and describe the set of points geometrically as well.

12. Draw a careful picture on $\mathbb{S}$ for the steps analogous to those in the argument for the Exterior Angle Inequality, but applied to $\triangle ABC$ on $\mathbb{S}$, as shown in Figure 2.15. Show the true relative positions of $\overrightarrow{CA}$, $\overrightarrow{CE}$, and $\overrightarrow{CD}$ (a ray here is half of a great circle). What can you conclude about $\angle ACD, \angle ACE$, and $\angle ECD$? (Use dotted lines to indicate parts of arcs on the "back side" of $\mathbb{S}$. Or, you can draw on the surface of a sphere (styrofoam ball, tennis ball, etc.).)

13. Let S be the part of the unit circle in $\mathbb{G}$ that lies to the right of the gap. That is, for $O = (0, 0)$ and $P = (x, y)$,

$$S = \{P \in \mathbb{G} : \ d_{\mathbb{G}}(OP) = 1 \text{ and } x > 1\}.$$

(Set notation is reviewed in Chapter 4). Find an equation in x and y for the elements of S. (Hint: If $y = mx$, use Problem 5 of this problem set and note that $m = y/x$.)

14. Let C be the unit circle in $\mathbb{H}$. That is, for $O = (0,0)$ and $P = (x,y)$,

$$C = \{P \in \mathbb{H} : \ d_{\mathbb{H}}(OP) = 1\}.$$

Find an equation in x and y for the elements of C. In particular, show that C is a circle (of what radius?) in $\mathbb{E}$. (Hint: If $y = mx$, find the coordinates of M and N in terms of m, and then find $d_{\mathbb{H}}(OP)$ in terms of x and m. Note that $y = m/x$ if $x \neq 0$. Alternatively, we can set Euclidean distance $e(OP) = r$, derive a formula for $d_{\mathbb{H}}(OP)$ in terms of r, then set this expression equal to 1 and solve for r.)

15. Let D be the circle in $\mathbb{H}$ with radius 1 and center $Q = (\frac{1}{2}, 0)$. That is, for $Q = (\frac{1}{2}, 0)$ and $P = (x, y)$,

$$D = \{P \in \mathbb{H} : \ d_{\mathbb{H}}(QP) = 1\}.$$

(a) Find all points P on the x-axis that are also in D.

(b) Find all points P on the line $x = \frac{1}{2}$ that are also in D.

(c) Is D a circle in $\mathbb{E}$? Explain.