# Geometry

**Mark J. Nielsen**
Professor of Mathematics, University of Idaho
Moscow, Idaho

March 26, 2008

# To the Instructor

Most all colleges and universities offer a course in geometry for an audience consisting primarily of secondary education and mathematics majors. This text was developed to use in just such a course at the University of Idaho. I have taught that course many times, and in so doing have come to two conclusions about what a good geometry text should do.

**1.** First, I saw a need for a renewed emphasis on learning proof. The targeted audience needs this skill. Mathematics majors cannot become true mathematicians without developing mastery of reading and writing proofs, and education majors with a poor understanding of proof will become poor teachers of mathematics.

**2.** I saw the importance of catching the imagination of the students right away. Many students enter this course believing geometry to be both difficult and boring. Unless that preconception is changed within the first couple weeks of a semester, the course will be a difficult experience for both student and instructor.

When I began assembling my own materials into a text, I tried to address these two issues. I decided that the old-fashioned axiomatic approach to geometry was the correct way to address the first goal. It is the only intellectually honest approach to the subject, and the best tool yet invented for developing skill at deductive reasoning.

The second goal was more difficult to address, in part *because* of my conclusion on the first issue! A traditional axiomatic treatment (in which the first goal encountered is that of proving routine facts from axioms that are not appreciably more obvious) loses too many students in its immediate formality. There had to be a better way, and I think I've found one in the way this text is structured.

The first five chapters of this book are built around what I call the "basic toolbox" theorems of Euclidean geometry – the standard facts usually covered in a high school geometry course. Rather than spend a lot of time up front proving these facts in full formality, we instead review them briefly in Chapter 1, then show their usefulness in Chapter 2 through short excursions to several dazzling theorems of more advanced Euclidean geometry. The first part of the course, then, becomes an extension and enhancement of the students' previous experience, rather than a re-hashing (albeit in more formal terms) of that same material. The advantages of this are:

- The students experience the "wow factor" of these theorems early in the course when their attention is still up for grabs.

- Because the more formal axiomatics are delayed, the students can warm up their proof skills in a friendlier setting.

- The fact that such impressive results can be derived from the "basic toolbox" theorems motivates the long task of the chapters that follow: putting that toolbox on solid footing by developing its theorems from axioms.

Chapter 2 concludes with a set of exercises that may be used throughout the course as reminders of the power of the "toolbox" theorems and the intrinsic appeal of Euclidean geometry.

Chapters 3 through 6 form the heart of the text. In Chapter 3 we give a more complete treatment of the axiomatic method than is usually contained in geometry texts. I believe that students will do better at following a development of geometry from axioms if they are told ahead of time *what* the rules are and *why* they are what they are. This chapter tries to accomplish that feat. In addition to setting out the rudiments of the axiomatic method, it includes practice at using the method in the simple (but nontrivial) setting of some finite geometries. It also gives a quick survey of the history of axiomatic treatments of Euclidean geometry and sets forth the fundamental role of the parallel postulate in differentiating Euclidean and hyperbolic geometry.

Chapters 4 through 6 then set out on the road map described in Chapter 3. Chapter 4 introduces axioms for neutral geometry and proves the neutral "toolbox theorems" from them. As is always the case in an axiomatic development, much of this early work is done in proving very basic facts, and the process can be conceptually challenging to many students. Chapter 5 is more straightforward, adding the parallel postulate to the set of axioms and proving the remaining "toolbox theorems".

Chapter 6 then develops hyperbolic geometry as an alternative extension to the neutral geometry from Chapter 4. The material is introduced alongside the story of its discovery. Aside from teaching the history of one of the great mathematical struggles of all time, this historical approach has another advantage: the relationship between Euclidean and hyperbolic geometry is emphasized by the stunning moment when the approach changes from trying to *disprove* the parallel postulate within neutral geometry to *accepting* its negation as a new axiom – all the attempts at finding a contradiction suddenly become a host of surprising theorems in a new geometry. The historical treatment concludes with the introduction of Poincaré's half-plane model for hyperbolic geometry and a discussion of its significance. The last part of Chapter 6 then turns to a more in-depth treatment of hyperbolic geometry, developing the basic facts of angle of parallelism, asymptotic parallels, and ideal points.

Chapter 7 provides a synthesis for the material from Chapters 4 to 6 by treating transformational geometry. The basic facts concerning transformations of the plane are developed in a neutral geometry setting, then applied to classify the transformations of both Euclidean and hyperbolic geometry.

This text should be more-or-less self contained, and should provide ample material for a one semester course. The amount of time spent on the introductory chapter will of course depend on the level of preparation of the students. The same is true to some extent of Chapter 3. In a few places, but most notably in the first part of Chapter 4 and the latter part of Chapter 6, we make use of some basic concepts from calculus – limits and continuity of functions. Though presumably the students taking a course in college geometry will have previously taken a calculus course, it may be advisable to review these concepts when needed. Previous exposure to linear algebra will make some of the calculations in Chapter 7 more meaningful to the students, but is not strictly necessary.

I have not tied the book to the use of dynamic geometry software, but I have suggested many places where explorations with such software can be easily incorporated into the course. I myself choose to use such software both for classroom demonstration and also for selected student homework problems. I don't believe that such technology should dominate one's study of geometry, but its power as a tool for exploration and illustration is impossible to dispute.

# To the Student

Geometry is perhaps the oldest branch of mathematics, its origins reaching some 5000 years back into human history. And the story of geometry is as rich as it is long. The first mathematical proofs were in geometry, and the great philosophers of ancient Greece regarded the study of geometry as essential to the development of wisdom. Geometry has been the setting of some of history's greatest intellectual challenges and achievements. Examples you may have heard of would include the straightedge and compass constructions of the early Greeks, Descartes' development of analytic geometry, and the two-millennium long battle over Euclid's "parallel postulate" culminating in the 19th century discovery of non-Euclidean geometries. If you haven't heard of these, don't fret – we'll encounter them all in the pages that follow! But despite geometry's ancient origins, the subject is far from stale. It remains today an important area of mathematical discovery and research, with applications ranging from robotic engineering to theoretical physics.

### Why study geometry?

From Plato (who inscribed "*Let no man ignorant of geometry enter here*" over the doorway of his Academy) to modern school boards, the study of geometry has long been regarded as a staple of the standard curriculum in our schools. Why is this, and should it remain so? I believe it should! In fact, I see three excellent reasons for *any* person to study geometry.

**Reason 1: Usefulness.** Geometric reasoning is something we all use. Many everyday actions involve reasoning with shape, distance, and volume, and adeptness at such reasoning is a learned skill. Also, geometry plays a part in the training necessary for many careers. Courses in physics and engineering as well as other areas of mathematics all presuppose some experience with geometric reasoning.

**Reason 2: Logic.** More important than its direct usefulness, though, is the fact that geometry has always been one of the best vehicles for indirectly teaching the process of deductive reasoning and logic. The construction and analysis of proofs has been central to the study of geometry since the time of Euclid, and no exercise has been found to sharpen our reasoning capabilities more effectively than working with proofs. Whether or not we ever use again the concepts of trapezoids, equilateral triangles, and perpendicular bisectors after leaving a course in geometry, we *all* use reasoning, and we are *all* made better by the improvement of our abilities in the art of reasoning.

**Reason 3: Beauty.** In my opinion, the most important reason for studying geometry is its beauty. Modern geometry is full of surprising results and elegant techniques, and even the "old fashioned" geometry of Euclid has an intellectual appeal great enough to have captured the attention of the world's greatest minds for centuries. I hope this book carries that message — geometry is beautiful!

## Why is geometry so difficult?

Despite the virtues listed above, geometry is regarded as a difficult and unpopular subject by many students. Why is that? Most often, "all those proofs" is the cited reason. Unfortunately, there has been a trend over the last generation to answer this objection (consciously or not) by diminishing the role of proof in the way geometry is taught. In many current geometry texts, proofs are reduced both in number and in sophistication. The formality of developing geometry from axioms is often replaced by a breezy informal treatment of geometric highlights. This text will follow no such trend! Our approach will be unapologetically axiomatic, and proof will be at the heart of all we do. In fact, our major goals will be

- to develop our ability to write proofs, and

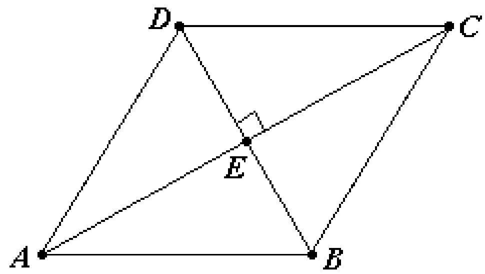
- to understand the workings of the axiomatic system.

Both of these are essential points for mathematicians and mathematics educators.

But our faithfulness to the prominence of mathematical proof need not mean that our road will be unpleasant or overly difficult. In my experience there are three sources for difficulty in mathematical proof. If we are careful in addressing these, we can make learning the art of proof a much easier task.

The difficulties are perhaps best illustrated by considering a specific example. So, before reading on in this introduction, take a moment with a sheet of scratch paper and try to prove the following theorem (probably one you proved back in that high school class):

THEOREM. *Every parallelogram with perpendicular diagonals is a rhombus. That is, if $ABCD$ is a parallelogram (so that $AB$ is parallel to $CD$ and $BC$ is parallel to $AD$) and $AC$ is perpendicular to $BD$ then all of the sides $AB$, $BC$, $CD$, and $DA$ are of equal length.*




Did you succeed? Was the process frustrating? Now read my list of factors contributing to the difficulty of writing proofs. Do they strike a chord with what you just experienced?

**Reason 1: Uncertain ground rules.** A common frustration in writing geometric proofs is expressed in the distress cry "What am I allowed to use in my proof?" In writing your proof of the above theorem, did you use the fact that $\angle CAB$ and $\angle ACD$ are congruent (because they are "alternate interior angles" of a transversed pair of parallel lines)? Did you use the fact that $AB$ and $CD$ have equal length, or did you actually prove that along the way? You weren't told what was a valid assumption and what needed proof, and there's no obvious list of what *should* be allowed. Did that make the process frustrating?

We can think of constructing proofs as a building process, where the building tools are the valid assumptions and facts we have at our disposal. If we don't know the make-up of our toolbox, the building process is indeed bewildering. This is where the old-fashioned axiomatic approach to geometry has a clear advantage (and where the newer "breezy" treatments have a clear disadvantage). In the axiomatic approach we make our assumptions explicit at the outset, setting them forth as the *axioms*. For the first proof we write, these are the only tools available for use. However, once a statement is proved, it becomes a tool available for use in building the next proof – to prove Theorem 37 we can use the axioms and any of Theorems 1 to 36. There should never be any doubt as to what can be used in a proof!

So, we can (and will!) avoid this first difficulty. Though we will not start an axiomatic treatment of plane geometry until Chapter 4 we will always make explicit what assumptions are allowed. It will still require care to avoid illegal steps in our proofs, but at least there should be no confusion about whether or not a particular step is allowable.

**Reason 2: Non-sequentiality.** We think sequentially, but geometry is inherently non-sequential. Our proofs are strings of statements, each following neatly from the one before it. Algebra is well adapted to this structure since reducing equation to equation to equation is itself a sequential process. But geometric thought processes are often highly *non*-sequential. In the diagram for our example above, there is a lot going on all at once! You may have scratched out several sets of congruences of various angles and sides, but that just makes for a messy diagram! We don't have a proof until we find a way to line up those diagram observations into a step-by-step path leading through the mess to the desired conclusion. That is why we sometimes experience the frustration of being able to "see why the theorem is true" while being at a loss for how to write down our reasons.

This difficulty is real and unavoidable. But it is not insurmountable. Translating the non-sequential reasoning from the diagram into a sequential string of sentences is something that *can* be learned with practice. Just being aware of the difficulty will probably help somewhat, since we know what process to focus on. You shouldn't expect to be a master at proof-writing immediately, but with practice anyone can improve his or her skill.

My last reason for the difficulty of the proof process is not illustrated by our little example, but it is certainly no less real.

**Reason 3: Formality.** A full-blown use of the axiomatic method is not an easy proposition. Its formality is sometimes intimidating if not stifling to students encountering it for the first time. The assumptions we start with as axioms usually constitute a very humble beginning and one often spends a great deal of effort in the first several theorems proving things that seem no less obvious than the axioms themselves. These first theorems constitute a "boot-up" process by which a collection of useful facts are established using only the meager set of axioms. These facts are then available for use in proving the more substantial and interesting theorems later in the development.

This initial phase of an axiomatic treatment is in many ways the most difficult because it is the most formal. Its placement at the beginning, though, is unfortunate, because inexperienced students struggle with its formality and fail to see its motivation. Where a more experienced student might be captivated by the power of deriving a large body of knowledge from a small set of assumptions, an uninitiated student might be completely lost in the process and see nothing of its beauty.

We will work around this difficulty by delaying the onset of our axiomatic treatment. Instead, after an initial introductory chapter we will set out in Chapter 2 on some truly captivating geometric excursions. We will set out as our allowable assumptions a "basic toolbox" of Euclidean geometry facts, given in Section B of Chapter 1. (If you had a high school geometry course, these facts will be familiar.) From them we will develop a few amazing theorems that you may not have seen before. These excursions will give you practice at proof and geometric reasoning. Then when we study the axiomatic method in Chapter 3 and begin applying it in subsequent chapters, the formality of the process might be overcome by an appreciation for its purpose.

## Prerequisites

We assume that the reader of this text has had a previous exposure to Euclidean geometry similar to what is usually included in a high school geometry course. There are a few references to topics from calculus; specifically, basic facts about limits and the notion of a continuous function. (These are most crucially used in Chapters 4 and 6.) Though we discuss the generalities of mathematical proof in Chapter 1, a previous exposure to a proof-based mathematics class is desirable. Developing the skill of reading and writing proofs is one of the goals of the text, so you need not feel like a "proof master" to begin. But in general, the more exposure you've had to proof in the past, the greater your advantage will be in this course of study.

## An invitation...

So now it's time to begin. Whatever your reasons for undertaking a study of geometry now, and whatever your prior experiences with the subject might be, open yourselves to the possibility that you may enjoy the journey we take. Prepare to work hard, to think deeply, and to explore a subject that lit the flame of western intellectual history over two millennia ago and continues now to chart our course to the future.

# Notation

Throughout the text we will use the following conventions for notation. Most are standard for mathematical descriptions of sets, and none should be too unfamiliar. Section 1B will review many of these terms and symbols in more detail.

### Characters

- Capital roman letters like $A$, $B$, and $C$ will usually denote points in the plane.
- Lower case roman letters like $a$, $b$, and $c$ will usually denote numerical values such as lengths.
- Capital Greek letters such as $\Gamma$, $\Lambda$, and $\Sigma$ will usually denote sets of points such as circles or lines.
- Lower case Greek letters such as $\alpha$, $\beta$, and $\gamma$ will usually denote the (numerical) measures of angles.

### Set theory notions

- $\Sigma_1 \cup \Sigma_2$ denotes the union of sets $\Sigma_1$ and $\Sigma_2$.
- $\Sigma_1 \cap \Sigma_2$ denotes the intersection of sets $\Sigma_1$ and $\Sigma_2$.
- $\Sigma_1 \setminus \Sigma_2$ denotes the set of points which are elements of $\Sigma_1$ but are not elements of $\Sigma_2$.
- $A \in \Sigma$ denotes the fact that point $A$ is an element of set $\Sigma$.
- $\Sigma_1 \subset \Sigma_2$ denotes the fact that $\Sigma_1$ is a subset of $\Sigma_2$.

### Geometric objects

- $AB$ will denote the segment with endpoints $A$ and $B$.
- $|AB|$ will denote the length of the segment $AB$.
- $\overrightarrow{AB}$ will denote the ray initiating at point $A$ and passing through point $B$.
- $\overleftrightarrow{AB}$ will denote the line determined by points $A$ and $B$.

- $\Lambda_1 \parallel \Lambda_2$ will denote that lines $\Lambda_1$ and $\Lambda_2$ are parallel. (The same notation may be used with segments or rays in place of lines.)
- $\Lambda_1 \perp \Lambda_2$ will denote that lines $\Lambda_1$ and $\Lambda_2$ are perpendicular. (The same notation may be used with segments or rays in place of lines.)
- $\angle APB$ will denote the angle consisting of the rays $\overrightarrow{PA}$ and $\overrightarrow{PB}$, and $m\angle APB$ will denote its angle measure.
- $P_1P_2 \ldots P_n$ will denote the $n$-sided polygon whose vertices are $P_1$, $P_2$, $\ldots$, $P_n$ (see p.13).
- $\text{area}(P_1P_2 \ldots P_n)$ will denote the area of the region inside the polygon $P_1P_2 \ldots P_n$.
- The polygon $ABC$ is (of course) a triangle. If we wish to specify the order of vertices (for example, to designate corresponding parts of congruent triangles) we will use the symbol $\triangle ABC$ for this triangle.
- For the triangle $ABC$ we will denote the length of a side by the lower case symbol of the opposite vertex. Thus, $|BC| = a$, $|AC| = b$, and $|AB| = c$.
- Also for a triangle $ABC$ we will shorten the names of angles as follows: $\angle CAB = \angle A$, $\angle ABC = \angle B$, and $\angle ACB = \angle C$.
- The symbol $\cong$ will be used to designate congruence of geometric figures.
- $A\overset{\frown}{P}B$ will denote the circular arc containing $P$ and determined by the chord $AB$. The measure of this arc will be denoted $m(A\overset{\frown}{P}B)$ (see p.20).

# Contents

# Chapter 1

# Preparations

Where should we begin our study of geometry? We have much ground to cover, and much work to accomplish along the way. We will devote this opening chapter to setting the stage for our journey.

- First, in Section A we will consider the origins of the subject matter we now call *geometry* in the civilizations of ancient Egypt, Babylon, and Greece. A thorough treatment of that topic would be well beyond the scope of this text, but we hope our brief summary gives at least some sense of geometry's place in the intellectual history of humankind.

- Section B reviews the basic facts you should remember from high school geometry. Our goal is to establish some terminology and set forth what we will call our "basic toolbox" of Euclidean geometry theorems.

- In Section C we cover the terms and techniques of deductive reasoning and proof, the principal tools in any systematic study of geometry.

- Finally, Section D presents an opportunity to brush up on the use of those tools as we consider the straightedge and compass constructions of early Greek geometry.

## A. A Brief History of Geometry to Euclid

It's impossible to point to a moment in history when geometry was born, but the subject's name itself may give some hint of its origins. *Geometry* means literally *"measurement of the earth"*, and while the derivation of the word is not entirely certain, it is probably a product of very practical origins. Indeed, surveying plots of land may have been the first "geometry" problems considered by humans. In both Egyptian and Babylonian mathematical writings, much attention is paid to the calculation of areas, and specific examples of surveying land plots are sometimes used. But most of what we now associate with geometry came much later in contributions by the great Greek thinkers. In this section we'll give a quick description of the development of geometry in these three ancient cultures.

### Geometry in Egypt

The Egyptian civilization's contributions to mathematics span the centuries from 3000 BC to about 1800 BC The exact level of achievement in Egyptian mathematics is impossible to ascertain from the few scrolls that survive to the present. For instance, some have suggested that the Egyptians were familiar with the relationship among the side lengths of a right triangle that we now call the Pythagorean Theorem, yet no evidence to support this exists in the surviving writings.[1]

It *is* clear that they had deduced a correct calculation for the area of an isosceles triangle. In fact, their derivation of this rule seems very much akin to geometric proof: divide the triangle into two right triangles using the bisector of the vertex angle, then arrange the two right triangles to form a rectangle (see Figure 1.1). That this was apparently not extended by the Egyptians to the calculation of area for an arbitrary triangle might seem curious to us, but is actually indicative of one of the weaknesses of Egyptian mathematics: the jump from specific examples to general theorems was never made. The problems considered in their treatises were always specific calculations. While it is clear that some rules of calculation were developed, these rules were not written as general formulae, and relationships among geometric objects were not studied systematically.

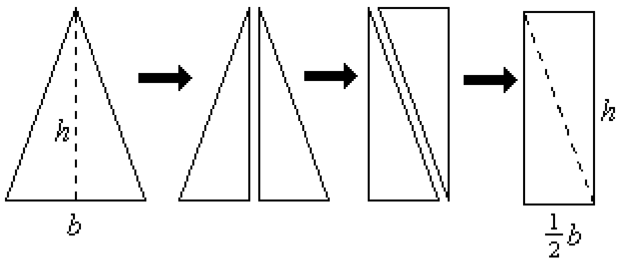



Figure 1.1: A geometric justification that the area of an isosceles triangle is the product of its height and half its base length

[1]It is clear that they were aware of the one case of a "3-4-5" right triangle; for they made use of it to measure right angles in their buildings. They simply divided a long loop of rope into 12 equal lengths. Then, with three people grasping the loop at positions 0, 3, and 7, would pull it taught to form the triangle. The angle formed by the rope at position 3 is then a right angle!

Another weakness of Egyptian mathematics is a failure to distinguish between exact calculation and approximation. For example, the apparent rule used for the area of a quadrilateral region was that the area would equal the product of the arithmetic means of the lengths of opposite sides. Thus, the area of a quadrilateral region with side lengths $a$, $b$, $c$, and $d$ (in order around the region's boundary — see Figure 1.2) would be calculated as $\left(\frac{a+c}{2}\right)\left(\frac{b+d}{2}\right)$. The fact that this gives only a rough approximation for the actual area is never stated if it was known at all.

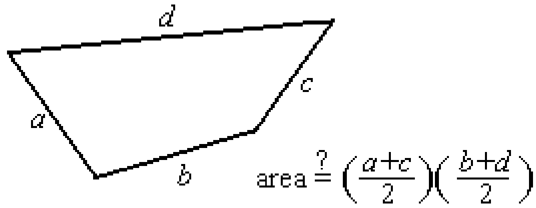


Figure 1.2: An incorrect area formula for quadrilaterals

One surviving Egyptian scroll (the Rhind papyrus) contains the assertion that the area of a circle of diameter 9 is equal to the area of a square of side length 8 (see Figure 1.3). Areas of other circles were then computed by dividing their diameters into nine equal parts and forming a square on eight of those nine parts. This rule anticipates the fact that there is a constant ratio relationship between the area of a circle and the square of its diameter. (The Greeks would later prove this fact, and then would themselves expend much effort in a search for a method to construct a square whose area is equal to that of a given circle.) In fact, the Egyptian's method amounts to an approximation of the ratio we call $\pi$ by the fraction $\frac{256}{81} \approx 3.16$ (see Exercise 1.7), though it is not clear that they understood this formula as an estimate.

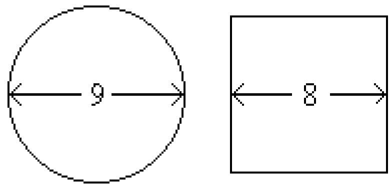


Figure 1.3: Equal areas?

Again, our modern approach of searching for general relationships stands in contrast to the Egyptian fixation with specific calculations. If the goal of the Egyptians was to develop a method good enough to calculate taxes on a circular plot of land, then their method is satisfactory and no distinction between exactness and approximation need be made. That the Egyptians apparently made no efforts to find increasingly accurate approximations for the ratio we call $\pi$ (as the Greeks later would do) is good evidence that this was indeed the mindset from which they worked.

### Geometry in Babylonia

The Babylonian empire flourished in the valley of Tigris and Euphrates rivers (in present day Iraq) from 2000 BC to 600 BC During these centuries the Babylonians made many great strides in mathematics. They, like the Egyptians, continued to be limited to specific calculations, but they did take steps toward a more general approach. They were familiar with the Pythagorean Theorem previous to 1000 BC, many hundred years before Pythagoras lived (though it is doubtful that they had a demonstration or "proof" of this fact). Also, they were clearly aware of the fact that we now call the Theorem of Thales (named after a Greek mathematician we will discuss shortly) that a triangle inscribed in a semicircle is a right triangle. Again, it is doubtful that they could have "proved" this theorem.

The Babylonians also continued the Egyptians' lack of distinction between approximation and exact calculation. In fact, the same inexact formula used by the Egyptians for the area of a quadrilateral (Figure 1.2) was used by the Babylonians. Yet the Babylonians did have a powerful tool for handling approximation not available to the Egyptians: their number system.

The Babylonian number system, like our own Hindu-Arabic number system, used place value for both whole and fractional parts. But whereas we use a base 10 system, the Babylonians used base 60.[2] Thus, the Babylonian number

$$a \quad b \quad c \quad ; \quad d \quad e \quad f$$

would have the value

$$(a \times 60^2) + (b \times 60) + c + \frac{d}{60} + \frac{e}{60^2} + \frac{f}{60^3}$$

(Note here that we have used the semicolon as a "decimal point".) This gave the Babylonians all the power of decimal approximations, and this power did seem to inspire a search for exactness. For example, because the Babylonians were familiar with the Pythagorean Theorem they were well aware that the diagonal of a square of side length 1 must be a number whose square is 2. In fact, they knew that for any square the ratio of diagonal length to side length would be this number. Though the Babylonians apparently used the number

$$1 \quad ; \quad 25 \quad = \quad 1.41\overline{666}$$

as a ready approximation for $\sqrt{2}$ when calculating, they clearly would have known that this was not an *exact* value. In fact, one tablet gives the much more accurate number

$$1 \quad ; \quad 24 \quad 51 \quad 10$$

[2]Why base 60? Aside from the fact that we have ten fingers, there is really no good reason for a base ten system! Base 60 (as the Babylonians used) or base 12 (such as is the basis for much of the English system of weights and measures) makes for easier computations since these numbers have many factors, making division easier for more divisors.

This search for accuracy was certainly a step beyond the level achieved by the Egyptians.

**Geometry among the Greeks**

While both Egyptian and Babylonian mathematicians made significant contributions to the origins of geometry (and in fact, to all of mathematics), neither culture made the crucial jump from calculation to proof. There is no evidence that either culture produced a mathematics in which general relationships were explored systematically the way we do in geometry today. While rules of calculation were certainly developed and used, nothing like our system of *theorem* and *proof* is evidenced. It would remain for the Greeks to develop that approach.

*Thales of Miletus*

The man credited with first making the jump from calculation to proof is Thales of Miletus.[3] Thales lived from 624 to 548 BC (though the exactness of those dates is uncertain). He was reportedly able to travel to both Egypt[4] and Mesopotamia (the seat of the Babylonian civilization) from which he learned of their mathematical achievements. It is probable that Thales learned the Babylonians' observed geometry facts and then set to work himself to reason out *why* these apparent coincidences were true. He is reported to have thus given the first proofs that angles formed by intersecting lines are congruent, that the three angles in a triangle add to the same as two right angles, and that the two base angles in an isosceles triangle are congruent He is also credited with the first proof of the fact that now bears his name:

Theorem of Thales. *If $AB$ is diameter to a circle and if $C$ is any other point on that circle, then $\angle ACB$ is a right angle.*

What Thales actually did is uncertain. None of his work survives to the present, if in fact he ever composed any mathematical treatises. Tradition, however, ascribes to him the first mathematical proofs, and the first organization of geometry into a sequence of theorems that build on each other. For this reason, he is sometimes said to be the world's first true mathematician. This much is certain: he is the first person to whom a specific mathematical result is attributed.

[3] Thā ′ lēz of Mī · lē ′ tes – also remembered as the father of Greek philosophy.

[4] One legend claims that Thales greatly impressed the Egyptians when he calculated the height of pyramids by merely measuring the length of their shadows at the moment when a stick's shadow equaled its length.

### *Pythagoras*

Some half a century after Thales came Pythagoras of Samos (about 580-500 BC). Samos is near Thales' home of Miletus, and some legends indicate that Pythagoras may have studied under Thales. The difference in their ages, however, makes this somewhat unlikely. Like Thales, Pythagoras is said to have traveled widely — to Egypt, Mesopotamia, and possibly even to India. Also like Thales, Pythagoras is credited with first proving the famous theorem that bears his name (though again no actual writings of Pythagoras survive to the present). Unlike the practical figure of Thales, however, Pythagoras was somewhat of a religious and mystical figure. His was an important time in the development of religion; he was a contemporary of Confucius, Siddharta Guatama (Buddha), and Lao-Tzu (the founder of Taoism). If indeed he traveled to Asia, the philosophical climate he found there may well have pushed his own thinking in the direction of mysticism.

Whatever the inspiration might have been, Pythagoras founded a secret society in which mathematical pursuits were intertwined with a strict moral code, communal lifestyle, vegetarianism, and a belief in reincarnation of souls. Because the group held everything in common, including their mathematical discoveries, no result can be attributed to any one person. We likely should refer to the achievements of the *Pythagoreans*, and not only of Pythagoras himself. Intellectual achievements in mathematics and music were regarded almost religiously by the Pythagoreans, and mathematics and religion were meshed in their number mysticism, or belief in mystical qualities of certain numbers.

The phrase *"All is number"* is attributed to Pythagoras, and the beliefs of the Pythagoreans apparently included the tenet that all things in nature can be accounted for in the properties of the whole numbers and their ratios. Thus it was somewhat of a disturbance to the Pythagorean school of thought when it was discovered that the ratio between the diagonal length and side length of a square cannot be expressed as a ratio of whole numbers. (In modern language this amounts to the fact that $\sqrt{2}$ is irrational.) In a bit of historical irony, this discovery was likely made by a Pythagorean mathematician using the fact we now call the Pythagorean Theorem, though it is unlikely that the discovery occurred during the lifetime of Pythagoras. We will discuss this discovery a bit more in Section C of this chapter.

### *Euclid*

Sometime near 300 BC the man who shaped the study of Geometry more than any other stepped onto the scene. Euclid's primary accomplishment was the composition of an introductory mathematics textbook called the *Elements*. This treatise consists of 13 chapters, the first six of which treat plane geometry. Theorems are arranged to build on each other, starting from a basic set of assumptions. The assumptions are called *axioms* or *postulates* and are limited to statements Euclid believed to be indisputable, requiring no proof. This method of deducing facts from a minimal set of assumptions, called the *axiomatic method*, was not an innovation of Euclid. There had been several previous attempts to organize mathematical knowledge in this way, but the remarkable success of Euclid's work (the *Elements* became the most widely read and studied mathematical work to its time) sets it apart. The facts proved in the *Elements* were not original to Euclid — it was, after all, a textbook, not a research tract. Euclid's work is noteworthy not because of its originality, but because of its success in putting mathematical knowledge into an axiomatic framework.

The *Elements* is sometimes called the most influential textbook in history. It was studied for centuries as the standard introduction to mathematics, and its success established its author's reputation. Euclid was awarded a position at the *Museum* (or university) at Alexandria[5], the foremost center of learning at that time. Little else is known about his life and personality.

In Chapter 3 we will study the use of the axiomatic method. In the course of this discussion we will examine Euclid's *Elements* more closely. Then in Chapters 4 and 5 we will follow in Euclid's footsteps, applying the axiomatic method to derive the facts of Euclidean geometry from a small set of assumptions. Our approach will not be exactly Euclid's, but should be sufficient to give us an appreciation for his work.

## Exercises

**1.1.** True or False? (Questions for discussion)

**(a)** The ancient Egyptians were aware of several geometry facts, but never attempted to prove them.

**(b)** The Egyptian number system was superior to that of the ancient Babylonians.

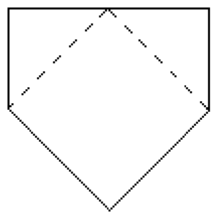
**(c)** It is not clear that the Egyptians distinguished between approximations and exact answers.

[5] Alexandria was established about that same time by Alexander the Great to be the seat of learning for his empire. The library at the Museum was unrivaled in the ancient world, and stood as the greatest repository of human knowledge until being closed by Christian authorities in 529 AD and subsequently burned by Islamic conquerors in 641 AD

**(d)** Pythagoras discovered the theorem named for him.

**(e)** Euclid invented the axiomatic method.

**1.2**. Give a geometric justification (in the spirit of Figure 1.1) for the fact that the area of a general triangle is $\frac{1}{2}ba$ where $b$ and $a$ are, respectively, the base length and altitude of the triangle. Feel free to use familiar facts about triangles. (*Hint:* It is always possible to find one altitude which lies in the triangle's interior [why?] and thus divides the triangle into two right triangles.)

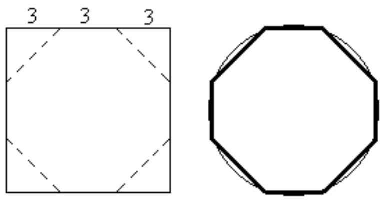


**1.3**. Inexact geometry didn't end with the Babylonian empire! Consider the following official regulation for the size and shape of home plate in Little League baseball: "*Home base shall be marked by a five-sided slab of whitened rubber. It shall be a 12-inch square with two of the corners filled in so that one edge is 17 inches long, two are 8 1/2 inches and two are 12 inches.*" Analyze the feasibility of these instructions.[6]

**1.4**. Suppose that you have a square cake frosted on all sides except the bottom. You want to cut this cake into 5 pieces so that each piece has the same amount of cake *and* the same amount of frosting. Can you find an easy way to do this? Does your method generalize to cutting the cake into $n$ pieces of equal size and equal frosting?

**1.5**. How accurate is the Babylonian approximation $\sqrt{2} \approx 1\ ;\ 24\ 51\ 10$ ? Find a Babylonian number that more closely approximates $\sqrt{2}$.

**1.6**. What area does the Egyptian formula from Figure 1.3 predict for a circle of radius 10?




**1.7**. One possible derivation of the Egyptian rule for circle areas in Figure 1.3 is based on the octagon shown above. (Each side of a square of side length

[6]See *Building Home Plate: Field of Dreams or Reality?* by M.J. Bradley, Mathematics Magazine **69** p.44-45 [Feb. 1996].

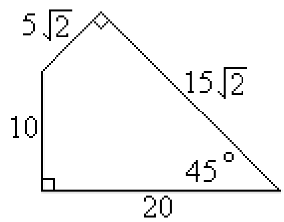
9 has been trisected, then the four corners removed using the trisection points to determine the cuts.) This octagon is visually close in area to the circle of diameter 9.

**(a)** Complete the justification for the approximation by showing that the area of the octagon is close to the area of a square of side length 8.

**(b)** What approximation for the value of $\pi$ is implied by equating the areas of a circle with diameter $a$ and a square of side length $b$?

**(c)** Can you find integers $a$ and $b$ so that using a circle of diameter $a$ and a square of side length $b$ gives a better approximation to $\pi$ than does $a = 9$ and $b = 8$?

**1.8**. What area does the formula from Figure 1.2 predict for the quadrilateral shown at right? What is the actual area of this quadrilateral?




**1.9**. Consider the incorrect formula illustrated in Figure 1.2.

**(a)** Use this formula to derive the (also incorrect) formula that the area of a triangle with two sides of length $a$ and $b$ is $\frac{1}{8}(a+b)^2$.

**(b)** Explain why it is clear that this formula for a triangle's area cannot possibly be correct for *all* triangles.

**(c)** Find a triangle for which the formula *does* give the correct area.

**1.10**. Early Greek mathematicians knew how to compute the areas of polygons by dissecting them into triangles. So, a natural method for computing the area of a circle was to approximate the circle by an $n$-gon. Specifically, when $n$ is large enough the circle looks quite similar to the $n$-gon created by connecting (in order) $n$ points evenly spaced around the circle.

**(a)** Show that the area of the $n$-gon created by connecting $n$ equally spaced points around a radius $r$ circle is $\frac{r^2 n}{2} \sin\frac{2\pi}{n}$.

**(b)** Show that the limit of this area as $n$ tends to infinity is $\pi r^2$.

**1.11**. Our brief historical survey has left out many important topics and characters (including *everything* after Euclid). Research an item from the following list and write a short report describing its history and significance to the development of mathematics.

- the golden section
- the Pythagorean pentagram
- Hippocrates of Chios (*not* Hippocrates of Cos)
- Hippias of Elis
- Algebra among the Greeks
- Eudoxus of Cnidus
- Archimedes of Syracuse
- Apollonius of Perga
- Aristarchus of Samos
- Eratosthenes of Cyrene
- Mathematics in ancient India
- Mathematics in ancient China

## B. The "Basic Toolbox" of Euclidean Geometry

A student's first exposure to geometry is usually in high school. The content of that course varies, but we will catalog here a set of basic facts that would be included in nearly all such treatments. These facts constitute the fundamental tools of plane Euclidean geometry, and we will henceforth refer to them as our "*basic toolbox theorems*". Your high school course may or may not have included proofs of them – we'll put off our own proofs until Chapters 4 and 5 (where we will show how they can be derived from a simple set of axioms). But shortly in Chapter 2 we will use these toolbox facts to prove some rather remarkable theorems. You might just be surprised at the power of what you learned in that high school class!

We want to avoid too much formality at this point. But though most of the terms we use in listing these basic facts should already be familiar, we will nonetheless give some of their definitions. First, we want to be certain there is no ambiguity in what our theorems say or in what the exercises ask you to demonstrate. But also, the formulation of exact definitions is part of any mathematical endeavor. We will be more particular with our definitions once we begin our axiomatic treatment in Chapter 4, but it is not too early now to warm up to the process.

Terms we will not define here include *line, segment, ray, parallel, perpendicular, angle, angle measure, distance, length, region* and *area.* Nor will we define *congruence* for segments or angles, instead relying on the intuitive notion that congruent segments have equal length and congruent angles have equal angle measure. Throughout, we'll use the notational conventions set forth in the introduction.

## Angles

Our first group of basic facts concerns properties of angles. We begin with some definitions.

**DEFINITIONS.** An angle measuring 90° is called a *right angle*. An angle measuring less than 90° is called an *acute* angle, and an angle measuring more than 90° is called an *obtuse* angle.

**DEFINITION.** In Figure 1.4 the marked angles are called *vertical angles* of the line intersection.

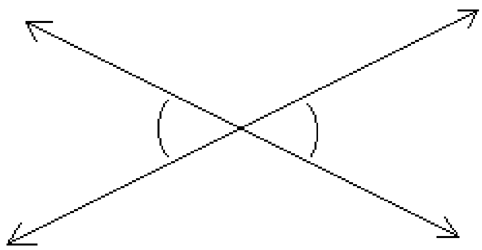


Figure 1.4: Vertical angles

**DEFINITION.** A *transversal* is a set of lines, one of which crosses all of the others. Figure 1.5 shows a transversal of two lines $\Lambda_1$ and $\Lambda_2$ by a line $\Lambda_3$. The angles labeled $\alpha$ and $\beta$ in this transversal are called *alternate interior angles* (because they lie on alternate sides of the transversing lines and between the other two lines). The angles labeled $\alpha$ and $\gamma$ are called *corresponding angles* of the transversal. (Note that they lie on the same side of the transversing line, with only one of them between the other two lines.)

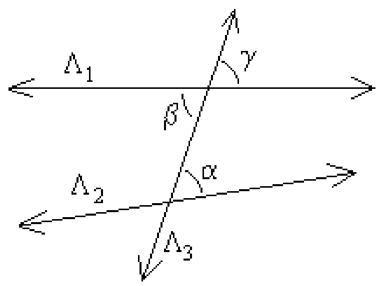



Figure 1.5: A transversal of two lines

The first three of our toolbox facts can now be stated. Note that each is an "if and only if" statement – if you are rusty on that concept we will brush up on it in the next section.

**VERTICAL ANGLES THEOREM.** *The vertical angles of an intersection of lines are congruent. In fact, if $C$ is a point on line $\overleftrightarrow{AB}$ between $A$ and $B$ and if points $D$ and $E$ are on opposite sides of that line, then $\angle ACD \cong \angle BCE$ if and only if $C$ lies on the line $\overleftrightarrow{DE}$ (see Figure 1.6).*

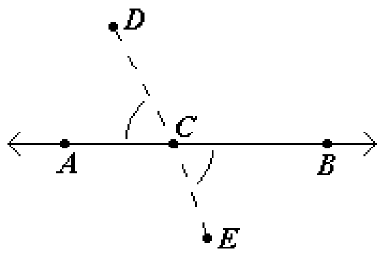



Figure 1.6:

**ALTERNATE INTERIOR ANGLES THEOREM.** *If lines $\Lambda_1$ and $\Lambda_2$ are transversed by a line $\Lambda_3$ (as above in Figure 1.5) then $\Lambda_1$ and $\Lambda_2$ are parallel if and only if the transversal has a congruent pair of alternate interior angles.*

**Corresponding Angles Theorem.** *If lines $\Lambda_1$ and $\Lambda_2$ are transversed by a line $\Lambda_3$ (as above in Figure 1.5) then $\Lambda_1$ and $\Lambda_2$ are parallel if and only if the transversal has a congruent pair of corresponding angles.*

The next basic fact describes an important property of right angles (via perpendicular lines). First, some definitions:

**Definitions.** Let $AB$ be a segment. The *midpoint* of $AB$ is the point $M$ on that segment such that $|AM| = |MB|$. The *perpendicular bisector* of $AB$ is the line through this point $M$ that is perpendicular to $AB$.

**Perpenicular Bisector Theorem.** *The perpendicular bisector of the segment $AB$ is exactly the set of points that are* equidistant *from $A$ and $B$. That is, $C$ is on the perpendicular bisector of $AB$ if and only if $|AC| = |BC|$. (In set theory notation, the perpendicular bisector of $AB$ is the set $\{C : |AC| = |BC|\}$.)*

Also relating to right angles is the notion of distance from a point to a line:

**Definition.** Let $\Lambda$ be a line and $P$ a point. The *distance from $\Lambda$ to $P$* is $|PQ|$ where $Q$ is the point on $\Lambda$ such that $PQ \perp \Lambda$.

## Polygons

**Definitions.** The *polygon* or *$n$-gon* $P_1P_2\cdots P_n$ is the union of the segments $P_1P_2, P_2P_3, \ldots, P_{n-1}P_n, P_nP_1$.

- The points $P_1, P_2, \ldots, P_n$ are called the *vertices* of the polygon. By convention we take the indices on the vertices modulo $n$ so that $P_0 = P_n$ and $P_{n+1} = P_1$.
- The segments $P_iP_{i+1}$ $(1 \leq i \leq n)$ are the *sides* of the polygon.
- The angles $\angle P_i = \angle P_{i-1}P_iP_{i+1}$ are the *angles* of the polygon.
- The *perimeter* of the polygon is $|P_1P_2|+|P_2P_3|+\cdots+|P_{n-1}P_n|+|P_nP_1|$.

- A segment $P_iP_j$ that is not a side of the polygon is called a *diagonal* of the polygon.
- $P_1P_2 \cdots P_n$ is a *simple polygon* if the points $P_1, P_2, \ldots, P_n$ are distinct and if no two sides intersect except at a shared vertex.
- $P_1P_2 \cdots P_n$ is a *convex polygon* if for $1 \leq i \leq n$ all vertices except $P_i$ and $P_{i+1}$ lie on the same side of the line $\overleftrightarrow{P_iP_{i+1}}$.
- A simple 3-gon is (of course) a triangle.
- A simple 4-gon is a *quadrilateral*.

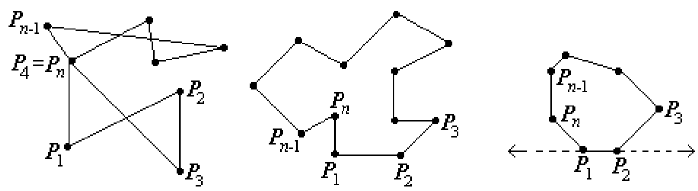



Figure 1.7: Left to right: a polygon, a simple polygon, and a convex polygon

While general polygons will come into some of our discussions (particularly in some of the exercises), triangles and quadrilaterals will be by far the most important types to us, and a large part of our toolbox will be devoted to facts about them. We begin with triangles.

## Triangles

You might recall from your previous experience with geometry that many geometry proofs involve showing the congruence of triangles. Some of the best-remembered facts from a first exposure to geometry are usually the various congruence criteria such as "side-angle-side". We'll certainly put these in our toolbox, but first let's set out the definition of triangle congruence.

**Definition.** We say that triangles $\triangle ABC$ and $\triangle DEF$ are *congruent* (written $\triangle ABC \cong \triangle DEF$) if the following congruences of their sides and angles are all true.

- $AB \cong DE$, $BC \cong EF$, and $CA \cong FD$.
- $\angle A \cong \angle D$, $\angle B \cong \angle E$, and $\angle C \cong \angle F$.

The congruent pairs of sides and pairs of angles listed in this definition are called *corresponding parts* of the triangles – thus the phrase "*corresponding parts of congruent triangles are congruent*". That is, quite simply, why triangle congruence is at the heart of so many geometry proofs: a wealth of information comes from proving a pair of triangles to be congruent.

A few words on notation are in order here. Ordinarily we will name a triangle merely by listing its vertices, such as $ABC$. However, when we wish to specify a congruence of triangles, we need to set a specific order to the vertices (for notice that corresponding parts are identified by their positions in the name of the triangle). To distinguish this "ordered" triangle name from the ordinary unordered one, we will use the prefix $\triangle$. Thus the names $ABC$ and $ACB$ denote the same triangle (as a set of points), but $\triangle ABC$ and $\triangle ACB$ have different meanings. In fact, these last two are congruent only if $\angle B \cong \angle C$ and $AB \cong AC$.

The methods to show congruence of triangles are supplied by those familiar congruence criteria. We list them here for inclusion in our toolbox.

**Triangle Congruence Criteria.** *The two triangles $\triangle ABC$ and $\triangle DEF$ are congruent if any one of the following conditions are met (see Figure 1.8).*

**SAS criterion:** $CA \cong FD$, $\angle A \cong \angle D$, *and* $AB \cong DE$

**ASA criterion:** $\angle A \cong \angle D$, $AB \cong DE$, *and* $\angle B \cong \angle E$

**SAA criterion:** $AB \cong DE$, $\angle B \cong \angle E$, *and* $\angle C \cong \angle F$

**SSS criterion:** $AB \cong DE$, $BC \cong EF$, *and* $CA \cong FD$

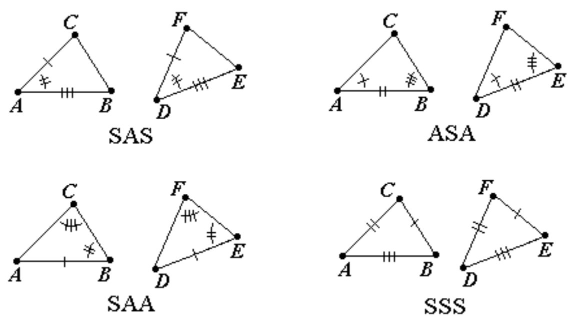



Figure 1.8: The four triangle congruence criteria

A few special varieties of triangles are important enough to give names to. The following definitions should be familiar (see Figure 1.9.)

**Definitions.** A triangle $ABC$ is

- an *equilateral* triangle if all three sides have the same length.
- an *isosceles* triangle if $|AB| = |AC|$. In this case, $A$ is called the *top vertex* and $BC$ the *base* of the triangle. The angles $\angle B$ and $\angle C$ are called the *base angles.*
- a *right* triangle if $\angle C$ is a right angle. In this case, $AC$ and $BC$ are called the *legs* and $AB$ the *hypotenuse* of the right triangle.

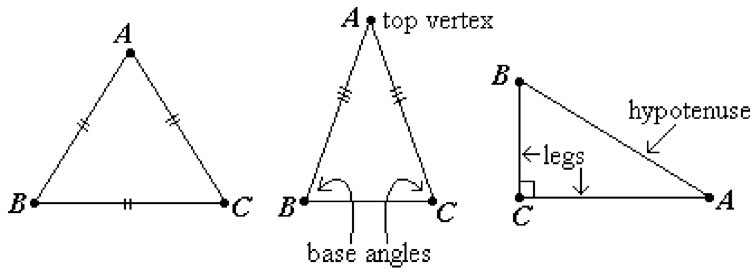


Figure 1.9: equilateral, isosceles, and right triangles

**Isosceles Triangle Theorem.** *Triangle $ABC$ is isosceles with top vertex $A$ if and only if $\angle B \cong \angle C$.*

The next basic fact needs no introduction. It is (deservedly!) one of the most famous of all results in geometry.

**Pythagorean Theorem.** *If $ABC$ is a right triangle with right angle $\angle C$ then $a^2 + b^2 = c^2$.*

Another famous fact about triangles, and one of the most useful, is the following.

**180° Sum Theorem.** *The measures of the three angles of any triangle sum to* 180°.

The concept of similarity of triangles is a surprisingly powerful tool for proving interesting geometric theorems. We give a definition of similarity to refresh your memory.

**Definition.** If $\triangle ABC$ and $\triangle DEF$ are triangles such that $\angle A \cong \angle D$, $\angle B \cong \angle E$, and $\angle C \cong \angle F$, then we say the triangles are *similar* and write $\triangle ABC \sim \triangle DEF$.

Just like congruence, there are criteria for determining similarity of triangles. And the payoff for knowing two triangles are similar is that the ratios of their corresponding side lengths are equal. The most useful facts on similarity are summarized in our next toolbox theorem.

**Similar Triangles Theorem.** *Let $\triangle ABC$ and $\triangle PQR$ be two triangles. Then*

- $\triangle ABC \sim \triangle PQR$ *if and only if* $p/a = q/b = r/c$.
- $\triangle ABC \sim \triangle PQR$ *if and only if* $q/b = r/c$ *and* $\angle A \cong \angle P$.

The first part of this theorem is the most frequently used, and (though technically an "if and only if" fact) is almost always used to say that once similarity of triangles is known, then the ratios of corresponding side lengths will be equal. The second part is a sort of "side-angle-side similarity criterion", and is used to give an alternative way of showing two triangles are similar: instead of showing all three pairs of corresponding angles are congruent (as the definition of similarity requires), we can show one pair of corresponding angles and the equality of ratios of lengths for the corresponding pairs of sides bounding those angles.

There are many interesting objects derived from triangles. The ones listed in the definition below come up in several of our results and exercises.

**Definitions.** Let $ABC$ be a triangle.

- The *median* of $ABC$ at vertex $A$ (or side $BC$) is the segment $AP$ where $P$ is the midpoint of $BC$.
- The *bisector* of $\angle A$ is the segment $AP$ (or ray $\overrightarrow{AP}$) where $P$ is the point on $BC$ so that $m\angle BAP = m\angle PAC$.
- The *altitude* of $ABC$ at vertex $A$ (or relative to side $BC$) is the segment $AP$ where $P$ is the point of $\overleftrightarrow{BC}$ so that $AP \perp \overleftrightarrow{BC}$. The distance $|AP|$ is then called the *height* of triangle $ABC$ relative to the *base* $BC$. (Note that this is just the distance from $A$ to line $\overleftrightarrow{BC}$.)

## Quadrilaterals

A few types of quadrilaterals become important in the study of geometry. Most everyone is familiar with the names of these quadrilaterals, but there is often confusion as to their exact definitions.

**Definitions.** A quadrilateral $ABCD$ is

- a *trapezoid* if $AB \; // \; CD$.
- a *parallelogram* if $AB \parallel CD$ and $BC \parallel DA$.
- a *rectangle* if all of its angles are right angles.
- a *rhombus* if all of its sides have equal length.
- a *square* if it is both a rectangle and a rhombus.

The following fact relating some of these quadrilateral types is important enough to put in our toolbox.

**Parallelogram Theorem.**

- *Every rectangle is a parallelogram and every rhombus is a parallelogram.*
- *If $ABCD$ is a parallelogram then $\angle A \cong \angle C$, $\angle B \cong \angle D$, $|AB| = |CD|$, and $|BC| = |DA|$. (That is, in a parallelogram opposite angles are congruent and opposite sides have equal length.)*

## Area

Area computations lie at the heart of some of the proofs we will consider. Our toolbox needs to include facts sufficient to allow us to calculate the area of simple figures. As usual, we begin with the necessary definitions.

**Definitions.** By the *interior* of the simple polygon $P_1P_2 \cdots P_n$ we mean the region of the plane bounded by $P_1P_2 \cdots P_n$. We will often say "*the area of [some polygon]*" to mean the area of the interior of that polygon.

**Area Formulae.** *(See Figure 1.10.)*

- *The area of triangle $ABC$ is $\frac{h}{2}|AB|$ where $h$ is the height of $ABC$ relative to base $AB$.*
- *The area of parallelogram $DEFG$ is equal to $h|DE|$ where $h$ is the height of $DEF$ relative to base $DE$. ($h$ is also called the* height *of parallelogram $DEFG$ relative to $DE$.)*

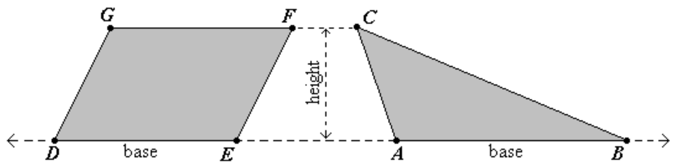



Figure 1.10:

## Circles

The circle is one of the basic building blocks in geometric figures. There are a host of defined terms involving circles that we wish to review here. These terms are probably familiar, but will be used in the results and exercises from here on. So take a moment to read them and refresh your memory.

**Definitions.** Let $C$ be a point and $r > 0$ a real number. The *circle* with *center* $C$ and *radius* $r$ is the locus of points $\Gamma = \{P : |PC| = r\}$. We say that a point $Q$ is *inside* $\Gamma$ if $|QC| < r$ and *outside* $\Gamma$ if $|QC| > r$.

**Definition.** Let $\Gamma$ be a circle with center $C$. A line or segment is *tangent* to the circle $\Gamma$ if it contains a point $A$ of $\Gamma$ and is perpendicular to $AC$.

**Definition.** We say that the simple polygon $P_1P_2\cdots P_n$ is *inscribed* in the circle $\Gamma$ (or that $\Gamma$ *circumscribes* $P_1P_2\cdots P_n$) if all vertices of $P_1P_2\cdots P_n$ lie on $\Gamma$.

**Definition.** We say that the circle $\Gamma$ is *inscribed* in the simple polygon $P_1P_2\cdots P_n$ (or that $P_1P_2\cdots P_n$ *circumscribes* $\Gamma$) if each side of $P_1P_2\cdots P_n$ is tangent to $\Gamma$.

**Definitions.** Refer to Figure 1.11 in the following items. Let $\Gamma$ be a circle with center $C$ and let $A$ and $B$ be points on $\Gamma$.

- The segment $AB$ is called a *chord* of $\Gamma$.
- The two arcs of $\Gamma$ lying respectively on either side of the line $\overleftrightarrow{AB}$ are called the arcs *determined* by the chord $AB$.
- If $P$ is a point on one of these arcs, that arc may be denoted by the symbol $A\overset{\frown}{P}B$.
- If the center $C$ is on $AB$ then $AB$ is called a *diameter* of $\Gamma$ and the segment $CA$ (and likewise $CB$) is called a *radius* of $\Gamma$. The two arcs of $\Gamma$ determined by a diameter are called *semicircles*.
- If the center $C$ is not on $AB$ then the arc of $\Gamma$ on the same side of $\overleftrightarrow{AB}$ as $C$ is called the *major arc* determined by $AB$, and the arc on the other side is called the *minor arc* determined by $AB$.
- Each arc $A\overset{\frown}{P}B$ has a *measure* between $0°$ and $360°$ which we will denote by $m(A\overset{\frown}{P}B)$. Note that the measures of the major and minor arcs determined by a chord will always add to $360°$.
- If $Q$ is a point of $\Gamma$ not on the arc $A\overset{\frown}{P}B$ then the angle $\angle AQB$ is called an *inscribed angle* for the arc $A\overset{\frown}{P}B$.

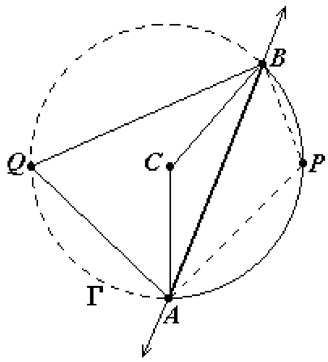



Figure 1.11: The major arc $\overset{\frown}{AQB}$ (dashed) and minor arc $\overset{\frown}{APB}$ (solid) determined by chord $AB$. $\angle AQB$ is an inscribed angle for the arc $\overset{\frown}{APB}$, and $\angle APB$ is an inscribed angle for the arc $\overset{\frown}{AQB}$.

After these many definitions, our toolbox will include only one fact involving circles, but its usefulness can hardly be overstated.

**Inscribed Angle Theorem.** *The measure of a arc is twice the measure of any inscribed angle for that arc.*

Thus, in Figure 1.11 above, $m(\overset{\frown}{APB}) = 2m\angle AQB$ and $m(\overset{\frown}{AQB}) = 2\angle APB$. Note that one consequence of the Inscribed Angle Thoerem is that any two inscribed angles for the same arc are congruent. Notice also that the Theorem of Thales described in Section A (see p.5) is a special case of the Inscribed Angle Theorem – just take the arc to be a semicircle. In that case, the arc measures 180° so any inscribed angle must measure 90° and is thus a right angle.

This completes our basic toolbox for Euclidean plane geometry. It is by no means a complete catalog of theorems – you're probably aware of several geometry facts we have not included here. But what we have included gives enough tools to prove some really spectacular things, as we'll see in Chapter 2. Some of the homework problems in this section will give you practice at using these toolbox theorems by doing some routine geometric calculations. We'll close the section with an example of this.

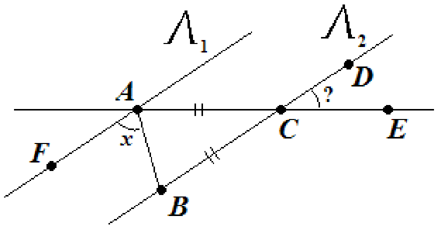
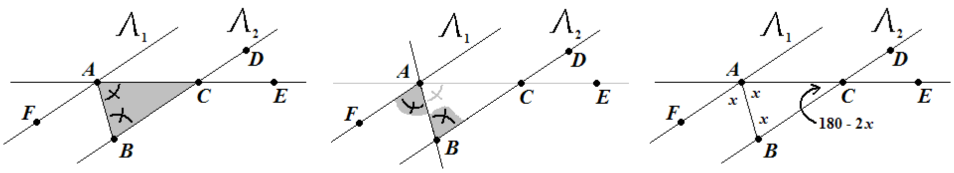



Figure 1.12:

**Example 1.1.** *In Figure 1.12 above, lines $\Lambda_1$ and $\Lambda_2$ are parallel and $AC \cong BC$. If $m\angle FAB = x$, calculate $m\angle DCE$.*

*Solution:* Note how many of our toolbox theorems we use in the computation steps below!




- First, from what we are given, we know that triangle $ABC$ is isosceles with top vertex $C$. Thus, by the Isosceles Triangle Theorem, its base angles $\angle CAB$ and $\angle CBA$ are congruent (see leftmost figure above).

- But now we have line $\overleftrightarrow{AB}$ transversing the parallel lines $\Lambda_1$ and $\Lambda_2$, so the alternate interior angles $\angle CBA$ and $\angle FAB$ are congruent by the Alternate Interior Angle Theorem (see middle figure above).

- Putting together these facts, we see that $x = m\angle FAB = m\angle CBA = m\angle CAB$ (see rightmost figure above).

- Now, applying the 180° Sum Theorem, we conclude that $m\angle ACB = 180° - m\angle CAB - m\angle CBA = 180° - 2x$ (again, see rightmost figure above).

- Finally, using the Vertical Angles Theorem we can see that $m\angle DCE = m\angle ACB$, so $m\angle DCE$ is also $180° - 2x$.

$\square$

## Exercises

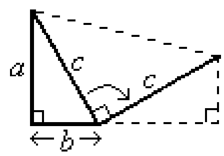
**1.12**. Show how the Corresponding Angles Theorem can be proved using the Vertical Angles Theorem and Alternate Interior Angles Theorem.

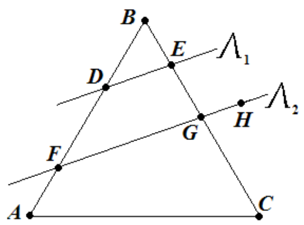
**1.13**. Show how the 180° Sum Theorem can be proved from the Alternate Interior Angles Theorem. (*Hint:* construct the line through $C$ parallel to line $\overleftrightarrow{AB}$.)

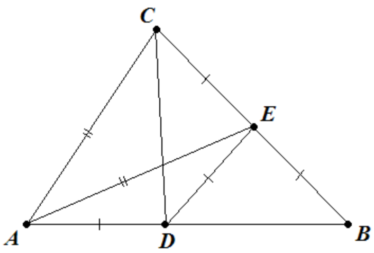
**1.14**. Use the other "toolbox theorems" to prove the part of the Parallelogram Theorem that says *opposite angles in a parallelogram are congruent*.

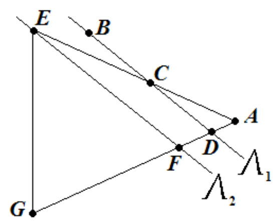
**1.15**. State and prove an area formula for trapezoids.

**1.16**. The trapezoid shown at right is constructed by starting with a right triangle and rotating its hypotenuse 90° outward from one of its endpoints. Based on this figure, give a proof of the Pythagorean Theorem using only the other basic tools. (*Hint:* compute the area of the trapezoid in two different ways.)



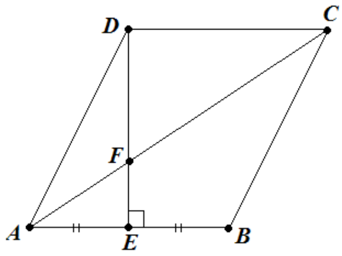

**1.17**. In the leftmost figure above, triangle $ABC$ is equilateral, lines $\Lambda_1$ and $\Lambda_2$ are parallel, and $m\angle BDE = 35°$. Compute $m\angle HGC$.

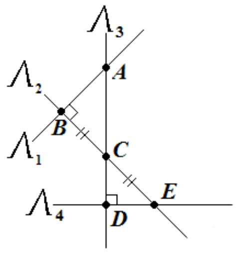
**1.18**. In the middle figure above, $AD \cong DE \cong BE \cong CE$, $AE \cong AC$, and $m\angle B = x$.

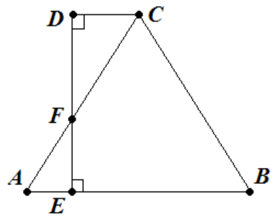
**(a)** Compute $m\angle ACB$.

**(b)** Compute $m\angle ACD$.

**1.19**. In the rightmost figure above, $\Lambda_1 \parallel \Lambda_2$, triangle $AEG$ is isosceles with top vertex $A$, and triangle $EFG$ is isosceles with top vertex $E$. If $m\angle EGF = x$, what is $m\angle BCA$?



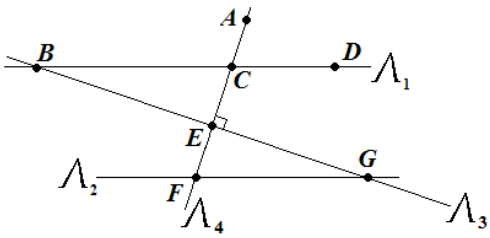



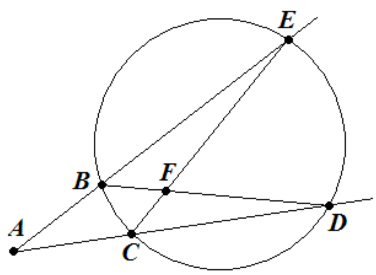




**1.20**. In the leftmost figure above, $ABCD$ is a parallelogram, $DE \perp AB$, $E$ is the midpoint of $AB$, $|AB| = 6$, and $|BC| = 10$. Compute area$(AEF)$.

**1.21**. In the middle figure above, $\Lambda_1 \perp \Lambda_2$, $\Lambda_3 \perp \Lambda_4$, $C$ is the midpoint of $BE$, $|AC| = 2|CD|$, and $|AB| = 3$. Compute $|DE|$.

**1.22**. In the rightmost figure above, triangle $ABC$ is equilateral with $|AE| = 3$ and $|CD| = 4$. Compute $|EB|$.




**1.23**. In the figure at left above, $\Lambda_1 \parallel \Lambda_2$, $\Lambda_3 \perp \Lambda_4$, and $m\angle EGF = 25°$. Compute $m\angle ACD$.

**1.24**. In the figure at right above, $|CE| = 8$, $|BD| = 5$, $|AB| = 3$, and $|AD| = 10$.

**(a)** Compute $|AC|$.

**(b)** Compute $|BE|$.

# C. Logic and Deductive Reasoning

When the Greeks introduced deduction and proof to mathematics they ignited a fire that would drive mathematical study from that day to the present. The importance of proof in mathematics is evident to any student who ventures beyond the rudiments of the subject. In this section we will learn some terminology from the study of deduction and proof, and we will discuss some of the most important methods of proof available to students of geometry. Depending on your previous experience with proof, you may need only to skim this section.

**Implications, contrapositives, converses, and equivalences**

The simplest kind of logical statement is an assertion of the form

*If condition* X *holds then condition* Y *also holds.*

We call such a statement an *implication* because it may be rephrased as "X *implies* Y." Condition X is called the *hypothesis* and condition Y is called the *conclusion.* Using shorthand notation we may write simply $\mathrm{X} \Longrightarrow \mathrm{Y}$.

Many theorems in mathematics may be phrased as implications. Examine, for instance, the triangle congruence criteria from the last section. Each is phrased in the form of a statement "*If* [something] *then* [something else]." Other examples of implications are less mathematical: "*If it is raining then the sidewalk is wet*" and "*If Rex is a dog then Rex has four legs*" are possibilities.

Variations on this are easy. For instance, if we let the symbol $\sim$X denote "*condition* X *does not hold*" then we could write $\sim\mathrm{Y} \Longrightarrow \sim\mathrm{X}$ for the statement

*If condition* Y *does not hold then condition* X *does not hold.*

You may notice that this last statement is *logically equivalent* to the original statement. That is, $\mathrm{X} \Longrightarrow \mathrm{Y}$ is a true statement *if and only if* $\sim\mathrm{Y} \Longrightarrow \sim\mathrm{X}$ is a true statement. (Both statements essentially say that condition Y must hold if condition X holds.) We say that $\mathrm{X} \Longrightarrow \mathrm{Y}$ is the *direct statement* of the implication and that $\sim\mathrm{Y} \Longrightarrow \sim\mathrm{X}$ is its *contrapositive statement.*

Any theorem in the form of a simple implication can be phrased in either its direct form or its contrapositive form. For example, the part of the Parallelogram Theorem (p.18) that says "*every rectangle is a parallelogram*" could be phrased as the implication "*If $ABCD$ is a rectangle then $ABCD$ is a parallelogram.*" Its contrapositive form is then "*If $ABCD$ is not a parallelogram then $ABCD$ cannot possibly be a rectangle.*" Sometimes there is a specific advantage (clarity or ease of proof) to phrasing a theorem in one or the other of these forms.

In addition to the contrapositive statement, there is another variation on X $\Longrightarrow$ Y that is important to consider, namely the *converse statement* Y $\Longrightarrow$ X. Unlike the contrapositive statement, **the converse statement is *not* logically equivalent to the direct statement!** In fact, our simple theorem "*If $ABCD$ is a rectangle then $ABCD$ is a parallelogram*" has a false converse, namely "*If $ABCD$ is a parallelogram then $ABCD$ is a rectangle*".

There is often some care required in giving a good statement of the converse or contrapositive to a theorem. Take the Theorem of Thales, for example. The original statement is "*If $A$, $B$, and $C$ are points on a circle $\Gamma$ so that $AB$ is a diameter of $\Gamma$ then $\angle C$ is a right angle.*" This certainly is worded in the form "*if* X *then* Y" so it shouldn't be difficult to give either contrapositive or converse. And, really, it isn't. But we should recognize the words "*If $A$, $B$, and $C$ are points on a circle $\Gamma$ ...*" as the *setting* for the theorem and not as part of either X or Y. Our statement has the form "*if* X *then* Y" where X is "*$AB$ is a diameter of $\Gamma$*" and Y is "*$\angle C$ is a right angle*", both in the setting of $A$, $B$, and $C$ being points on the circle $\Gamma$. An elegant statement of contrapositive or converse will establish the setting first, just as the direct statement did. Thus:

**Contrapositive:** *If $A$, $B$, and $C$ are points on a circle $\Gamma$ and $\angle ACB$ is not a right angle, then $AB$ is not a diameter of $\Gamma$.*

**Converse:** *If $A$, $B$, and $C$ are points on a circle $\Gamma$ and $\angle ACB$ is a right angle then $AB$ is a diameter of $\Gamma$.*

As pointed out above, the converse of a true statement need not be a true statement itself. In some cases, however, both the direct statement and its converse may be true. This is in fact the case with Thales' Theorem – its converse is also a theorem in geometry (which you are asked to prove in Exercise 1.37). Now X $\Longrightarrow$ Y can be rephrased "X *is true only if* Y *is also true*" whereas Y $\Longrightarrow$ X can be stated as "X *is true if* Y *is true*". So we may indicate both the direct statement and the converse with the single phrase "X *if and only if* Y". This is written in symbols as X $\Longleftrightarrow$ Y. Such a statement is called an *equivalence*, for it demonstrates that X and Y are *logically equivalent* conditions – if one of them holds then so must the other. Thales' Theorem and its converse may be stated together as the single statement

> *Let $A$, $B$ and $C$ be points on a circle $\Gamma$. Then $AB$ is a diameter of $\Gamma$ if and only if $\angle ACB$ is a right angle.*

Equivalences are very important in mathematics, and are easily identified by the phrase *if and only if*. You should recognize the Alternate Interior Angles Theorem, Corresponding Angles Theorem, Perpendicular Bisector Theorem, Isosceles Triangle Theorem, and Similar Triangles Theorem from Section B as equivalences. As we pointed out above, Thales' Theorem may be turned into an equivalence, but tradition applies the name of Thales only to the one implication. The situation is the same with the Pythagorean Theorem, which also has a mathematically correct converse.

Proving an equivalence always involves two steps, for an equivalence is really two different implications combined into a single statement. To prove X $\Longleftrightarrow$ Y we would first prove X $\Longrightarrow$ Y (by assuming X as the hypothesis and deriving the conclusion Y) then prove Y $\Longrightarrow$ X (by assuming Y as the hypothesis and deriving X). We will get plenty of practice at this in the coming chapters!

**Direct Proofs**

Often we phrase our theorems as statements involving quantifiers like *"all", "every", "some", "none"*, or *"at least one"*. Thus, instead of "*If $ABCD$ is a rectangle then $ABCD$ is a parallelogram*" we might say "*Every rectangle is a parallelogram.*" In setting out to prove or disprove such a statement, it is important to consider the type of quantifiers involved.

*Universal quantifiers* are words like *"all"* or *"none"*. They imply some condition (or lack of a condition) for every item in some class. To prove a statement involving universal quantifiers we must show the conclusion of the statement holds universally for *every* member of the class under consideration. It is not good enough to show that some specific rectangle is a parallelogram – we must somehow show that *every* rectangle is a parallelogram.

But how can we possibly examine and verify a property (such as having pairs of opposite sides parallel to each other) for *every* member of an infinite class of objects (such as the class of all rectangles)? This is the essential step not taken by the Egyptians and Babylonians: the abstraction from considering specific examples to a universal examination of all members of some class. The needed innovation was the concept of a *generic representative.* Generic representatives present a way of examining not just one particular example, but rather *every* example of a certain type of object *simultaneously.* To prove that every rectangle is a parallelogram we would begin not with a specific rectangle (such as one whose side lengths are specified), but rather with a generic representative for the class of all rectangles. We might call this representative $ABCD$, and because it must be generic for the entire class, the only facts we could assume about it would be those common to all rectangles – namely that each of the angles $\angle A$, $\angle B$, $\angle C$, and $\angle D$ are right angles. Proofs using generic representatives usually begin with the phrase "*Let* _____ *be a* _____" – in the case of proving every rectangle is a parallelogram it would be "*Let $ABCD$ be a rectangle*". The proof would be completed by using only the generic properties of a rectangle to show that $ABCD$ also satisfies the definition of a parallelogram. Thus, in this case we would need to prove that $AB$ is parallel to $CD$ and $BC$ is parallel to $AD$. This would establish that every member of the class of rectangles is also a parallelogram – exactly what is claimed in the statement we are trying to prove.

Of course, to *disprove* a statement involving universal quantifiers is much easier. Consider for example the (false) statement: *"Every parallelogram is a rectangle"*. To disprove this statement we need only display a specific example of a parallelogram which is *not* a rectangle. An example used in this way to disprove a statement is called a *counterexample* to the statement. In summary, examples can disprove a universal statement, but never prove one.

The situation is quite opposite for statements involving *existential quantifiers* like *"some"* or *"at least one"*. Here, an example is exactly what is needed for proof. To prove the statement *"Some parallelograms are rectangles"* is merely to produce an example of a parallelogram all of whose angles are right angles. On the other hand, disproving an existential statement calls for a universal argument using a generic representative. For to disprove the statement "*There exists an object of type X such that condition Y holds*" we need an argument that is valid for *every* object of type X showing that *none* of them satisfy the condition Y. Thus, we would start with a generic representative for the objects of type X and proceed to show that condition Y fails to hold for this object.

### Writing Proofs

This is a good time to discuss how a proof can be written down. A proof is basically a sequence of statements, beginning with the hypothesis of the theorem and ending with the conclusion of the theorem, such that each statement is justified by the statements preceding it or other known facts. Frequently beginning geometry students are taught to write these proofs in a "T diagram" format. Here, a statement is placed on the left side of the "T" opposite its justification.

It is important that every step be justified! When you write a proof, think of yourself as an attorney building your case before a jury (your reader). Each statement you make in your case must be supported by some witness or piece of evidence. Previously proved facts, definitions, and given hypotheses are all valid witnesses, whereas a statement you have not yet proved cannot be called to testify. Any claim you make that is *not* backed up by the testimony of a valid witness has the effect of invalidating the entire proof in the eyes of your jury.

Here is a simple example in which we prove a well-known geometric fact.

**THEOREM.** *Opposite sides of a parallelogram have equal lengths.*

*Proof:* Let $ABCD$ be a parallelogram. We will prove that $|AB| = |CD|$ and $|BC| = |AD|$. (See Figure 1.13.)

| | |
|---|---|
| $AB \parallel CD$ | Definition of parallelogram |
| $\angle ABD \cong \angle CDB$ | Alternate Interior Angles Theorem |
| | ($BD$ transverses the parallel lines $\overleftrightarrow{AB}$ and $\overleftrightarrow{CD}$) |
| $BC \parallel AD$ | Definition of parallelogram |
| $\angle ADB \cong \angle CBD$ | Alternate Interior Angles Theorem |
| | ($BD$ transverses the parallel lines $\overleftrightarrow{BC}$ and $\overleftrightarrow{DA}$) |
| $\|BD\| = \|DB\|$ | Trivial |
| $\triangle ABD \cong \triangle CDB$ | ASA Congruence Criterion (see lines above) |
| $\|AB\| = \|CD\|$ and $\|AD\| = \|CB\|$ | Corresponding parts of congruent triangles |

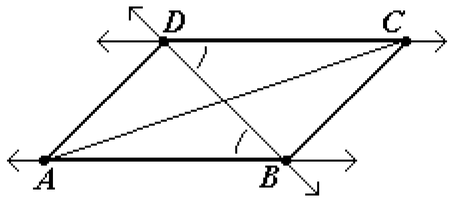



Figure 1.13:

This is a valid proof (provided we are allowing ourselves to use the Alternate Interior Angles Theorem and ASA Congruence Criterion as witnesses), and there is nothing technically wrong with writing proofs in this format. But it isn't much of a leap from here to the more pleasing format of using complete sentences. In practice, a "T diagram" provides a good way to sketch out and organize a proof. Then, once it is in this form, each line of the "T" can simply be turned into a sentence of the proof. We could of course organize these sentences into paragraphs. But to emphasize the connection with the "T diagram" (as well as to add clarity and organization) we will usually itemize our sentences in a list. For example, the above "T diagram" proof can easily be translated as follows:

*Proof:* Let $ABCD$ be a parallelogram. We will show that $|AB| = |CD|$ and $|BC| = |AD|$.

- We know $AB \parallel CD$ from the definition of a parallelogram.
- So $\angle ABD \cong \angle CDB$ by the Alternate Interior Angles Theorem because they are alternate interior angles in the transversal of $\overleftrightarrow{BD}$ over the parallel lines $\overleftrightarrow{AB}$ and $\overleftrightarrow{CD}$ (see Figure 1.13).

- Similarly, we know $BC \parallel AD$ from the definition of a parallelogram.
- So, $\angle ADB \cong \angle CBD$ (again by the Alternate Interior Angles Theorem).
- Also, $|BD| = |DB|$ trivially.
- Combining the 2nd, 4th, and 5th lines above with the ASA Congruence Criterion, we see that $\triangle ABD \cong \triangle CDB$.
- Examining corresponding sides of these triangles we see that $|AB| = |CD|$ and $|AD| = |CB|$. $\square$

The above proof illustrates a few more aspects of good proof writing. We list them here as guidelines to follow in writing your own proofs.

**1 - Opening.** We began our proof with a couple of introductory sentences to set the stage. One sentence describes the starting point by stating what is given – in this case it is just the introduction of $ABCD$ as a generic representative for the class of all parallelograms. Another sentence indicates where the proof is headed by stating what needs to be demonstrated. These sentences serve to orient the reader for what is to come (as well as to orient your own thinking for constructing the proof).

**WARNING:** Don't skip these opening sentences! Most errors by students writing proofs come from not getting the proof off to a good start. Writing this preamble to your proof will help prevent common mistakes such as confusing the hypothesis with the conclusion.

**2 - Justification of Steps.** Each step of the proof clearly states what is being established and by what rationale the step is justified. Note that when we use a fact such as the "Alternate Interior Angle Theorem" we should clarify its use for the reader by specifying the setting to which we are applying that fact (such as the appropriate transversal of lines).

**3 - Use of Diagrams.** Students are often unsure of what they can and cannot do with a diagram, so either end up avoiding diagrams altogether (which can make a proof needlessly wordy and confusing) or else using them in invalid ways. The rule is this: a diagram *can* be used to:

**clarify** – to make plain some object being described in the proof (such as in the second itemized step in the above proof), or

**simplify** – such as saying "*Let the points be named as in the diagram*". A single reference to a diagram can save a paragraph of wordy descriptions.

A diagram *cannot* be used, however, to supply any justification for a step in the proof. No matter how tempting, we can not allow a line like "*As can be seen from the figure,* $\angle JKL$ *is obtuse*" to enter in our proof.

**4 - Conclusion.** Finally, the end of the proof is marked by some symbol, in our case an open box at the right margin.

Look back at Example 1.1 in Section B. Though this example was posed as a computation, you'll note that the solution we give there is actually written as an itemized proof – and indeed, it actually *is* a proof, for the computation relies on deductive reasoning. If you didn't write your solutions to Exercises 1.17 through 1.24 as proofs, you could gain some proof-writing practice by re-writing those solutions now (see Exercise 1.30).

### Indirect Proofs

In the centuries between Pythagoras and Euclid a technique of proof known as *reductio ad absurdum* (or *reduction to an absurdity*) was championed by the school of the philosopher Plato. In modern terms, we call this technique "indirect proof", or "proof by contradiction". The technique consists of this: to prove a statement correct it is good enough to establish that it cannot possibly be incorrect. So, we begin by assuming that our given statement is false, and then reduce this assumption to an absurdity — that is, we derive from our assumption some impossible conclusion. This, in effect, proves that we must have been wrong to assume the statement is false.

More formally, an indirect proof of statement $S$ begins with the assumption ∼S (the statement *"S does not hold"*) and ends successfully when an impossibility is reached from this assumption. The assumed statement ∼S is called the *negation* of $S$. Constructing negations to statements is a necessary skill in the task of deduction, and fortunately comes easily for most people. While negations (like any statement) can be worded many different ways, one safe method for constructing the negation of a statement is merely to place the words *"It is not true that ..."* before the statement. Thus, the negation of *"Every parallelogram is a rectangle"* is *"It is not true that every parallelogram is a rectangle"*. This might then be reworded to the logically equivalent statement *"There is at least one parallelogram which is not a rectangle"*. Note that the negation uses an existential quantifier whereas the original statement was universal. This is true in general: the negations of general statements are existential, and the negations of existential statements are universal.

Since the time of the Greeks, the technique of indirect proof has become a mainstay of modern deductive reasoning and argument. It was this technique which led to a discovery which disturbed the prevailing theories of the Pythagoreans and caused a major scandal in the mathematics of the time. The Pythagoreans held as an axiom the belief that given any two positive numbers $x$ and $y$ there would be a smaller value $z$ that divided evenly into both. (That is, there would be integers $j$ and $k$ so that $x = jz$ and $y = kz$.) This property is called *commensurability* – the Pythagoreans believed any two positive quantities were commensurable. It was more than just a matter of philosophy. They used their axiom of commensurability freely in their mathematical proofs, assuming that such a $z$ could indeed be found for any given pair $(x, y)$.

But, alas, the axiom is false! This shouldn't surprise modern readers familiar with modern number systems, for the commensurability assumption is equivalent to the belief that all numbers are rational numbers (see Excercise 1.29). Anciently, it was quite a shock when a member of the Pythagorean school made the discovery known as the "*incommensurability of the diagonal and side of a square*": if $ABCD$ is a square then the quantities $|AB|$ and $|AC|$ are *not* commensurable. The effect of this was devastating because suddenly every proof that used this assumption of commensurability had to be discarded. It took many decades for the Greeks to repair the mathematical superstructure by replacing these fallacious proofs. Indeed, it was not until about 360 BC when Eudoxus of Cnidus[7] was able to give a satisfactory theory of ratios that the issue was finally put to rest. (It is probably only a legend – but a good story nonetheless – that the discoverer[8] of the incommensurability showed his result to Pythagoras during a sailing voyage. The ship returned to shore that day with a crew diminished by one. The consequences of mathematical heresy were apparently quite severe!)

Let's see now how a very simple application of indirect proof can be used to establish the infamous incommensurability result.

**Theorem 1.2.** *If $ABCD$ is a square then $|AB|$ and $|AC|$ are not commensurable.*

[7] In addition to his theory of ratios that resolved the incommensurability crisis, Eudoxus is remembered for his pioneering of the "*method of exhaustion*", the most sophisticated of techniques used by the early Greeks and a forerunner of the calculus.

[8] In the legend, a man named Hippaus.

*Proof:* Suppose $ABCD$ is a square. Assume (for the purpose of reaching a contradiction) that $|AB|$ and $|AC|$ *are* commensurable.

- Let's say that $|AB| = r$. The Pythagorean Theorem then shows that $|AC| = \sqrt{2}r$ (since $|AC|^2 = |AB|^2 + |BC|^2 = r^2 + r^2 = 2r^2$).
- By the commensurability assumption, there is a number $z$ and integers $j$ and $k$ so that $r = jz$ and $\sqrt{2}r = kz$.
- But then $\sqrt{2} = \frac{\sqrt{2}r}{r} = \frac{kz}{jz} = \frac{k}{j}$.

**Note:** this, of course, says that $\sqrt{2}$ is a fraction of integers and thus a rational number. You probably already know that $\sqrt{2}$ is irrational – the rest of our proof amounts to showing this.[9].

- We can assume that this fraction $k/j$ for $\sqrt{2}$ is in lowest terms. In particular, we can assume that at most one of $j$ or $k$ is even.
- Modifying the above equation only slightly, we have

$$k = \sqrt{2}j \quad \text{so } k^2 = 2j^2 \,.$$

- This means $k^2$ is an even number, so $k$ is also even and $j$ must be odd.
- But since $k$ is even, there is then an integer $m$ so that $k = 2m$. So:

$$2m = k = \sqrt{2}j \,.$$

- Squaring both sides, we have $4m^2 = 2j^2$ and thus $2m^2 = j^2$.
- This shows $j^2$ is even. But $j$ is odd, and an odd number can't have an even square! This is our contradiction. The trouble could only have arisen from our assumption that $|AB|$ and $|AC|$ are commensurable, so in fact they must *not* be commensurable.

$\square$

[9]For a geometric proof of the irrationality of $\sqrt{2}$, see the short note *Irrationality of the square root of two – a geometric proof* by T.M. Apostol, American Mathematical Monthly **(107)** p.841 [Nov. 2000].

**Proof by Induction**

One more important technique of proof is worth mentioning here – that of proof by mathematical induction. Mathematical induction is often useful in proving statements that must be demonstrated to hold true for an infinity of values of some integer variable. Suppose $S(n)$ is a statement involving the integer $n$ and we wish to prove $S(n)$ true for all integer values of $n$ greater than or equal to some fixed integer $n_0$. Mathematical induction accomplishes this with two steps:

**Step 1.** Verify that the statement $S(n_0)$ is true.

**Step 2.** Prove that if the statements $S(n_0),\ S(n_0+1),\ S(n_0+2), \ldots, S(k)$ are all true then the statement $S(k+1)$ is also true.

Step 2 now sets up a chain reaction for which step 1 is the spark: since $S(n_0)$ is true, step 2 implies $S(n_0+1)$ is also true. But then since both $S(n_0)$ and $S(n_0+1)$ are true, step 2 implies $S(n_0+2)$ is true. This in turn implies that $S(n_0+3)$ is true, and so on. The conclusion is that $S(n)$ is true for all values $n \geq n_0$.

The assumption (in step 2) that $S(n)$ is true whenever $n_0 \leq n \leq k$ is called the *inductive hypothesis*. Step 2, which in practice is almost always the bulk of the work in a proof by induction, amounts to taking the inductive hypothesis as an assumption and using it to prove the statement $S(k+1)$. For a concrete example, we'll now use induction to prove a well-known formula for the sum of the first $n$ integers.

THEOREM **1.3.** *For any integer $n \geq 1$ we have*

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2} \ .$$

*Proof:* In the notation we used to introduce the concept of proof by induction, the mathematical equality

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$

is the statement $S(n)$ which we must prove valid for all $n$.

**Step 1.** The statement $S(1)$ is simply

$$\sum_{i=1}^{1} i = \frac{1(1+1)}{2}$$

which is certainly true (since both sides clearly reduce to 1).

**Step 2.** We take as our inductive hypothesis the assumption that statement $S(n)$ is true for $n = 1, 2, 3, \ldots, k$. In particular, we assume the statement $S(k)$:

$$\sum_{i=1}^{k} i = \frac{k(k+1)}{2} \ .$$

We must use this assumption to prove the statement $S(k+1)$. That is, we must prove the formula

$$\sum_{i=1}^{k+1} i = \frac{(k+1)([k+1]+1)}{2} = \frac{(k+1)(k+2)}{2} \ .$$

This is easy if we remember that we can use our inductive hypothesis! The left side of this formula can be rewritten by "breaking off" the last term in the sum:

$$\sum_{i=1}^{k+1} i = \left( \sum_{i=1}^{k} i \right) + (k+1) \ .$$

The sum that remains is now covered by our inductive hypothesis:

$$\begin{aligned}
\sum_{i=1}^{k+1} i &= \left( \sum_{i=1}^{k} i \right) + (k+1) \\
&= \frac{k(k+1)}{2} + (k+1) \\
&= (k+1)\left( \frac{k}{2} + 1 \right) \\
&= \frac{(k+1)(k+2)}{2} \ .
\end{aligned}$$

It now follows that the formula must be valid for all positive integer values of $n$. $\square$

Lest you think that this technique could never be useful in geometry, we'll give an easy inductive proof here for the following interesting theorem.

**THEOREM 1.4.** *Given any triangle $T$ and any integer $n \geq 4$, we can cut $T$ into exactly $n$ isosceles triangles.*

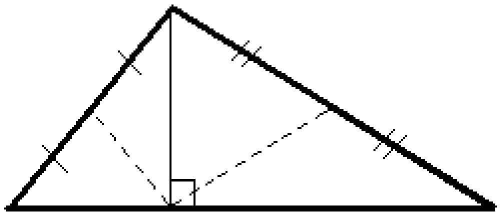


Figure 1.14: Cutting a triangle into four isosceles triangles

*Proof:* This proof is a bit less straightforward than our last example. We'll use induction to prove the non-equilateral case, leaving the equilateral case for later. So, let $S(n)$ be the statement "*if $T$ is a non-equilateral triangle then we can cut $T$ into exactly $n$ isosceles triangles.*" We will prove $S(n)$ is true for all $n \geq 4$.

For step 1 of the induction we prove the statement $S(4)$ – that is, we prove that any non-equilateral triangle $T$ may be cut into exactly four isosceles triangles. To do this, we first cut $T$ into two right triangles. Now the midpoint of the hypotenuse of a right triangle is equidistant from the three vertices of the triangle (see Exercise 1.38), so if we cut each of the two right triangles from its right angle vertex to the midpoint of its hypotenuse, the result is a decomposition of $T$ into four isosceles triangles (see Figure 1.14).

For step 2 of the process, our inductive hypothesis is that statements $S(4), S(5), \ldots, S(k)$ are all true. That is, we assume that any non-equilateral triangle may be cut into any number of isosceles triangles between 4 and $k$. We need to prove statement $S(k+1)$: given a non-equilateral triangle $T$ we must show that $T$ can be cut into exactly $k+1$ isosceles triangles.

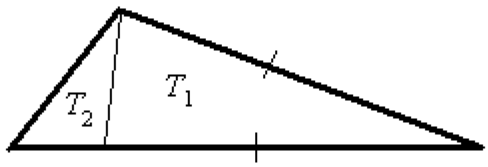



Figure 1.15:

The key to doing this is illustrated in Figure 1.15. Since $T$ is not equilateral it has two adjacent sides of different lengths. We may then cut $T$ so as to trim off a part of the longer side and thus create two new triangles, say $T_1$ and $T_2$ where $T_1$ is isosceles. Now by our inductive hypothesis we can cut $T_2$ into exactly $k$ isosceles triangles. (Note that $T_2$ is not equilateral since one of its angles must measure more than $90°$.) This cuts $T$ into $k+1$ isosceles triangles, as desired, and so completes the proof for the non-equilateral case.

To prove the theorem in the case that $T$ is equilateral, we need only show that an equilateral triangle can be cut into either four, five, or six isosceles triangles. (We leave this to you – see Exercise 1.41 at the end of this section.) Then, provided that one of the four isosceles triangles is not equilateral, we can cut an equilateral triangle into any number $k \geq 7$ of isosceles triangles as follows:

- First, cut the equilateral triangle into four isosceles triangles $T_1$, $T_2$, $T_3$, and $T_4$ where $T_1$ is not equilateral.
- Now use the non-isosceles case of the theorem (already proved above!) to cut $T_1$ into $k-3$ isosceles triangles.

This completes the proof of the theorem. $\square$

## Exercises

**1.25**. True or False? (Questions for discussion)

**(a)** "*If B has wings then B is a bird*" is the converse of "*If B is not a bird then B does not have wings*".

**(b)** "*If B does not have wings then B is not a bird*" is the contrapositive of "*If B is a bird then B has wings*".

**(c)** "*No bird has wings*" is the negation of "*Every bird has wings*".

**(d)** "$X$ is true only if $Y$ is true" is the same as $\mathrm{X} \Longrightarrow \mathrm{Y}$.

**(e)** A single counterexample can prove a universal statement to be false.

**(f)** A single example can prove a universal statement to be true.

**1.26**. Consider the following half of the Alternate Interior Angles Theorem: *The alternate interior angles formed by a line transversing two parallel lines are congruent.*

**(a)** State this theorem in the form "*If $X$ then $Y$*."

**(b)** State the contrapositive and converse of your answer to part (a). Be certain that your wording is clear.

**(c)** State this theorem using a universal quantifier, then state its negation.

**1.27**. State the converse to the Pythagorean Theorem, then combine both the Pythagorean Theorem and its converse into a single equivalence ("if and only if" statement).

**1.28**. For each of the following statements, give the contrapositive, converse, and negation. Word these as clearly as possible.

**(a)** Every dog has four legs.

**(b)** If a politician said it, it's a lie.

**(c)** Every equilateral triangle is isosceles.

**1.29**. Prove that if all pairs of numbers are commensurable then all numbers are rational numbers. (*Hint*: start with an arbitrary number $x$ and use the fact that $x$ is [by assumption] commensurable with 1 to prove that $x$ is rational.)

**In the exercises below, your proofs should use only the "toolbox theorems" as outlined in Section B.**

**1.30**. Rewrite your solutions to Exercises 1.17 through 1.24 as proofs.

**1.31**. Prove that a parallelogram is a rhombus if and only if its diagonals are perpendicular.

**1.32**. Prove that $ABCD$ is a parallelogram if and only if both $\angle A \cong \angle C$ and $\angle B \cong \angle D$.

**1.33**. Give both a direct proof and a proof by contradiction that if line $\Lambda_1$ is perpendicular to both lines $\Lambda_2$ and $\Lambda_3$, then $\Lambda_2 \parallel \Lambda_3$.

**1.34**. Prove that if $ABC$ is a triangle with $D$ and $E$ the midpoints of sides $AB$ and $CA$, respectively, then $DE \parallel BC$ and $|DE| = \frac{1}{2}|BC|$.

**1.35**. Prove that if $\triangle ABC \sim \triangle DEF$ with $a/d = t$, and if the height of $ABC$ relative to side $AB$ is $h$ then the height of $DEF$ relative to side $DE$ is $t \cdot h$ and $\text{area}(DEF) = t^2 \cdot \text{area}(ABC)$.

**1.36**. Imitate the last part of the proof of Theorem 1.2 to show by contradiction that $\sqrt{5}$ is irrational.

**1.37**. Prove the converse of Thales' Theorem.

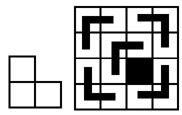
**1.38**. Use Exercise 1.37 to prove that if $ABC$ is a right triangle with right angle $\angle C$ and if $D$ is the midpoint of the hypotenuse $AB$ then $|DA| = |DB| = |DC|$.

**1.39**. Let $ABC$ be a triangle, $\Gamma_1$ the circle through $A$ and $B$ with center at the midpoint of $AB$, and $\Gamma_2$ the circle through $B$ and $C$ with center at the midpoint of $BC$. Prove that $\Gamma_1$ and $\Gamma_2$ intersect at a point of $CA$.

**1.40**. Use mathematical induction to prove the formula $\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$.

**1.41**. Complete the proof of Theorem 1.4 by showing that an equilateral triangle can be cut into either four, five, or six isosceles triangles where in each case, at least one of the isosceles triangles is not equilateral.

**1.42**. Demonstrate that an equilateral triangle may be cut into ten isosceles triangles by following the procedure outlined in the proof of Theorem 1.4. (*Note:* there are many ways to cut an equilateral triangle into ten isosceles triangles. However, there is only one correct answer to this exercise! The idea is to check that you understand the proof of Theorem 1.4 by carrying out the very specific procedure outlined in that proof.)



**1.43**. Suppose you have a grid of squares of size $2^n \times 2^n$ (for some positive integer $n$) with one square marked and a supply of tiles consisting of three squares in an "L" shape (as at right). Use induction on $n$ to prove that you can always place tiles on the grid so as to cover all of the unmarked squares with no tiles overlapping. (The example here shows how this might be done for one case of a four-by-four grid.) *Hint:* in the inductive step, divide the $2^{k+1} \times 2^{k+1}$ grid into four grids of equal size.

# D. Constructions

The early Greek mathematicians did not have the notational benefits we enjoy today. The development of algebra was centuries away, and dealing with non-integer quantities was a clumsy process in their number system. Perhaps that is why they turned to geometry as the primary way of doing and expressing mathematics. What we accomplish today with algebraic formulae or pocket calculators, the Greeks did with straightedge and compass. Geometric constructions were the means by which mathematics was computed and expressed. The diagonal of a square with side length 1 became the preferred way to represent the value $\sqrt{2}$, and computing the area of a figure was equated with constructing a square with equal area.

Because these constructions were so important to how early Greek mathematics was done, and because the early Greek mathematicians were sticklers for logical consistency, the process became a very exact science with very specific ground rules.

- The starting point was merely the existence of a unit length, so each construction began with just two points marked on an otherwise blank slate – the distance between these two points then became the unit of measurement, or the value "1".
- A valid construction then consisted of a finite sequence of steps, each of which was one of the following:
  - Drawing a straight line through two existing points.
  - Drawing a circle with center at an existing point and passing through another existing point.

The tools by which these two construction steps were accomplished consisted of a straightedge (for drawing lines) and a compass (for drawing circles). For this reason, the process became known as "straightedge and compass construction."

Note that the straightedge does *not* allow us to measure distances. In particular, we could not construct the midpoint between two points merely by measuring the distance with the straightedge and then marking the point halfway between them. (However, we'll show below how the midpoint *can* be constructed using the straightedge and compass together.)

Also, the compass the Greeks intended would collapse when lifted from the paper, so that it could construct a circle only once a point on that circle (as well as its center) was known. In particular, their compass did not allow one to mark a distance between two points with the compass and then transfer the compass (while holding this distance) to construct a circle with that radius centered at a third point. We, however, will allow ourselves a rigid compass that *can* transfer lengths. Exercise 1.46 at the end of this section shows that this is not really cheating: anything that can be done with our rigid compass can also be done with the Greeks' floppy compass (but with a lot more work).

A good construction should include two parts, of which the sequence of construction steps is just the first. Following the construction there should be given a justification of how we know the construction accomplishes the intended task – a *proof* of its validity. These justifications provide an excellent initiation into geometric proof.

We will begin with a sequence of three constructions to illustrate the process.

**EXAMPLE 1.5.** *Given two points $A$ and $B$, construct a point $C$ so that $\triangle ABC$ is equilateral (that is, has all three sides of equal length).*

*Construction:* (See Figure 1.16.)

- First, construct the circle $\Gamma_1$ centered at $A$ and passing through $B$.
- Next, construct the circle $\Gamma_2$ centered at $B$ and passing through $A$.
- Finally, let $C$ be one of the two points where the circles $\Gamma_1$ and $\Gamma_2$ meet.

*Justification:* If the distance from $A$ to $B$ is called $d$, then every point on $\Gamma_1$ is at distance $d$ from $A$, and every point on $\Gamma_2$ is at distance $d$ from $B$. So, the point $C$ is at distance $d$ from both $A$ and $B$, which clearly makes the triangle $\triangle ABC$ equilateral. □

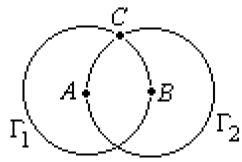



Figure 1.16:

**EXAMPLE 1.6.** *Given two points $A$ and $B$, construct the perpendicular bisector to the segment $AB$.*

*Construction:* (See Figure 1.17.)

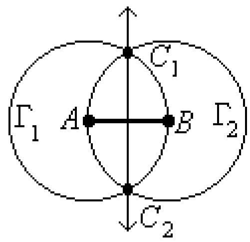



Figure 1.17:

- Construct the circles $\Gamma_1$ and $\Gamma_2$ as in the example above.
- Label the two points at which these circles meet as $C_1$ and $C_2$.
- The line $\overleftrightarrow{C_1C_2}$ is the perpendicular bisector of $AB$.

*Justification:* By the Perpendicular Bisector Theorem, the perpendicular bisector of segment $AB$ is exactly the locus of points equidistant from $A$ and $B$. The last construction example shows that $C_1$ and $C_2$ are each points of this locus. Since the locus is a line, it must be the line through these two points. So, $\overleftrightarrow{C_1C_2}$ is the perpendicular bisector of $AB$. □

Note that this last construction tacitly gives us the midpoint of the segment $AB$ as well. We need only mark where $\overleftrightarrow{C_1C_2}$ intersects $AB$.

Note also that once a construction is known, its steps need not be repeated each time it is used. For instance, we may now justifiably say "construct the midpoint of $AB$" or "construct the perpendicular bisector of $AB$" in future constructions, since we have already demonstrated how these steps can be accomplished. Our next example illustrates this.

**Example 1.7.** *Given a point $P$ and a line $\Lambda$, construct the line passing through $P$ that is perpendicular to $\Lambda$.*

*Construction:* (See Figure 1.18.)

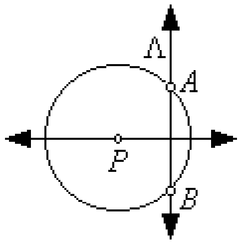



Figure 1.18: Construction of a line through $P$ perpendicular to $\Lambda$

- Choose a point $A$ lying on the line $\Lambda$.
- Construct the circle centered at $P$ and passing through $A$.
- Mark both points at which this circle meets $\Lambda$. One will be $A$, label the other one as $B$. (If there is only one point of intersection between the circle and $\Lambda$, then the circle is tangent to $\Lambda$ at $A$. In this case, $\overleftrightarrow{PA}$ will be perpendicular to $\Lambda$.)
- Finally, construct the perpendicular bisector to segment $AB$ (as in Example 1.6). This line will be perpendicular to $\Lambda$ and will pass through $P$.

*Justification:* The verification of this construction consists of checking the accuracy of the last sentence. Since $\Lambda$ is the line determined by $A$ and $B$, the perpendicular bisector of $AB$ will be perpendicular to $\Lambda$. Furthermore, $P$ is on this perpendicular bisector by the Perpendicular Bisector Theorem (since $P$ is clearly equidistant from $A$ and $B$). □

### Constructible numbers

As mentioned, constructions were the method of calculation to the early Greeks. Lengths and areas were considered to be understood once they could be constructed in a figure, and the Greeks' calculations thus led to many challenging geometric construction problems. The following are particularly interesting examples of this:

**Challenge 1:** Given a circle $\Gamma$ (with known radius), construct a line segment with length equal to the circumference of $\Gamma$.

**Challenge 2:** Given a circle $\Gamma$ (with known radius), construct a square with area equal to the area inside $\Gamma$.

If we take the radius of $\Gamma$ to be 1, these amount to constructing a line segment of length $2\pi$ and a square of side length $\sqrt{\pi}$ beginning only with a unit length. Today we say that a number is *constructible* if a segment of that length can be constructed by the usual Greek rules. (For instance, the number $\sqrt{2}$ is constructible because beginning with two points $A$ and $B$ at distance 1 apart, we can construct a square $ABCD$ and the diagonal $AC$ then has length $\sqrt{2}$.) Note that the feasibility of the above two challenges boils down to whether or not $\pi$ is a constructible number. For if length $\pi$ can be constructed, it's easy to then double it to $2\pi$. And if a length $x$ can be constructed, then so can $\sqrt{x}$ – see Exercise 1.47.

These challenges occupied a considerable amount of mathematical attention for over 2000 years! It was not until the 19th century that the issue was finally resolved with the proof that $\pi$ is *not* constructible. And despite the geometric roots to the problem, the ultimate resolution came from algebra – the two key elements of the proof are as follows:

- First, René Descartes' invention of analytic geometry (in the first half of the 17th century) led to the conclusion that all constructible numbers are solutions to polynomials with integer coefficients. (Such numbers are called *algebraic*.) The connection isn't difficult to see – by Descartes' method, the circles and lines in a construction became polynomials, and their points of intersection became solutions to polynomial equations.

- All that remained, then, was for the German mathematician Ferdinand Lindemann to prove (in 1882) that $\pi$ is not algebraic (a number that is not algebraic is called *transcendental*). The means by which he accomplished this are quite technical, so we will not go into them here.

We now understand the hierarchy of numbers to be as follows:

constructible numbers $\subset$ algebraic numbers $\subset$ real numbers

with none of these inclusions being equalities. As we have mentioned, $\pi$ is an example of a real number that is not algebraic (Euler's constant $e$ is another that you have probably encountered). For an example of an algebraic number that is not constructible, we'll digress for a moment to another famous challenge from early Greek geometry.

The problem of "*duplication of the cube*" was to construct the side length for a cube of volume equal to exactly twice that of a unit cube. (The corresponding two dimensional problem is easy – given a square of area 1, it is easy to construct a square of area 2 since we can use as side length of the larger square the diagonal of the unit square!) Duplication of the cube, then, amounts to constructing the length $\sqrt[3]{2}$, and while this number is certainly algebraic (being a solution to the equation $x^3 = 2$), it is *not* constructible (though see Exercise 1.49). This challenge also withstood solution for over two millennia, for the non-constructibility of $\sqrt[3]{2}$ was only proved (using Galois theory from abstract algebra) in the 19th century.

**Quadratures**

Challenge 2 given above is an example of a problem of *quadrature* of a region, or constructing a square with area equal to the area of a given region. If a quadrature is possible for a given region, we say that the region is *quadrable*. While the problem of quadrature of a circle's interior (or as it is sometimes called, "squaring the circle") is now known to be impossible, there are many interesting quadrature problems that *can* be successfully accomplished. We will here outline a proof that the interior of any simple polygon is quadrable. Some of the steps in this proof will be left for you as exercises.

LEMMA **1.8.** *The interior of any rectangle is quadrable.*

*Proof:* The proof takes the form of a construction. We begin our construction with a generic rectangle $ABCD$. The goal will be to construct a segment of length $\sqrt{|AB||BC|}$ (since once we have a segment of this length it is easy to construct a square having that segment as a side – see Exercise 1.45(a)). See Figure 1.19 for reference in the following construction steps.

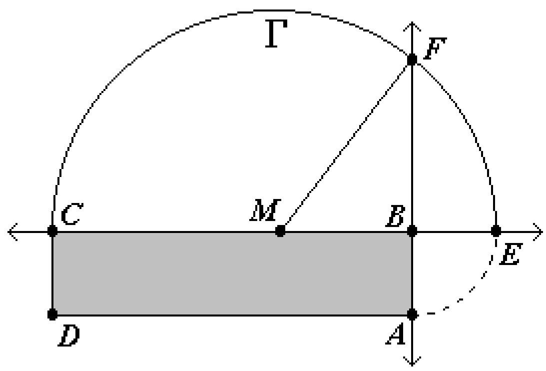



Figure 1.19:

- First, extend the line $\overleftrightarrow{BC}$.

- Find a point $E$ (with $B$ between $C$ and $E$) on this line so that $|BE| = |AB|$. (This can be done by intersecting the line $\overleftrightarrow{BC}$ with the circle centered at $B$ and having radius $|AB|$.)

- Find the midpoint $M$ of segment $CE$ (as in Example 1.6).
- Construct the circle $\Gamma$ with center $M$ and radius $|ME|$.
- Construct the line through $B$ that is perpendicular to $\overleftrightarrow{BC}$ (as in Example 1.7).
- Let $F$ be the intersection of this line with the circle $\Gamma$ (with $B$ between $A$ and $F$).

We leave to you (see Exercise 1.50) the justification that segment $BF$ has the desired length (so that a square built on this segment would have area $|AB||BC|$).

□

**THEOREM 1.9.** *Every polygonal region (the interior of a simple polygon) is quadrable.*

*Proof:* The proof is by induction on the number of sides to the bounding polygon. The base case is a polygon with 3 sides – that is, a triangle. We leave it to you (see Exercise 1.51(a)) to use Lemma 1.8 to show that the interior of any triangle is quadrable.

We now make our inductive hypothesis: assume that for some $k \geq 3$ it is true that quadrature can be done on the interior of any simple polygon with up to $k$ sides. We will complete the induction by showing every $k+1$-sided simple polygon has a quadrable interior.

So let $\Pi$ be a polygonal region of $k+1$ sides. Using one of the diagonals we may express $\Pi$ as the union of two polygonal regions $\Pi_1$ and $\Pi_2$, each of which has $k$ or fewer sides (see Figure 1.20). We may then apply our inductive hypothesis to each of these smaller regions to construct squares $A_1B_1C_1D_1$ and $A_2B_2C_2D_2$ with areas equal to the areas of $\Pi_1$ and $\Pi_2$, respectively. The lengths $|A_1B_1|$ and $|A_2B_2|$ can be used (with the Pythagorean Theorem) to produce the side length of a square with area equal to the area of $\Pi$ – see Exercise 1.51(b).

□

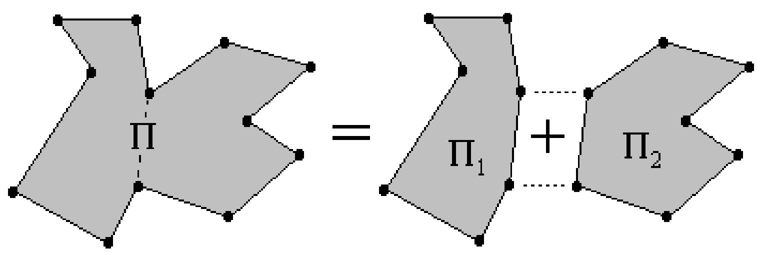



Figure 1.20:

## Exercises

**1.44**. True or False? (Questions for discussion)

**(a)** Every constructible number is an algebraic number.
**(b)** Every algebraic number is a constructible number.
**(c)** Every rational number is a constructible number.
**(d)** $\sqrt{n}$ is constructible for every positive integer $n$.
**(e)** If a region $\Pi$ can be cut into two pieces, each of which is quadrable, then $\Pi$ itself is quadrable.
**(f)** The number $\pi^2$ is not constructible.

**1.45**. Give constructions (similar to our examples above) for each of the following tasks. Include a justification for the accuracy of the construction. If you have access to a geometry software package, use it to test your construction.

**(a)** Given two points $A$ and $B$, construct a square with $AB$ as one side.

**(b)** Given two points $A$ and $B$, construct a square with $AB$ as a diagonal.

**(c)** Given an angle $\angle ABC$, construct a ray $\overrightarrow{BD}$ which bisects this angle.

**(d)** Given a line $\Lambda$ and a point $P$, construct the line through $P$ that is parallel to $\Lambda$.

**(e)** Given a line $\Lambda$ and a point $P$, construct the line through $P$ that meets $\Lambda$ at a $45°$ angle.

**(f)** Given a line $\Lambda$ and a point $P$ not on $\Lambda$, construct an equilateral triangle $PAB$ where $A$ and $B$ are points on $\Lambda$.

**(g)** Given two points $A$ and $B$, construct a rhombus with $AB$ as one side and with one angle measuring $45°$.

**(h)** Given two points $A$ and $B$, construct a rhombus with $AB$ as one side and with one angle measuring $30°$.

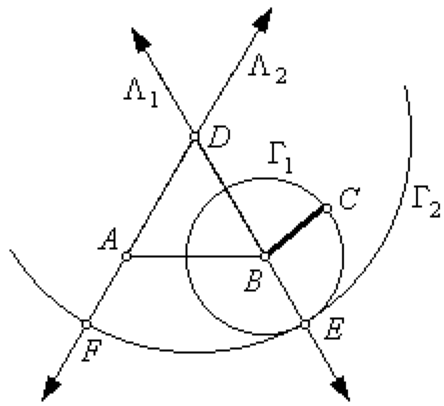
**(i)** Given two points $A$ and $B$ and a line $\Lambda$, construct a circle through $A$ and $B$ with center on $\Lambda$. (Under what circumstances will your construction **not** work?)

**(j)** Given three points $A$, $B$, and $C$ (not all on a common line), construct the circle passing through all three points.

**(k)** Given a line $\Lambda$, a point $A$ on $\Lambda$, and a point $B$ not on $\Lambda$, construct the circle tangent to $\Lambda$ at $A$ and passing through $B$.

**(l)** Given a circle $\Gamma$ construct an equilateral triangle $ABC$ with all three vertices on $\Gamma$.

**1.46**. The following construction is the proof of Theorem 2 from Euclid's *Elements*, which states that given a point $A$ and a segment $BC$ it is possible to construct the circle with center $A$ and radius $|BC|$. This shows that a collapsing compass (together with a straightedge) can still construct a circle with a given center and radius.




- Construct segment $AB$.
- Construct an equilateral triangle $ABD$ on segment $AB$
- Construct the circle $\Gamma_1$ with center at $B$ that passes through $C$.
- Construct the line $\Lambda_1$ through $D$ and $B$.
- Let $E$ be the point at which $\Lambda_1$ intersects $\Gamma_1$ such that $B$ is between $D$ and $E$ (remember that $B$ is the center of $\Gamma_1$).
- Construct the circle $\Gamma_2$ with center at $D$ that passes through $E$.
- Construct the line $\Lambda_2$ through $A$ and $D$.
- Let $F$ be the point at which $\Lambda_2$ intersects $\Gamma_2$ such that $A$ is between $D$ and $F$.
- The desired circle has center at $A$ and passes through $F$.

Give a justification for this construction by explaining why $|AF| = |BC|$.

**1.47**. Show that if $x$ is a constructible length then $\sqrt{x}$ is also constructible. (*Hint:* use a quadrature on a rectangle.)

**1.48**. Show (by outlining a construction) that each of the following numbers is constructible.

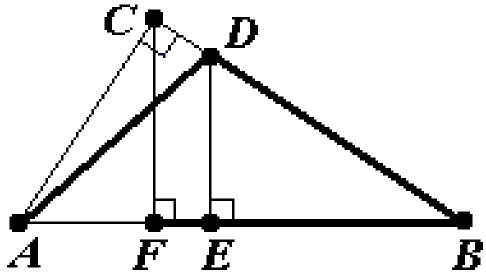
**(a)** $\sqrt{5}$

**(b)** $\frac{\sqrt{5}-1}{2}$ (the "*golden ratio*")

**(c)** $\sqrt{2+\sqrt{5}}$ (consider using a quadrature!)

**(d)** $\sqrt[4]{2}$

**1.49**. As mentioned on p.44, the number $\sqrt[3]{2}$ is not constructible. However, if we "cheat" just a little by using a *ruler* with marks instead of an unmarked straightedge, then we *can* construct $\sqrt[3]{2}$.




**(a)** In the figure at right, $|AD| = |DB| = |BF| = 1$. Show that $|BC| = \sqrt[3]{2}$. (*Hint:* use three similar right triangles.)

**(b)** Give a sequence of steps for constructing $BC$ using only a compass and a straightedge with marks 1 unit apart.

**1.50**. Complete the proof of Lemma 1.8 by giving a justification for the correctness of the construction. (*Hint:* Consider the right triangle $MBF$.)

**1.51**. Complete the proof of Theorem 1.9 as follows.

**(a)** Show how the quadrature of a triangle's interior may be accomplished. First construct a rectangle with area equal to the triangle's interior, then apply Lemma 1.8.

**(b)** Show how to construct a square area equal to the sum of the areas of the two given squares. (Use the Pythagorean Theorem.)

**1.52**. Prove that if region $\Sigma_1$ is a subset of region $\Sigma_2$ and both regions are quadrable then the region $\Sigma_2 \setminus \Sigma_1$ (the part of $\Sigma_2$ that is not in $\Sigma_1$) is also quadrable.

**1.53**. Carry out an actual quadrature on the interior of an isosceles triangle with sides of length 4, 6, and 6.

# Chapter 2

# Beyond the Basics

In Section 1B we described our "basic toolbox" of Euclidean geometry facts. Our goal for this chapter is to show just how useful those tools are. Here, we will investigate some interesting theorems in Euclidean geometry that go beyond the basics. We will prove them (as well as some of their extensions and relatives) using only our basic toolbox theorems. This should help you develop a sense for *why* the facts set out in Section 1B are regarded as the "core" of Euclidean geometry, and why it is a worthwhile endeavor to develop this core from a set of axioms, as we will do in Chapters 4 and 5.

Euclidean geometry abounds in theorems that *could* appear in this chapter – we have selected these few for the appeal of their statements and proofs. These theorems are not needed in the remainder of the text, so you may choose to do whichever sections seem most appropriate. You may even choose to do some of these topics later, perhaps after Chapter 5. Section G consists of additional short exercises that can be done using only the basic toolbox facts. These can be done on an occasional basis as you continue through the remainder of the text.

## A. Distances In Triangles

There is hardly a more familiar object in geometry than the equilateral triangle. Yet equilateral triangles figure prominently in two of the topics we will present in this chapter – there is often magic in even the most routine of geometric topics.

The starting point for this section is a simple but striking property of equilateral triangles that you may not have seen before. Using a geometry computer software package (or paper, pencil, ruler, and compass or protractor) form an equilateral triangle $ABC$ and let $P$ be any point inside this triangle. Measure the distances from $P$ to each of the three sides of $ABC$ (by forming a perpendicular segment from $P$ to each side) and then add these three distances. Now move the point $P$ to another location inside the triangle and measure again. What do you get?

You might be surprised that the sum of these distances apparently does not depend on where $P$ is located. The proof is amazingly simple.

**THEOREM 2.1.** *Let $ABC$ be an equilateral triangle and let $P$ be any point in the interior of $ABC$. As in Figure 2.1, let $D$, $E$, and $F$ be the points (on $AB$, $BC$, and $CA$ respectively) so that $PD \perp AB$, $PE \perp BC$, and $PF \perp CA$. Finally, let $G$ be on $AB$ so that $CG \perp AB$. Then $|PD| + |PE| + |PF| = |CG|$. In other words, no matter where $P$ is located, the sum of its distances to the three sides of $ABC$ is equal to the height of $ABC$.*

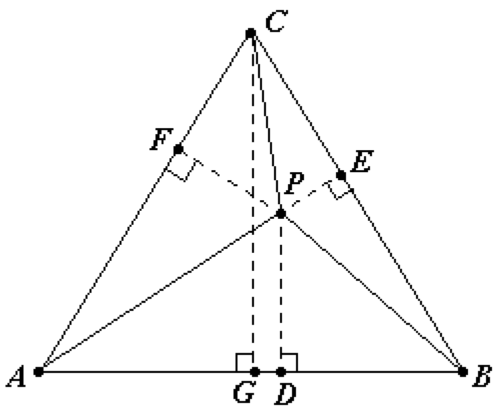



Figure 2.1:

*Proof:* The only tool we need for this theorem is the area formula for a triangle. We have:

$$\begin{aligned}\text{area}(ABC) &= \text{area}(PAB) + \text{area}(PBC) + \text{area}(PCA)\\ &= \frac{1}{2}|AB||PD| + \frac{1}{2}|BC||PE| + \frac{1}{2}|CA||PF|\end{aligned}$$

But since $ABC$ is equilateral, we have $|AB| = |BC| = |CA|$, so this equation can be rewritten as

$$\text{area}(ABC) = \frac{1}{2}|AB|\Big(|PD| + |PE| + |PF|\Big).$$

Applying the area formula to the left side, we get

$$\begin{aligned}\frac{1}{2}|AB||CG| &= \frac{1}{2}|AB|\Big(|PD| + |PE| + |PF|\Big)\\ |CG| &= |PD| + |PE| + |PF|\end{aligned}$$

which is exactly what we were to prove! □

We can now apply this simple property of equilateral triangles to a very natural distance problem for more general triangles.

**DEFINITION.** Let $ABC$ be a triangle. Define the function $d_{ABC}(P) = |PA| + |PB| + |PC|$ for each point $P$ in the interior of $ABC$. We say that $Q$ is a *minimum distance point* for $ABC$ if $Q$ is in the interior of $ABC$ and $d_{ABC}(Q)$ is minimum.

In other words, a minimum distance point is a point for which the sum of the distances to the vertices of $ABC$ is as small as possible. Can we identify the minimum distance points for a given triangle? Can a triangle have more than one minimum distance point? The key to answering these questions is another definition.

**DEFINITION.** A point $Q$ in the interior of triangle $ABC$ is the *Steiner point* for $ABC$ if $m\angle AQB = m\angle BQC = m\angle CQA = 120°$.

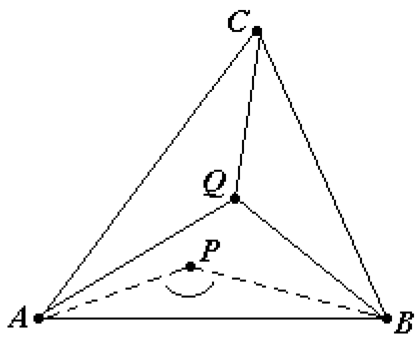



Figure 2.2:

It is easy to see that the Steiner point for a triangle (if it exists) must be unique – no triangle can have more than one Steiner point. (For if $Q$ is a Steiner point for $ABC$ and $P$ is any other point in the interior of $ABC$ then $P$ lies in the interior of one of the triangles $ABQ$, $BCQ$, or $CAQ$. But if, say, $P$ is in the interior of $ABQ$ then $m\angle APB > m\angle AQB = 120°$ – see Figure 2.2 and Exercise 2.4.) It is also easy to see that some triangles don't have Steiner points. In particular, if any of the angles of $ABC$ measure 120° or more then $ABC$ has no Steiner point. But if each angle measures less than 120° then the construction in Exercise 2.5 shows that the triangle *does* have a Steiner point. And for such triangles, the Steiner point provides the answer to our questions.

**THEOREM 2.2.** *Let $ABC$ be a triangle with all angles of measure less than 120°. Then the unique minimum distance point for $ABC$ is its Steiner point.*

*Proof:* Let $Q$ be the Steiner point for $ABC$ and let $P$ be any other point in the interior of $ABC$. We will prove that $d_{ABC}(P) > d_{ABC}(Q)$. Refer to Figure 2.3 in the following steps.

- Form the lines through $A$, $B$, and $C$ perpendicular to the segments $QA$, $QB$, and $QC$.
- These lines meet at points $D$, $E$, and $F$ (as in Figure 2.3), and triangle $DEF$ is equilateral. (See Exercise 2.9.)
- Let $h$ be the height of $DEF$ – then by Theorem 2.1 we see that for both $P$ and $Q$, the sum of the distances to the sides of $DEF$ is $h$.

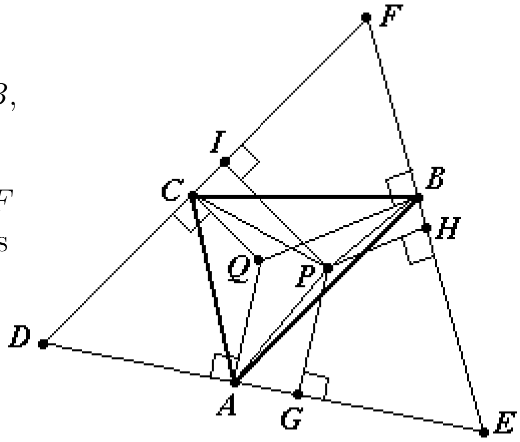



Figure 2.3:

- Let $G$, $H$, and $I$ be the points (as in Figure 2.3) giving the distances from $P$ to the sides of $DEF$.
- Then $APG$ is a right triangle (with right angle at $G$), so by the Pythagorean Theorem we see

  $$|PA| = \sqrt{|PG|^2 + |AG|^2} \geq |PG|$$

  with equality if and only if $G = A$.
- Similarly, $|PB| \geq |PH|$ and $|PC| \geq |PI|$ (with equality if and only if $H = B$ or $I = C$ respectively).
- Since $P \neq Q$, we cannot have $G = A$, $H = B$, and $I = C$ all true. So:

$$\begin{aligned} d_{ABC}(P) &= |PA| + |PB| + |PC| \\ &> |PG| + |PH| + |PI| \\ &= h \\ &= |QA| + |QB| + |QC| \\ &= d_{ABC}(Q)\,. \end{aligned}$$

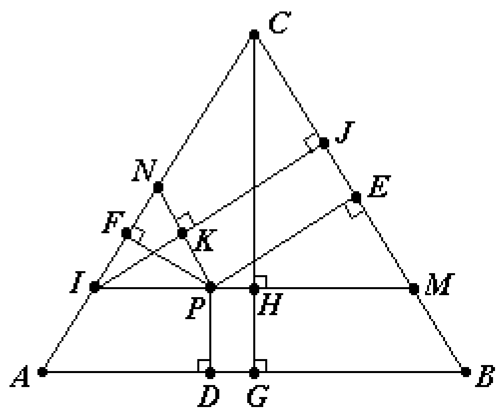
□

## Exercises

**2.1**. Using a dynamic geometry software package, illustrate Theorem 2.1. (You should be able to move the point $P$ within the equilateral triangle $ABC$ while the software keeps track of the perpendicular distances to each side of the triangle.)

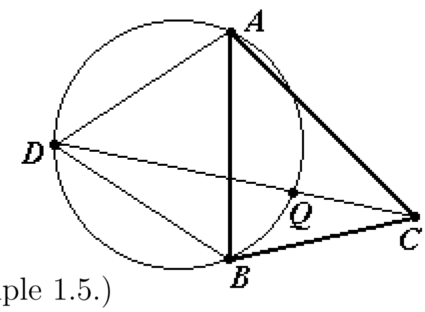
**2.2**. Does Theorem 2.1 hold if the point $P$ is on one of the sides of $ABC$? Justify your answer.

**2.3**. An alternate proof of Theorem 2.1 can be based on the figure at right. Here, $D$, $E$, $F$, and $G$ are as in the theorem statement, and the other segments are as indicated. Complete the proof by showing that $|GH| = |PD|$ and $|CH| = |PE| + |PF|$.




**2.4**. Prove in Figure 2.2 that $m\angle APB > 120°$.

**2.5**. Let $ABC$ be a triangle with all angles measuring less than 120°. Give a justification that the following construction produces the Steiner point for $ABC$.




- Construct an equilateral triangle $ABD$ with $D$ on the opposite side of $\overleftrightarrow{AB}$ from $C$. (Recall the construction in Example 1.5.)
- Construct the circle $\Gamma$ through the points $A$, $B$, and $D$. (The center can be found by intersecting the perpendicular bisectors of $AD$ and $BD$ – see Example 1.6.)
- Construct the segment $CD$.
- The Steiner point is the point $Q$ where $CD$ crosses $\Gamma$.

**2.6**. Use a dynamic geometry software package to construct an illustration of Theorem 2.2. (You should be able to construct the Steiner point $Q$ using Exercise 2.5. You should now be able to move another point $P$ within the triangle $ABC$, keeping track of $d_{ABC}(P)$ and comparing it to $d_{ABC}(Q)$.)

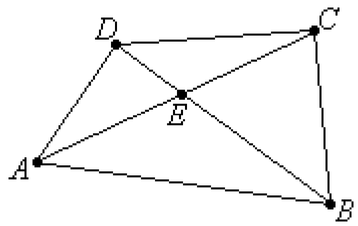
**2.7**. Let $ABC$ be an equilateral triangle with sides of length 8, and let $P$ be a point in the interior of $ABC$. Find the sum of the distances from $P$ to the sides of $ABC$.

**2.8**. Let $ABC$ be an isosceles triangle with sides of length 8, 8, and 12. Find the minimum value of $d_{ABC}(P)$ for any point $P$ in the interior of $ABC$. (*Hint:* see the construction in Exercise 2.5 above.)

**2.9**. Prove that in the proof of Theorem 2.2 the triangle $DEF$ is equilateral.

**2.10**. Suppose $ABCD$ is a quadrilateral with diagonals $AC$ and $BD$ that intersect at a point $E$. Prove that the sum $|PA| + |PB| + |PC| + |PD|$ is minimized when $P = E$.

**2.11**. Consider the following theorem:




*Let $ABCD$ be a quadrilateral with diagonals $AC$ and $BD$ that intersect at a point $E$. Then*

$$\text{area}(ABE)\cdot\text{area}(CDE) = \text{area}(BCE)\cdot\text{area}(ADE)\ .$$

**(a)** Use a geometry computer software package to demonstrate this theorem.

**(b)** Prove the theorem using only the area formula for triangles.

## B. Ceva's Theorem

In this section we study a theorem first discovered by (and named for) the Italian mathematician Giovanni Ceva (1648-1734). It is a deeper result than those in the last section, and its proof is more difficult. But while the statement of Ceva's Theorem may not seem striking, its consequences are! In fact, while you may have never heard of Ceva's Theorem, you probably *have* heard of some facts we will derive from it as easy corollaries.

You may recall that a collection of lines or segments are said to be *concurrent* if they share a common point. The "classical" version of Ceva's Theorem is as follows (see Figure 2.4):

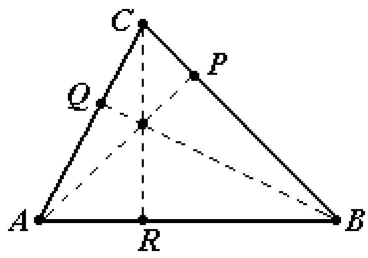



Figure 2.4:

**Ceva's Theorem.** *Let $ABC$ be any triangle and let $P$, $Q$, and $R$ be points on the sides $BC$, $CA$, and $AB$ respectively (other than the points $A$, $B$, and $C$ themselves). Then $AP$, $BQ$, and $CR$ are concurrent if and only if* $\frac{|AR|}{|RB|} \cdot \frac{|BP|}{|PC|} \cdot \frac{|CQ|}{|QA|} = 1.$

From this theorem comes the term *cevian*:

**Definitions.** If $P$ is a point on $BC$ then the segment $AP$ is called a *cevian* of triangle $ABC$.

Of course, if (in the above definition) $P$ is the midpoint of $BC$ then the cevian $AP$ is a median of the triangle, and if $\angle BAP \cong \angle PAC$ then $AP$ is the angle bisector for $ABC$ at vertex $A$ (see p.17). Concurrence for these two types of cevians are well-known facts, but there is usually some hard work involved in their proofs. But once we have Ceva's theorem (whose proof we will give shortly) there is very little additional work needed, as we now observe.

**Corollary 2.3.** *The three medians of any triangle are concurrent.*

*Proof:* If $AP$, $BQ$, and $CR$ are the medians of $ABC$ then each of the fractions $\frac{|AR|}{|RB|}$, $\frac{|BP|}{|PC|}$, and $\frac{|CQ|}{|QA|}$ equal 1. So by Ceva's Theorem, $AP$, $BQ$, and $CR$ are concurrent. □

**Corollary 2.4.** *The three angle bisectors for any triangle are concurrent.*

*Proof:* See Exercise 2.18.

**Definitions.** The point common to the medians of a triangle is called the *centroid* of the triangle. The point common to the angle bisectors is called the *incenter* of the triangle.

(See Exercise 2.19 for the justification of the name “incenter”.)

Before proceeding, we’ll pause to note an interesting (and again, well-known) fact about the medians and centroid of a triangle. This fact will prove useful in Sections C and D.

**Theorem 2.5.** *If $ABC$ is any triangle, $AP$ is its median at vertex $A$, and $V$ its centroid, then $|AV| = 2|VP|$. (In other words, the centroid cuts each median into lengths of ratio 2:1.)*

*Proof:* Let the medians of $ABC$ be $AP$, $BQ$, and $CR$ and let $V$ be the centroid.

- We leave it as Exercise 2.15 to show that the six triangles $ARV$, $RBV$, $BPV$, $PCV$, $CQV$, and $QAV$ have equal areas.

- Let $d$ be the distance from $B$ to the line $\overleftrightarrow{AP}$ (as in Figure 2.5).

- Then

$$\begin{aligned}
\frac{1}{2}|AV|d &= \text{area}(ABV) \\
&= \text{area}(ARV) + \text{area}(RVB) \\
&= 2\,\text{area}(BPV) \\
&= d|VP|
\end{aligned}$$

which proves $|AV| = 2|VP|$.

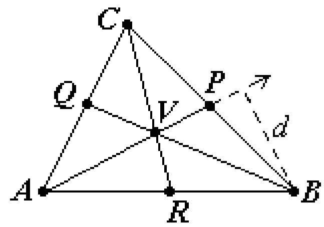



Figure 2.5:

□

We now turn our attention to proving Ceva’s Theorem. Our proof will use the following property of distances on a line.

**Fact.** *Given a line $\overleftrightarrow{AB}$ and a positive real number $t \neq 1$ there are exactly two points $X_1$ and $X_2$ on $\overleftrightarrow{AB}$ satisfying the condition $\frac{|AX_i|}{|X_iB|} = t$. Furthermore, exactly one of these points is on the segment $AB$. The midpoint of $AB$ is the only point $X$ on $\overleftrightarrow{AB}$ satisfying $\frac{|AX|}{|XB|} = 1$.*

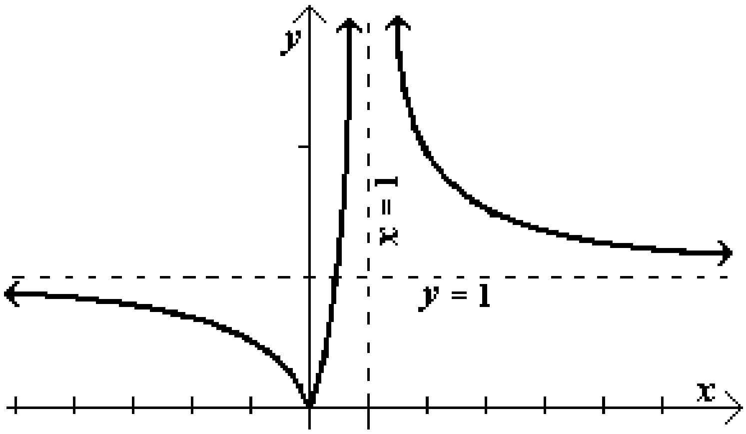



Figure 2.6: The graph of $y = \left|\frac{x}{1-x}\right|$

To see a justification for this, note that distances on a line in the plane behave exactly like distances on the real number line $\mathbb{R}$. Then simply think of points $A$ and $B$ as the points 0 and 1 on that number line. If $X$ is then the point $x \in \mathbb{R}$, the ratio $\dfrac{|AX|}{|XB|}$ is $\left|\frac{x}{1-x}\right|$. The graph of $y = \left|\frac{x}{1-x}\right|$ is shown in Figure 2.6. (Note the vertical asymptote at $x = 1$ and the horizontal asymptote at $y = 1$ – can you explain these geometrically?) A formal justification of the above fact could be made by applying the Mean Value and Intermediate Value Theorems from calculus to this function. But it is probably sufficient to convince oneself from examining the graph that every positive $y$-value is achieved exactly once in the interval $(0, 1)$ and every positive $y$-value except $y = 1$ is achieved exactly once in $(-\infty, 0) \cup (1, \infty)$.

*Proof of Ceva's Theorem:* The theorem states an equivalence, so its proof has two parts. Refer to Figure 2.4 in what follows, taking $V$ to be the point where the three cevians cross.

**Part 1:** First assume that $AP$, $BQ$, and $CR$ are concurrent, meeting at a point $V$. We will prove that $\dfrac{|AR|}{|RB|} \cdot \dfrac{|BP|}{|PC|} \cdot \dfrac{|CQ|}{|QA|} = 1$.

- Triangles $ARC$ and $RBC$ have equal height $h$ relative to the sides $AR$ and $RB$, so by the area formula for triangles,

$$\frac{\text{area}(ARC)}{\text{area}(RBC)} = \frac{\frac{1}{2}h|AR|}{\frac{1}{2}h|RB|} = \frac{|AR|}{|RB|}\,.$$

- The same argument applied to triangles $ARV$ and $RBV$ shows

$$\frac{\text{area}(ARV)}{\text{area}(RBV)} = \frac{|AR|}{|RB|}\,,$$

so

$$\frac{\text{area}(ARC)}{\text{area}(RBC)} = \frac{\text{area}(ARV)}{\text{area}(RBV)} = \frac{|AR|}{|RB|}\,.$$

- An elementary property of ratios is as follows: if $x/y = z/w = t$ then $x = ty$ and $z = tw$, so $(x \pm z)/(y \pm w) = (ty \pm tw)/(y \pm w) = t$ also.

- Applying this rule to the ratios we obtained above, we have

$$\frac{|AR|}{|RB|} = \frac{\text{area}(ARC) - \text{area}(ARV)}{\text{area}(RBC) - \text{area}(RBV)} = \frac{\text{area}(CAV)}{\text{area}(BCV)}\,.$$

- Parallel arguments to this can be given to show

$$\frac{|BP|}{|PC|} = \frac{\text{area}(ABV)}{\text{area}(CAV)} \quad \text{and} \quad \frac{|CQ|}{|QA|} = \frac{\text{area}(BCV)}{\text{area}(ABV)}\,.$$

- Putting these facts together, we get the desired result:

$$\frac{|AR|}{|RB|} \cdot \frac{|BP|}{|PC|} \cdot \frac{|CQ|}{|QA|} = \frac{\text{area}(CAV)}{\text{area}(BCV)} \cdot \frac{\text{area}(ABV)}{\text{area}(CAV)} \cdot \frac{\text{area}(BCV)}{\text{area}(ABV)} = 1\,.$$

**Part 2:** Now assume that $AP$, $BQ$, and $CR$ are cevians of a triangle $ABC$ such that $\frac{|AR|}{|RB|} \cdot \frac{|BP|}{|PC|} \cdot \frac{|CQ|}{|QA|} = 1$. We will prove that they are concurrent. The battle plan is simple: $AP$ and $BQ$ will meet at some point – call it $W$. We will show that $\overleftrightarrow{CW}$ meets $AB$ at point $R$, meaning that $AP$, $BQ$, and $CR$ are concurrent at $W$.

- Because the point $W$ is interior to the triangle $ABC$, the line $\overleftrightarrow{CW}$ must meet side $AB$ at a point we will call $R'$.

- Since $AP$, $BQ$, and $CR'$ are concurrent at $W$, part 1 above proves that $\frac{|AR'|}{|R'B|} \cdot \frac{|BP|}{|PC|} \cdot \frac{|CQ|}{|QA|} = 1$.

- Since we are assuming $\frac{|AR|}{|RB|} \cdot \frac{|BP|}{|PC|} \cdot \frac{|CQ|}{|QA|} = 1$, it must be true that $\frac{|AR|}{|RB|} = \frac{|AR'|}{|R'B|}$.

- From our preliminary fact, $R$ and $R'$ must be identical!

- So $W$ is on all of $AP$, $BQ$, and $CR$, and the proof is complete. $\square$

An altitude of a triangle (again, see p.17) may or may not be a cevian for the triangle – the point $P$ on $\overleftrightarrow{BC}$ so that $AP \perp \overleftrightarrow{BC}$ may not be on segment $BC$. But you are probably aware that, similar to the medians and angle bisectors, the *lines* containing the altitudes of a triangle are concurrent. That fact can be proved from Ceva's Theorem in the case that all angles of the triangle are acute (so that the altitudes *are* cevians). To prove this fact in the general case requires a more general version of Ceva's Theorem. The basic ideas of the above proof can be used to prove the theorem below, though there are a few subtleties involved in the adaptation (see Exercise 2.21).

THEOREM **2.6.** **(Generalized Ceva's Theorem)** *Let $ABC$ be any triangle and let $P$ be a point on $\overleftrightarrow{BC}$ (other than $B$ or $C$), $Q$ a point on $\overleftrightarrow{CA}$ (other than $C$ or $A$), and $R$ a point on $\overleftrightarrow{AB}$ (other than $A$ or $B$) so that no two of the lines $\overleftrightarrow{AP}$, $\overleftrightarrow{BQ}$, and $\overleftrightarrow{CR}$ are parallel. Then $\overleftrightarrow{AP}$, $\overleftrightarrow{BQ}$, and $\overleftrightarrow{CR}$ are concurrent if and only if the following are true:*

**(i)** *the sides of $ABC$ together contain an odd number (one or three) of the points $\{P, Q, R\}$, and*

**(ii)** $\dfrac{|AR|}{|RB|} \cdot \dfrac{|BP|}{|PC|} \cdot \dfrac{|CQ|}{|QA|} = 1.$

COROLLARY **2.7.** *If $ABC$ is a triangle with altitudes $AP$, $BQ$, and $CR$ then $\overleftrightarrow{AP}$, $\overleftrightarrow{BQ}$, and $\overleftrightarrow{CR}$ are concurrent.*

*Proof:* First we note that we may assume $ABC$ is not a right triangle, for the altitudes of a right triangle are (trivially!) concurrent at the right angle vertex. For $ABC$ a non-right triangle, we need only verify that **(i)** and **(ii)** in the statement of Theorem 2.6 hold true for the altitudes $AP$, $BQ$, and $CR$. We first verify **(i)**.

- Clearly, $P$ is on side $BC$ if and only if neither $\angle B$ nor $\angle C$ is an obtuse angle (see Figure 2.7). Similarly, angles $\angle A$ and $\angle C$ determine whether or not $Q$ is on side $AC$, and angles $\angle A$ and $\angle B$ determine whether or not $R$ is on side $AB$.

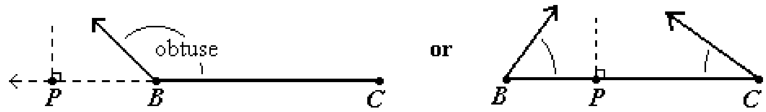



Figure 2.7:

- But by the 180° Sum Theorem, $ABC$ has either zero obtuse angles or exactly one obtuse angle.
- If $ABC$ has no obtuse angles, then clearly all three points of $\{P, Q, R\}$ lie on its sides.
- If $ABC$ has exactly one obtuse angle, then exactly one point of $\{P, Q, R\}$ lies on its sides. (For instance, if $\angle A$ is obtuse then neither $Q$ nor $R$ lie on sides of $ABC$, but since both $\angle B$ and $\angle C$ are acute in this case, $P$ *would* lie on side $BC$.)

Thus, statement **(i)** holds.

We now show that the equality in statement **(ii)** holds also. Figure 2.8 illustrates both possibilities of either one or all points of $\{P, Q, R\}$ lying on the sides of $ABC$. You should verify from this that the steps in the proof below do not depend on how the figure is drawn.

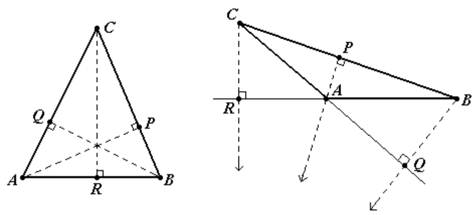



Figure 2.8:

- A consequence of the 180° Sum Theorem is that two right triangles are similar if a corresponding pair of their acute angles are congruent.
- Using this fact, we have the following pairs of similar triangles:
  - $\triangle ABQ \sim \triangle ACR$
  - $\triangle BCR \sim \triangle BAP$
  - $\triangle CAP \sim \triangle CBQ$
- From these and the Similar Triangles Theorem we get:

$$\frac{|AR|}{|AQ|} = \frac{|AC|}{|AB|}, \quad \frac{|BP|}{|BR|} = \frac{|BA|}{|BC|}, \text{ and } \frac{|CQ|}{|CP|} = \frac{|CB|}{|CA|}.$$

- Applying these ratios to the expression in statement **(ii)** we have

$$\begin{aligned}\frac{|AR|}{|RB|}\cdot\frac{|BP|}{|PC|}\cdot\frac{|CQ|}{|QA|} &= \frac{|AR|}{|AQ|}\cdot\frac{|BP|}{|BR|}\cdot\frac{|CQ|}{|CP|} && \left(\begin{array}{c}\text{rearranging}\\ \text{denominators}\end{array}\right)\\ &= \frac{|AC|}{|AB|}\cdot\frac{|BA|}{|BC|}\cdot\frac{|CB|}{|CA|} && \text{(above equalities)}\\ &= 1\,.\end{aligned}$$

Thus, statements **(i)** and **(ii)** from Theorem 2.6 both hold, so by that theorem, the lines $\overleftrightarrow{AP}$, $\overleftrightarrow{BQ}$, and $\overleftrightarrow{CR}$ are concurrent. □

**Definition.** The point at which the lines in Corollary 2.7 meet is called the *orthocenter* of the triangle $ABC$.

## Exercises

**2.12**. Use a dynamic geometry software package to construct illustrations of the following facts. (In each case, you should be able to change the shape of the triangle $ABC$ and observe that the medians, angle bisectors, or altitudes remain concurrent.)

**(a)** Corollary 2.3

**(b)** Corollary 2.4

**(c)** Corollary 2.7

**2.13**. What happens to the statement of Ceva's Theorem if we allow $P = C$? In what sense is the theorem still true in this case?

**2.14**. Show that it is possible for **(ii)** in the statement of Theorem 2.6 to be true while **(i)** is false. (*Hint:* it might be easiest to work with an equilateral triangle $ABC$. Can you arrange $P$ and $Q$ so that $\frac{|BP|}{|PC|}\cdot\frac{|CQ|}{|QA|} = 1$ while only one is on a side of $ABC$?)

**2.15**. Let $ABC$ be a triangle with medians $AP$, $BQ$, and $CR$ and centroid $V$. Prove that the six triangles $ARV$, $RBV$, $BPV$, $PCV$, $CQV$, and $QAV$ have equal area.

**2.16**. Carry out the steps to show (in the proof of Theorem 2.6) that

$$\frac{|BP|}{|PC|} = \frac{\text{area}(ABV)}{\text{area}(CAV)}.$$

**2.17**. Let $ABC$ be any triangle and let $P$, $Q$, and $R$ be points on the sides $BC$, $CA$, and $AB$ respectively so that $|AB|+|BP| = |AC|+|CP|$ (that is, $P$ is "half way around the perimeter of the triangle from $A$"), $|BC|+|CQ| = |BA|+|AQ|$, and $|CA|+|AR| = |CB|+|BR|$. Prove that the segments $AP$, $BQ$, and $CR$ are concurrent. (*Hint:* label the six distances around the perimeter of the triangle and work with equations in these six variables to apply Ceva's Theorem.)

**2.18**. In this exercise we will prove Corollary 2.4.

**(a)** Prove that if $AP$ is the angle bisector for $ABC$ at $A$ then $\dfrac{|BP|}{|PC|} = \dfrac{|AB|}{|AC|}$. (*Hint:* construct the ray $\overrightarrow{AP}$ and the line through $C$ parallel to $AB$.)

**(b)** Use part (a) and Ceva's Theorem to prove Corollary 2.4.

**2.19**. Let $ABC$ be a triangle and $\Gamma$ a circle inscribed in $ABC$. Prove that the center of $\Gamma$ is the point common to the angle bisectors (the *incenter*).

**2.20**. Let $ABC$ be a triangle and let $\Gamma$ be its inscribed circle, with $\Gamma$ tangent to $BC$ at point $P$, to $CA$ at point $Q$, and to $AB$ at point $R$. Use Ceva's Theorem to show that $AP$, $BQ$, and $CR$ are concurrent. (*Hint:* what can you say about $|AQ|$ and $|AR|$?)

**2.21**. This exercise extends the proof of Ceva's Theorem to a proof of Theorem 2.6.

**(a)** First, assuming that $\overleftrightarrow{AP}$, $\overleftrightarrow{BQ}$, and $\overleftrightarrow{CR}$ are concurrent at a point $V$, prove that **(i)** from the statement of Theorem 2.6 is true. (*Hint:* Consider the seven regions determined by the three lines $\overleftrightarrow{AB}$, $\overleftrightarrow{BC}$, and $\overleftrightarrow{AC}$. The point $V$ must be in one of these regions.)

**(b)** With the same assumption as part **(a)**, prove statement **(ii)** from Theorem 2.6 by showing how the first part of the proof of Ceva's Theorem can be adapted to handle the case where the sides of $ABC$ contain only one of the points $\{P, Q, R\}$.

**(c)** Finally, adapt the second part of the proof of Ceva's Theorem to the more general case. (Note that if $W$ is the point at which $\overleftrightarrow{AP}$ and $\overleftrightarrow{BQ}$ meet, it is no longer clear that $\overleftrightarrow{CW}$ is not parallel to $\overleftrightarrow{AB}$.)

## C. Napoleon's Theorem

Yes, *that* Napoleon! It was not with mathematics that Napoleon Bonaparte (1769-1821) earned his fame, but he *was* a great admirer of science and scientists, and at least in his own mind a man of some mathematical achievement (at one point arranging his own election to a prestigious academic position in geometry). The main theorem of this section carries his name by tradition, though it is uncertain whether he actually discovered or proved the result.

We will begin by recalling something that should be familiar from elementary trigonometry. Its proof is left as an easy exercise (see Exercise 2.23). For the sake of simplicity, we will refer to a right triangle with other angles of 30° and 60° by the often-used informal title "*30-60-90 triangle*".

**LEMMA 2.8.** *If $ABC$ is a 30-60-90 triangle with $m\angle A = 30°$ and $m\angle B = 60°$ then $b = \sqrt{3}a$ and $c = 2a$.*

Napoleon's Theorem involves the construction illustrated in Figure 2.9 at right. Starting with any triangle $ABC$ we form equilateral triangles $ABD$, $BCE$, and $CAF$ outward from the sides of $ABC$. (This means, for instance, that $D$ is chosen on the side of line $\overleftrightarrow{AB}$ not containing point $C$.) We first prove that the segments shown dashed in this figure have equal length. This (somewhat surprising) fact will play a crucial role in the proof of Napoleon's Theorem itself.

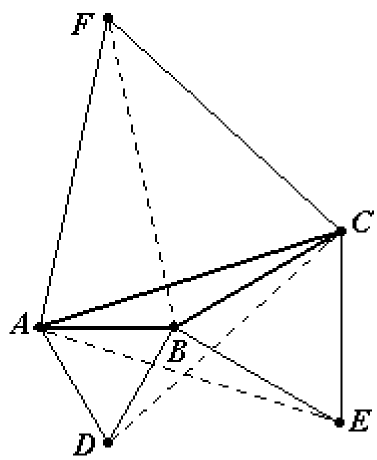


Figure 2.9:

**THEOREM 2.9.** *In the construction described above, $|AE| = |BF| = |CD|$.*

*Proof:* We will prove that $|AE| = |BF|$ by showing that $\triangle ECA \cong \triangle BCF$. (The proof that $|AE| = |CD|$ is similar, using triangles $\triangle EBA \cong \triangle CBD$.)

- $EC \cong BC$ since $BCE$ is equilateral.
- $CA \cong CF$ since $CAF$ is equilateral.
- $m\angle ECA = 60° + m\angle BCA = m\angle BCF$.

So, by the SAS congruence criterion, $\triangle ECA \cong \triangle BCF$, as claimed. □

**Note:** In the above proof we used the fact that all angles of an equilateral triangle measure 60°. While we did not feel it merited lengthening the proof with a justification of this, you should quickly remind yourself of how this follows easily from the definition of equilateral (all three sides the same length) using the Isosceles Triangle Theorem and the 180° Sum Theorem.

We're now ready to prove Napoleon's Theorem. This is one of the most impressive of geometry theorems to illustrate with computer software (see Exercise 2.22). Construct points $A$ through $F$ as described above. Then, as in Figure 2.10 at right, locate the points $K$, $L$, and $M$ as the centroids of the equilateral triangles $ABD$, $BCE$, and $CAF$ respectively. The magic should now be evident, especially if you have the freedom to move points $A$, $B$, and $C$ on your computer screen – no matter how $A$, $B$, and $C$ are placed, the triangle $KLM$ is always equilateral!

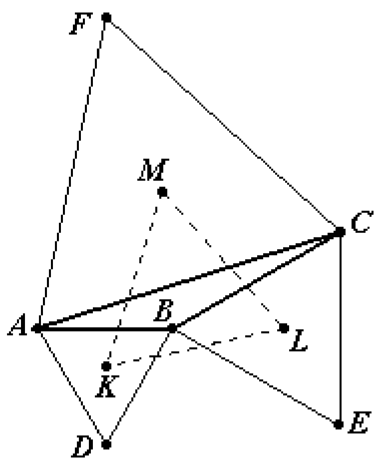



Figure 2.10:

Napoleon's Theorem. *With the construction described above, triangle $KLM$ is equilateral.*

*Proof:* We will prove directly that the sides of $KLM$ have equal length. Refer to Figure 2.11 in the following steps. As in that figure, let $AP$ and $FQ$ be medians of $CAF$.

- By Theorem 2.5 applied to the equilateral triangle $CAF$, $|AM| = \frac{2}{3}|AP|$. (See also Exercise 2.24.)
- But $CAP$ is a 30-60-90 triangle, so by Lemma 2.8, $\dfrac{|AC|}{|AP|} = \frac{2}{\sqrt{3}}$.

- Combining these facts, we get

$$\begin{aligned}\frac{|AC|}{|AM|} &= \frac{|AC|}{\frac{2}{3}|AP|} = \frac{3}{2}\left(\frac{|AC|}{|AP|}\right) \\ &= \frac{3}{2}\cdot\frac{2}{\sqrt{3}} = \sqrt{3}\,.\end{aligned}$$

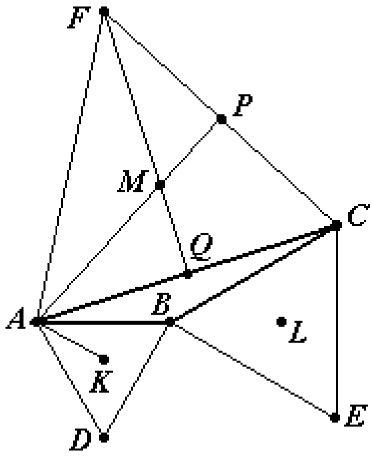



Figure 2.11:

- A similar argument shows $\dfrac{|AD|}{|AK|} = \sqrt{3}$.
- The angles $\angle MAK$ and $\angle CAD$ are congruent because $m\angle MAK = m\angle MAQ + m\angle QAK$ $= 30^\circ + m\angle QAK = m\angle KAD + m\angle QAK$ $= m\angle QAD = m\angle CAD$.
- Applying the second part of the Similar Triangles Theorem (p.17) to the facts above, we have $\triangle MAK \sim \triangle CAD$.
- The first part of the Similar Triangles Theorem then gives us

$$\frac{|CD|}{|MK|} = \frac{|AC|}{|AM|} = \sqrt{3}\,.$$

- Similar steps can be done to show that

$$\triangle LCM \sim \triangle BCF \quad \text{and} \quad \triangle KBL \sim \triangle ABE\,,$$

yielding

$$\frac{|BF|}{|LM|} = \frac{|BC|}{|CL|} = \sqrt{3} \quad \text{and} \quad \frac{|AE|}{|KL|} = \frac{|AB|}{|BK|} = \sqrt{3}\,.$$

- So, $|MK| = \frac{1}{\sqrt{3}}|CD|$, $|LM| = \frac{1}{\sqrt{3}}|BF|$, and $|KL| = \frac{1}{\sqrt{3}}|AE|$.
- But we know from Theorem 2.9 that $|CD| = |BF| = |AE|$. So the above equalities actually show $|MK| = |LM| = |KL|$, which is exactly what we needed to prove. $\square$

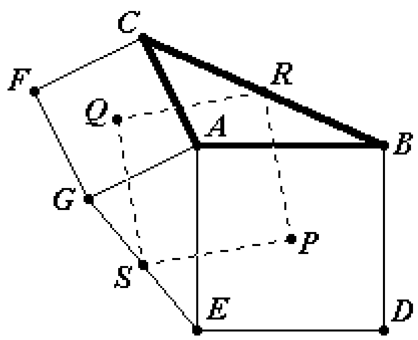
## Exercises

**2.22**. Use a dynamic geometry software package to construct an illustration of Napoleon's Theorem.

**2.23**. Prove Lemma 2.8.

**2.24**. Use Lemma 2.8 to prove directly (without using Theorem 2.5) that in the case of an equilateral triangle, the centroid cuts each median into lengths of ratio 2:1.

**2.25**. Carry out the steps to show (as in the proof of Napoleon's Theorem) that $\triangle LCM \sim \triangle BCF$.




**2.26**. The following result, known as the *Finsler-Hadwiger Theorem*, has somewhat of the same spirit as Napoleon's Theorem:

> *Let $ABC$ be any triangle. Construct squares $ABDE$ and $ACFG$ outward from two of its sides as in the figure. Let $P$ and $Q$ be the centers of these squares and let $R$ and $S$ be the midpoints of the segments $BC$ and $EG$ (as shown). Then $PRQS$ is a square.*

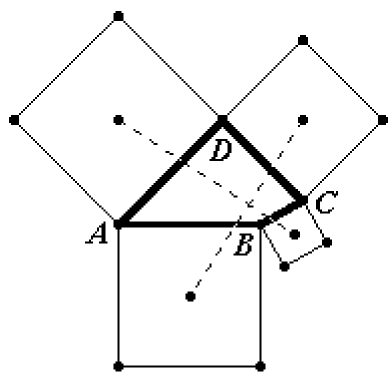
Prove this theorem in the following steps.

**(a)** Show that $BG \cong EC$.

**(b)** Show that $BG \perp EC$.

**(c)** Now use the result of Exercise 1.34 to show that $PRQS$ is a square.

**2.27**. Use the theorem from the previous exercise to prove that if $ABCD$ is any convex quadrilateral and if squares are constructed outward on its sides, then the segments connecting the centers of opposite pairs of these squares are perpendicular and congruent (see the figure).



## D. Cevian Lengths

The Pythagorean Theorem is an example of a formula involving lengths in a triangle – in that case the side lengths of a right triangle. In this section our goal is to develop several interesting formulae involving the lengths of various cevians (p.57) of an arbitrary triangle. Recall from Section B that medians and angle bisectors are interesting examples of cevians, so we will particularly focus on formulae involving their lengths. However, our most powerful result will be Stewart's Theorem, which gives a formula for the length of a general cevian in an arbitrary triangle.

The first stop on our cevian tour will be angle bisectors. The following lemma (useful in many settings) is needed to prove our length formula in this case. We leave its proof as an easy exercise (see Exercise 2.31).

LEMMA **2.10.** *Let $\Gamma$ be a circle with points $A$, $B$, $C$, and $D$ in cyclic order on $\Gamma$. Also let the chords $AC$ and $BD$ intersect at point $P$. Then $|AP| \cdot |PC| = |BP| \cdot |PD|$.*

THEOREM **2.11.** *Let $ABC$ be a triangle and let $AP$ be the angle bisector cevian at vertex $A$. Let $|BP| = a_1$ and $|PC| = a_2$. Then $|AP|^2 = bc - a_1 a_2$.*

*Proof:* Let $\Gamma$ be the circle through $A$, $B$, and $C$, and let $D$ be the point at which ray $\overrightarrow{AP}$ meets $\Gamma$ (as in Figure 2.12).

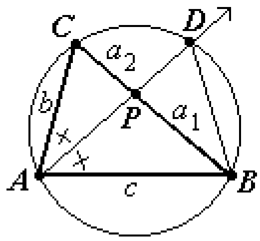


Figure 2.12:

- By the Inscribed Angle Theorem, $\angle ADB \cong \angle ACB$.
- So by the 180° Sum Theorem, $\angle ABD \cong \angle APC$, so $\triangle ABD \sim \triangle APC$.
- By the Similar Triangles Theorem, $\dfrac{|AB|}{|AP|} = \dfrac{|AD|}{|AC|}$, so

$$\begin{aligned} |AB| \cdot |AC| &= |AP| \cdot |AD| \\ cb &= |AP|(|AP| + |PD|) \\ &= |AP|^2 + |AP| \cdot |PD| \end{aligned}$$

- But $|AP| \cdot |PD| = |BP| \cdot |PC| = a_1 a_2$ by Lemma 2.10, so

  $$cb = |AP|^2 + a_1 a_2$$

  which proves the theorem. □

We next turn our attention to lengths of medians for a triangle. Our first result does not give a formula for the length of a single median, but rather puts bounds on the sum of the lengths of all three medians.

THEOREM **2.12.** *Let $AP$, $BQ$, and $CR$ be the medians of triangle $ABC$. Then $\frac{3}{4}(a+b+c) < |AP| + |BQ| + |CR| < a+b+c$.*

*Proof:* We first prove the lower bound. Let $V$ be the intersection of the medians (as in Figure 2.13 at right).

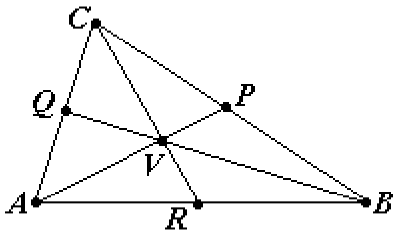


Figure 2.13:

- Recall the triangle inequality for distance: $|XZ| \leq |XY| + |YZ|$ with equality if and only if $Y$ is on segment $XZ$.
- From this inequality comined with Theorem 2.5 we have

$$\begin{aligned} a &= |BC| < |BV| + |CV| = \frac{2}{3}|BQ| + \frac{2}{3}|CR| \\ b &= |CA| < |CV| + |AV| = \frac{2}{3}|CR| + \frac{2}{3}|AP| \\ c &= |AB| < |AV| + |BV| = \frac{2}{3}|AP| + \frac{2}{3}|BQ| \end{aligned}$$

- Adding these, we get $a+b+c < \frac{4}{3}(|AP| + |BQ| + |CR|)$, or $\frac{3}{4}(a+b+c) < |AP| + |BQ| + |CR|$.

It remains for us to prove the upper bound $|AP| + |BQ| + |CR| < a+b+c$. Refer to Figure 2.14 in the following steps.

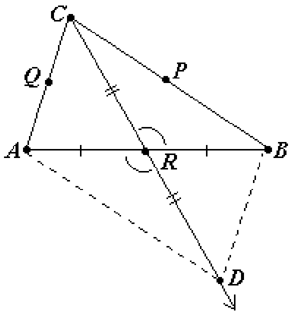


Figure 2.14:

- Find the point $D$ on $\overrightarrow{CR}$ so that $R$ is the midpoint of segment $CD$.
- Then $\angle ARD \cong \angle BRC$ by the Vertical Angles Theorem.
- So $\triangle ARD \cong \triangle BRC$ by the SAS criterion (remember that $R$ is the midpoint of both $CD$ and $AB$).
- Now $AD$ and $BC$ are corresponding sides of these triangles, so $|AD| = |BC| = a$.
- But then from the triangle inequality,

$$2|CR| = |CD| < |CA| + |AD| = b + a\,.$$

- Similar steps would show $2|AP| < b + c$ and $2|BQ| < a + c$.
- Adding these, we get $2(|AP| + |BQ| + |CR|) < 2(a + b + c)$, and dividing both sides by 2 gives our desired upper bound. □

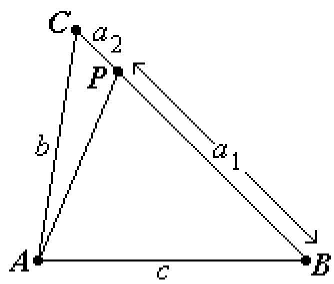



Figure 2.15:

We will give a formula for the length of a median in Theorem 2.13, but to get there we will actually need a more powerful theorem. In fact, this next result is really very remarkable, for it allows us to give the length of an arbitrary cevian $AP$ in an arbitrary triangle $ABC$ in terms of four easily-measured quantities: the side lengths $b$ and $c$ and the two lengths $a_1$ and $a_2$ into which $P$ cuts side $BC$ (see Figure 2.15). The formula was discovered by the Scottish mathematician Matthew Stewart (1717-1785) for whom it is named.

**Stewart's Theorem.** *In the above figure (with $a = a_1 + a_2$) we have $a_1 b^2 + a_2 c^2 = a(|AP|^2 + a_1 a_2)$.*

*Proof:* Let $D$ be the point on $\overleftrightarrow{BC}$ so that $AD \perp \overleftrightarrow{BC}$. The proof of Stewart's Theorem comes from repeated application of the Pythagorean Theorem to various right triangles in the figures below. (We include depictions of both the case where $D$ is on side $BC$ and the case where $B$ is between $D$ and $C$. You can check that the steps below are valid for both of these. The case where $C$ is between $D$ and $B$ can be reduced to one of these by merely reversing the names of $B$ and $C$.) So with distances as labeled in the diagrams, we proceed.

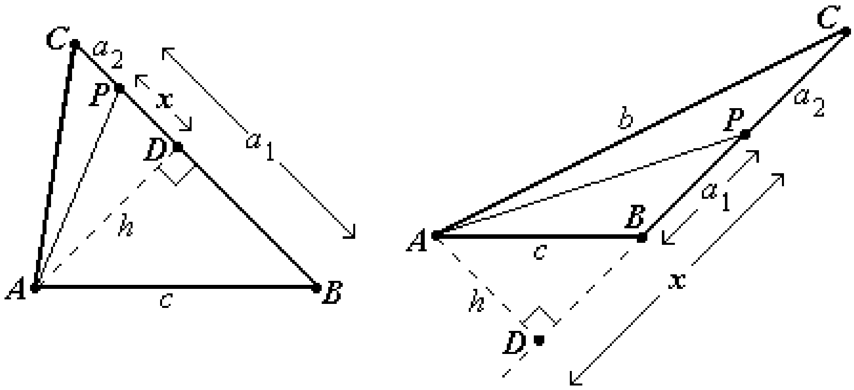



Figure 2.16:

- Considering the right triangle $BAD$ we have

$$(\mathbf{1}) \quad h^2 + |x - a_1|^2 = c^2$$

- From the right triangle $PAD$ we have

$$(\mathbf{2}) \quad h^2 + x^2 = |AP|^2$$

- And from the right triangle $CAD$ we have

$$(\mathbf{3}) \quad h^2 + (x + a_2)^2 = b^2$$

- Combining **(1)** and **(2)** we get

$$\begin{aligned} c^2 - |x - a_1|^2 &= |AP|^2 - x^2 \\ c^2 + 2xa_1 - a_1^2 &= |AP|^2 \\ (\mathbf{4}) \qquad c^2 &= |AP|^2 + a_1^2 - 2xa_1 \end{aligned}$$

- Combining **(2)** and **(3)** we get

$$\begin{aligned} |AP|^2 - x^2 &= b^2 - (x + a_2)^2 \\ |AP|^2 &= b^2 - a_2^2 - 2xa_2 \\ (\mathbf{5}) \qquad b^2 &= |AP|^2 + a_2^2 + 2xa_2 \end{aligned}$$

- Finally, multiplying both sides of **(4)** by $a_2$, multiplying both sides of **(5)** by $a_1$, and adding the results, we obtain our desired formula:

$$\begin{aligned} a_1 b^2 + a_2 c^2 &= (a_1 + a_2)|AP|^2 + a_1^2 a_2 + a_1 a_2^2 \\ &= a|AP|^2 + a(a_1 a_2) \\ &= a(|AP|^2 + a_1 a_2) \end{aligned}$$

□

THEOREM **2.13.** *If $AP$ is a median of $ABC$ then $|AP|^2 = \frac{1}{2}(b^2+c^2) - \frac{1}{4}a^2$.*

*Proof:* We have $|BP| = |PC| = \frac{1}{2}a$, so applying Stewart's Theorem we have

$$\begin{aligned} \frac{1}{2}a(b^2 + c^2) &= a(|AP|^2 + \frac{1}{4}a^2) \\ b^2 + c^2 &= 2|AP|^2 + \frac{1}{2}a^2 \end{aligned}$$

which is clearly equivalent to the claimed formula. □

COROLLARY **2.14.** *If $AP$, $BQ$, and $CR$ are the medians of $ABC$ then $|AP|^2 + |BQ|^2 + |CR|^2 = \frac{3}{4}(a^2 + b^2 + c^2)$.*

*Proof:* See Exercise 2.32.

## Exercises

**2.28**. Use a dynamic geometry software package to construct illustrations of the following theorems. Your illustrations should allow you to change the shape of triangle $ABC$ while keeping track of the lengths involved.

**(a)** Theorem 2.11

**(b)** Theorem 2.13

**(c)** Corollary 2.14

**2.29**. Use a dynamic geometry software package to construct an illustration of Theorem 2.12. Your illustration should allow you to change the shape of $ABC$, all the time displaying the three quantities in the theorem's inequality. Experiment with your illustration to discover how to make the quantity $|AP| + |BQ| + |CR|$ close to $a + b + c$ or $\frac{3}{4}(a + b + c)$.

**2.30**. Use a dynamic geometry software package to construct an illustration of Stewart's Theorem. Your illustration should allow you to move point $P$ along the side $BC$, showing that the equality is maintained at all times.

**2.31**. Prove Lemma 2.10.

**2.32**. Use Theorem 2.13 to prove Corollary 2.14.

**2.33**. Let $ABC$ be a right triangle with right angle at $C$, and suppose $a = 3$, $b = 4$, and $R$ is a point on $AB$ with $|AR| = 2$ and $|RB| = 3$. Find $|CR|$.

**2.34**. Suppose $ABC$ is an isosceles triangle with $b = c = 13$ and $P$ is a point on $BC$ so that $|BP| = |PC| + 6$. Find the length $a$ if $|AP| = 12$.

**2.35**. Let $ABC$ be a right triangle and let $P$ and $Q$ be points on the hypotenuse $AB$ with $|AP| = |PQ| = |QB| = \frac{1}{3}c$. Prove that $|CP|^2 + |CQ|^2 = \frac{5}{9}c^2$.

**2.36**. Prove the "*Parallelogram Identity*": If $ABCD$ is a parallelogram then $|AC|^2 + |BD|^2 = 2(|AB|^2 + |BC|^2)$.

# E. Ptolemy's Theorem and Trigonometry

The exact origins of trigonometry are obscure, and it is difficult to date the first uses of ideas we now describe in trigonometric language. Sometime in the 2nd century AD the mathematician and astronomer Hipparchus is reported to have published (though the work does not survive to the present) a work on the calculation of chord lengths in a circle – certainly a trigonometric endeavor. But Hipparchus' work was soon supplanted by that of Claudius Ptolemy of Alexandria. In 150 AD Ptolemy published a book so superior to its predecessors that (though the name given by its author was *Syntaxis Mathematica*) later mathematicians called it *Magiste*, or "*the greatest*". Arabic scholars in subsequent centuries would attach the article *al*, and today we identify Ptolemy's book by the name *Almagest*.[1]

Like the work of Hipparchus, Ptolemy's *Almagest* included a method for calculating chord lengths. In fact, Ptolemy's work on that subject was likely an adaptation or expansion of what Hipparchus had done. But Ptolemy's calculations were apparently more thorough (he carries his computations through to the point of giving chord lengths for arcs in increments of one-half degree – an amazing achievement for his time) and exceedingly well organized. Ptolemy's method was to apply a clever geometric theorem that today carries his name. In this section we will prove Ptolemy's Theorem, then show how it can be employed to calculate chord lengths. Finally, we will further demonstrate the geometric roots of trigonometry by defining the basic trigonometric functions and proving a geometric version of the familiar "Law of Sines".

Ptolemy's Theorem deals with a special type of quadrilateral specified in the following definition.

**Definition.** The quadrilateral $ABCD$ is a *cyclic quadrilateral* if all of its vertices lie on a common circle.

**Ptolemy's Theorem.** *If $ABCD$ is a cyclic quadrilateral then* $|AC||BD| = |AB||CD| + |AD||BC|$. (In other words, the product of the diagonal lengths is equal to the sum of the products of the lengths of opposite sides.)

*Proof:* The proof is short, but quite ingenious in that it relies on the construction of a seemingly obscure point (we will call $P$) on diagonal $AC$. Though its role is not at all obvious at first, it is what makes the pieces fit together. Refer to Figure 2.17 in the following steps.

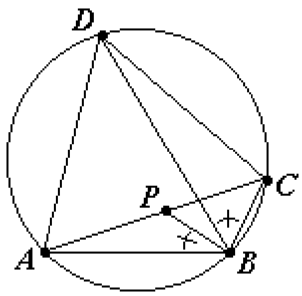


Figure 2.17:

- Let $P$ be the point on $AC$ so that $\angle ABP \cong \angle DBC$. (There is such a point because, since $ABCD$ is cyclic, $m\angle DBC < m\angle ABC$.)
- By the Inscribed Angle Theorem, $\angle PAB \ (= \angle CAB) \cong \angle CDB$.
- So then (using the 180° Sum Theorem and the definition of similarity) $\triangle ABP \sim \triangle DBC$.

[1]The *Almagest* was more than a geometry or trigonometry text. It also set forth an earth-centered model of the solar system. This model served remarkably well until it was eventually supplanted in the 16th century by the work of Copernicus (who, like Ptolemy, advanced the art of trigonometry as a tool in his astronomical work).

- From the Similar Triangles Theorem, we then have $\dfrac{|AB|}{|DB|} = \dfrac{|AP|}{|DC|}$, or

$$(*) \quad |AB||DC| = |AP||DB| \,.$$

- Now $m\angle ABD = m\angle ABP \pm m\angle DBP = m\angle DBC \pm m\angle DBP = m\angle PBC$ (where the choice of addition or subtraction depends on whether $P$ and $A$ are on the same side or opposite sides of $\overleftrightarrow{BD}$). So $\angle ABD \cong \angle PBC$.
- By the Inscribed Angle Theorem, $\angle BCP \;\; (= \angle BCA) \cong \angle ADB$.
- As above, then, $\triangle ABD \sim \triangle PBC$.
- By the Similar Triangles Theorem, $\dfrac{|AD|}{|PC|} = \dfrac{|BD|}{|BC|}$, or

$$(**) \quad |AD||BC| = |BD||PC| \,.$$

- Adding the equations $(*)$ and $(**)$ we have

$$|AD||BC| + |AB||DC| = |BD|(|AP| + |PC|) = |BD||AC| \,.$$

□

We will now show how to apply Ptolemy's Theorem to calculate chord lengths in circles. We will use the notation chord($\theta$) for the length of a chord with minor arc measuring $\theta$ in a circle of radius 1. (It is easy to relate this function to the usual modern trigonometric functions – see Exercise 2.42.) Figure 2.18 illustrates this function and shows the following starting values for our computations:

- chord$(60°) = 1$
- chord$(36°) = x$ where $x$ satisfies the equation $\frac{1}{x} = \frac{x}{1-x}$. We obtain this equation by applying the Similar Triangles Theorem to the similar isosceles triangles at the right of Figure 2.18. Solving this equation, we have chord$(36°) = \frac{\sqrt{5}-1}{2}$, the famous "golden mean".

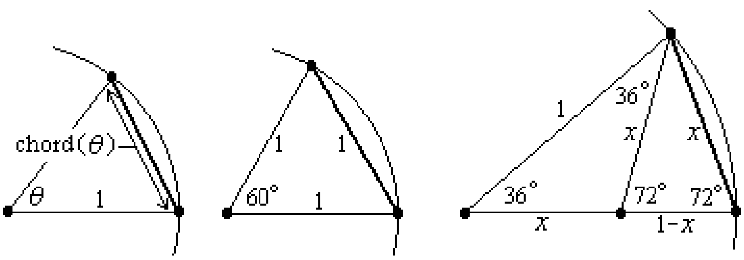



Figure 2.18: Definition of the function chord($\theta$) (left), chord$(60°) = 1$ (middle), chord$(36°) = \frac{\sqrt{5}-1}{2}$ (right)

From these starting values, we can easily calculate chord($\theta$) for other values of $\theta$ using Ptolemy's Theorem. The corollaries below outline the method.

**Corollary 2.15.** *In a circle $\Gamma$ of radius 1 let $AB$ be a chord of length $x$ and let the subtended arc of this chord be $\overset{\frown}{ACB}$. Suppose the chord $AC$ has length $y$. Then $|BC| = \frac{1}{2}(x\sqrt{4-y^2} - y\sqrt{4-x^2})$.*

*Proof:* Let $A'$ be the point of $\Gamma$ so that $AA'$ is a diameter. The proof is a direct application of Ptolemy's Theorem to the cyclic quadrilateral $ACBA'$ (see Figure 2.19).

- Note that $\angle ABA'$ and $\angle ACA'$ are right angles by the Inscribed Angle Theorem (or the Theorem of Thales).
- So by the Pythagorean Theorem, $|BA'| = \sqrt{4-x^2}$ and $|CA'| = \sqrt{4-y^2}$.
- Applying Ptolemy's Theorem, we have $x\sqrt{4-y^2} = 2|BC| + y\sqrt{4-x^2}$, which is clearly equivalent to the stated equation. $\square$

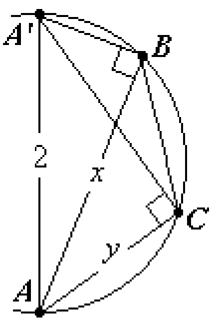



Figure 2.19:

**Corollary 2.16.** *In a circle $\Gamma$ of radius 1 let $AB$ be a chord length $x$. Let $C$ be the point of the subtended arc of this chord such that the arc subtended by $AC$ has measure one-half that of $\overset{\frown}{ACB}$. (That is, $AC$ subtends "half" of the subtended arc of $AB$.) Then $|AC| = \sqrt{2-\sqrt{4-x^2}}$.*

*Proof:* See Exercise 2.43.

Applying Corollary 2.16 to a chord subtending an arc of $60°$ we have

$$\text{chord}(30°) = \sqrt{2 - \sqrt{4 - \text{chord}^2(60°)}} = \sqrt{2-\sqrt{3}}\,.$$

Then we can apply Corollary 2.15 to the arc $\overset{\frown}{ACB}$ where $AB$ subtends an arc of $36°$ and $AC$ subtends an arc of $30°$ to obtain

$$\begin{aligned}\text{chord}(6°) &= \frac{1}{2}\Big[\text{chord}(36°)\sqrt{4-\text{chord}^2(30°)} \\ &\qquad\qquad - \text{chord}(30°)\sqrt{4-\text{chord}^2(36°)}\,\Big] \\ &= \frac{1}{2}\left[\frac{\sqrt{5}-1}{2}\sqrt{2+\sqrt{3}} \;-\; \sqrt{2-\sqrt{3}}\sqrt{4-\left(\frac{\sqrt{5}-1}{2}\right)^2}\,\right] \\ &\approx 0.1046719\end{aligned}$$

We could then continue, using Corollary 2.16 to evaluate chord$(3°)$, chord$(1\frac{1}{2}°)$, and chord$(\frac{3}{4}°)$. From this point, Ptolemy was able to get good approximations of the chord function for increments of $\frac{1}{2}°$ by deriving the inequality $\frac{2}{3}\text{chord}(1\frac{1}{2}°) \leq \text{chord}(1°) \leq \frac{4}{3}\text{chord}(\frac{3}{4}°)$ and then calculating chord$(\frac{1}{2}°)$ based on this approximation for chord$(1°)$.

Such was trigonometry in the days of Ptolemy. But what of modern trigonometry? Can we recapture the geometric roots of the subject using only our basic toolbox of Euclidean geometry facts? Indeed we can! The Similar Triangles Theorem is exactly the fact we need to justify the definitions of the more customary *sine* and *cosine* functions.

**Definition.** For $0° \leq \theta \leq 90°$ we define the functions $\sin(\theta)$ and $\cos(\theta)$ by the following rules.

- $\sin(0°) = 0$, $\cos(0°) = 1$, $\sin(90°) = 1$, and $\cos(90°) = 0$.
- If $0° < \theta < 90°$ then $\sin(\theta) = a/c$ and $\cos(\theta) = b/c$ for any right triangle $ABC$ with right angle $\angle C$ and with $m\angle A = \theta$.

This definition is valid because the ratios $a/c$ and $b/c$ do not depend on what triangle is used for $ABC$. If both $ABC$ and $A'B'C'$ are right triangles with right angles at $C$ and $C'$ and $m\angle A = m\angle A' = \theta$, then $a/c = a'/c'$ and $b/c = b'/c'$ by the Similar Triangles Theorem.

And though these definitions are restricted to angle measures between $0°$ and $90°$, we can easily extend them to $180°$ as follows.

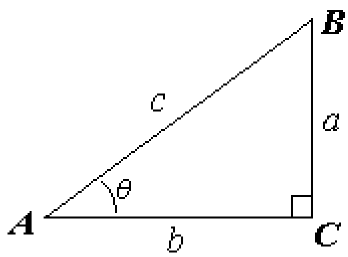



Figure 2.20: $\sin(\theta) = a/c$ and $\cos(\theta) = b/c$

**Definition.** For $90° < \theta \leq 180°$ we define $\sin(\theta)$ and $\cos(\theta)$ by using the above definition in conjunction with the rule that $\sin(\theta) = \sin(180° - \theta)$ and $\cos(\theta) = -\cos(180° - \theta)$.

Keeping these geometric roots in mind, some of the otherwise-unmotivated aspects of trigonometry come to life. To illustrate this, consider the well-known "Law of Sines". The usual statement merely equates the three mysterious fractions $a/\sin(m\angle A)$, $b/\sin(m\angle B)$, and $c/\sin(m\angle C)$; the geometry at the theorem's source, however, indicates that the common value of these ratios is actually a very natural number associated with the triangle $ABC$.

**The Law of Sines.** *For any triangle* $\triangle ABC$ *we have* $a/\sin(m\angle A) = b/\sin(m\angle B) = c/\sin(m\angle C)$, *with each of these ratios being equal to the diameter length of the circle through* $A$, $B$, *and* $C$.

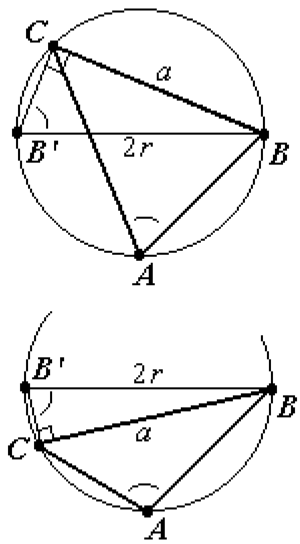



Figure 2.21:

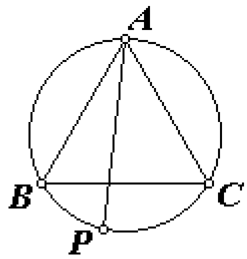
*Proof:* Let $r$ be the radius of the circle $\Gamma$ through $A$, $B$, and $C$. We need only prove that $a/\sin(m\angle A) = 2r$ (since similar arguments would show that $b/\sin(m\angle B)$ and $c/\sin(m\angle C)$ are likewise equal to $2r$). Refer to Figure 2.21 for illustration in these steps.

- We may assume that $BC$ is not a diameter of $\Gamma$, for otherwise $\angle A$ is a right angle by the Inscribed Angle Theorem (or Theorem of Thales), meaning $a/\sin(m\angle A) = a/\sin(90°) = a/1 = a = |BC| = 2r$.
- Let $B'$ be the point so that $BB'$ is a diameter of $\Gamma$.
- If $A$ and $B'$ are on the same side of $\overleftrightarrow{BC}$ then $m\angle A = m\angle B'$ by the Inscribed Angle Theorem. Otherwise, $m\angle A = 180° - m\angle B'$ by the Inscribed Angle Theorem. (Both cases are depicted in Figure 2.21.)
- In either case, then, we have $\sin(m\angle A) = \sin(m\angle B') = a/2r$, so the conclusion follows immediately. $\square$

## Exercises

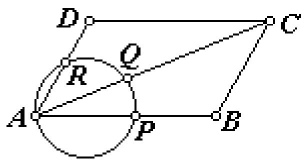
**2.37**. Use a dynamic geometry software package to construct an illustration of Ptolemy's Theorem. Your illustration should allow you to move the vertices of $ABCD$ around a circle.

**2.38**. Prove that if $ABCD$ is a cyclic quadrilateral if and only if $m\angle A + m\angle C = m\angle B + m\angle D = 180°$.




**2.39**. Let $\triangle ABC$ be an equilateral triangle with its vertices on a circle $\Gamma$, and let $P$ be a point on $\Gamma$ not on arc $\overset{\frown}{CAB}$. Use Ptolemy's Theorem to prove that $|AP| = |BP| + |CP|$.

**2.40**. Generalizing Exercise 2.39, let $\triangle ABC$ be an isosceles triangle with top vertex $A$ and all its vertices on a circle $\Gamma$, and let $P$ be a point on $\Gamma$ not on arc $\overset{\frown}{CAB}$. Prove that $|AP|/(|BP| + |CP|) = |AC|/|BC|$.




**2.41**. Suppose $ABCD$ is a parallelogram, $\Gamma$ is a circle through $A$, and suppose $\Gamma$ meets $AB$, $AC$, and $AD$ at points $P$, $Q$, and $R$ respectively. Use Ptolemy's Theorem to prove $|AC||AQ| = |AB||AP| + |AD||AR|$.

**2.42**. Give an expression for chord($\theta$) in terms of the sine and cosine functions for $0° \le \theta \le 180°$.

**2.43**. Prove Corollary 2.16.

**2.44**. Use Corollaries 2.15 and 2.16 to evaluate the following:

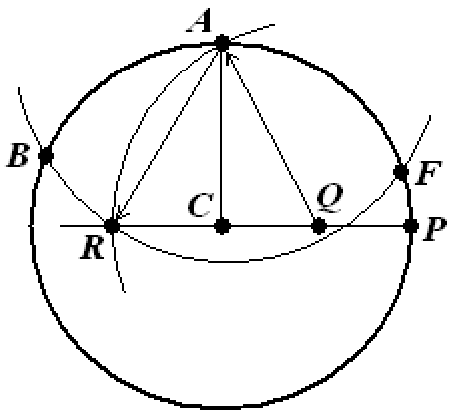
**(a)** chord($45°$)

**(b)** chord($27°$)

**(c)** chord($1\frac{1}{2}°$)

**(d)** chord($4\frac{1}{2}°$)

**2.45**. Let $\Gamma$ be a circle with radius 1 and center $C$, and let $A$ be a point on $\Gamma$. We wish to construct (using only straightedge and compass) points $B$, $D$, $E$, and $F$ on $\Gamma$ so that $ABDEF$ is a regular pentagon. Consider the following construction.




- Construct the perpendicular to $\overleftrightarrow{CA}$ through $C$ and let $P$ be a point where this line meets $\Gamma$.
- Construct the midpoint $Q$ of radius $CP$.
- Construct the circle with center $Q$ and passing through $A$, and let $R$ be the point common to this circle and $\overleftrightarrow{CP}$ so that $C$ is on segment $RQ$.
- Construct the circle centered at $A$ and passing through $R$.
- The points where this last circle meets $\Gamma$ will be vertices $B$ and $F$ of the pentagon.
- Vertices $D$ and $E$ may be found by intersecting $\Gamma$ with the circles through $A$ with centers at $B$ and $F$.

In this exercise we use Ptolemy's trigonometry to verify that this construction produces a regular pentagon.

**(a)** Determine the length $|AB|$ from the steps in the construction.

**(b)** Clearly the construction works if and only if the arc subtended by $AB$ measures $72°$ (one-fifth of $360°$). Show that this is the case by demonstrating that your answer to part (a) agrees with the value chord($72°$). (To find chord($72°$), use Corollary 2.16 and the known value for chord($36°$).)

**2.46**. Prove that the identity $\sin^2(\alpha) + \cos^2(\alpha) = 1$ holds for the functions sine and cosine as we have defined them.

**2.47**. Prove that if $\triangle ABC$ is any triangle then $\text{area}(ABC^\circ) = \frac{1}{2}\, bc \sin(m\angle A)$.

**2.48**. Let $\triangle ABC$ be a triangle and let $r$ be the radius of the circle through $A$, $B$, and $C$. Use the Law of Sines to prove $\text{area}(ABC^\circ) = abc/4r$.

# F. Heron's Formula

We devote this section to a single theorem. It's proof is extremely clever, and equally *non*-transparent. In fact, it seems difficult to imagine how someone would stumble upon this proof – it wanders (to all appearances, without aim) through a maze of similar triangle computations until without warning the pieces fit tidily together to give the result.

It is to Heron of Alexandria that we owe this gem. Though some of his writings have survived to the present, we know almost nothing about him. Not even the century in which he lived is certain, though the first century AD seems most likely. Many of Heron's other writings are rather mundane, but this one theorem (with its dazzling proof) is certainly enough to rank him among the most clever of the ancient mathematicians.

Recall that the usual formula for the area of a triangle is "one-half the base times the height with respect to that base". Of course, the height of a triangle is a somewhat awkward notion since it requires knowing a point (the foot of the perpendicular from the third vertex to the base) that is not immediately given as part of the triangle. Heron's Formula gives the area of an arbitrary triangle in terms of lengths that *are* easily measured and immediately available: its three side lengths. It should not be surprising that there would be such a formula, for we know (from the SSS criterion) that the three side lengths determine the triangle's congruence class, and thus also determine its area. What *is* striking about Heron's Formula is its simplicity.

**Heron's Formula.** *The area of any triangle $ABC$ is given by the formula* $\text{area}(ABC) = \sqrt{s(s-a)(s-b)(s-c)}$ *where $s$ (called the semiperimeter) is $\frac{1}{2}(a+b+c)$.*

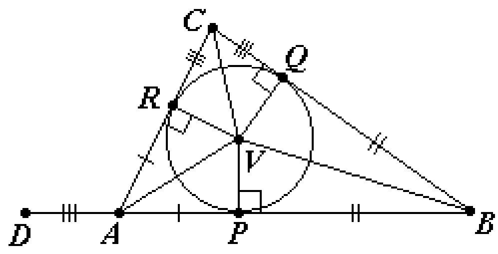



Figure 2.22:

*Proof:* Let $V$ be the incenter[2] of $ABC$ and let $P$, $Q$, and $R$ be the points of tangency of the sides of the triangle with the inscribed circle, as in Figure 2.22. If $r$ is the radius of this inscribed circle, then it is easy to prove (see Exercise 2.52) that $\text{area}(ABC) = sr$. In what follows we will prove that $sr = \sqrt{s(s-a)(s-b)(s-c)}$.

- First, note that $\triangle AVP \cong \triangle AVR$ (by the SAS criterion), so $|AP| = |AR|$.

[2]That is, $V$ is the intersection of the angle bisectors for $ABC$ – see p.57.

- Similarly, $|BP| = |BQ|$, and $|CQ| = |CR|$.
- Construct the point $D$ as in the figure, so that $A$ is on segment $DB$ and $|AD| = |CQ| = |CR|$.
- Note that

$$\begin{aligned} |DB| &= |AD| + |AP| + |BP| \\ &= \frac{1}{2}(2|AD| + 2|AP| + 2|BP|) \\ &= \frac{1}{2}([|CR| + |CQ|] + [|AP| + |AR|] + [|BP| + |BQ|]) \\ &= \frac{1}{2}(a+b+c) \quad = \quad s \end{aligned}$$

- Also,

$$\begin{aligned} s - a &= |DB| - |BC| \\ &= (|AD| + |AP| + |BP|) - (|BQ| + |CQ|) \\ &= (|CQ| + |AP| + |BQ|) - (|BQ| + |CQ|) \\ &= |AP| \,. \end{aligned}$$

- Similarly, $s - b = |BP|$ and $s - c = |AD|$
- Now construct the lines perpendicular to $AB$ through $A$ and perpendicular to $BV$ through $V$, and let $E$ be the point at which these lines meet. (See Figure 2.23.)
- Construct the circle with diameter $EB$. Then since $\angle EAB$ and $\angle EVB$ are right angles, $A$ and $V$ are on this circle. [This converse to the Theorem of Thales is easy to check: the ray $\overrightarrow{EA}$ will meet this circle at some point, say $A'$. But then $EA'B$ is a right triangle with right angle $\angle A'$ by the Theorem of Thales, so $\triangle EA'B \cong \triangle EAB$ by SAA (they share hypotenuse $EB$ and angle $\angle E$). Thus on ray $\overrightarrow{EA}$ we have $|EA| = |EA'|$, so $A = A'$. Thus $A$ is on the given circle, and similarly so is $V$.]

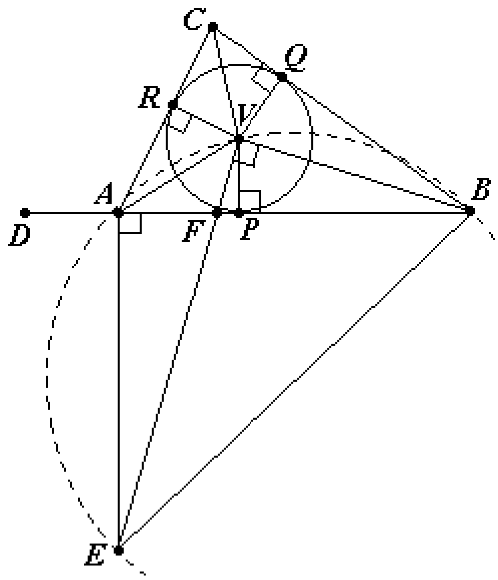



Figure 2.23:

- So $EAVB$ is a cyclic quadrilateral, so by Exercise 2.38 we have $m\angle AVB + m\angle BEA = 180°$.
- But (as we noted above) $\triangle AVP \cong \triangle AVR$, $\triangle BVP \cong \triangle BVQ$, and $\triangle CVQ \cong \triangle CVR$. Thus,

$$\begin{aligned} 2m\angle AVP + 2m\angle BVP + 2m\angle CVR &= 360° \\ m\angle AVP + m\angle BVP + m\angle CVR &= 180° \\ m\angle AVB + m\angle CVR &= 180° \end{aligned}$$

- These last two equations together imply $\angle BEA \cong \angle CVR$.

We now apply the Similar Triangles Theorem to three similar pairs of triangles.

- $\triangle BEA \sim \triangle CVR$, so $\dfrac{|AB|}{|RC|} = \dfrac{|AE|}{|RV|}$, or $\dfrac{|AB|}{|AD|} = \dfrac{|AE|}{r}$.
- $\triangle AFE \sim \triangle PFV$, so $\dfrac{|AE|}{|PV|} = \dfrac{|AF|}{|PF|}$, or $\dfrac{|AE|}{r} = \dfrac{|AF|}{|PF|}$.
- Combining these last two steps, we get

$$(*) \quad \frac{|AB|}{|AD|} = \frac{|AF|}{|PF|}\,.$$

- $\triangle PFV \sim \triangle PVB$, so $\dfrac{|PF|}{|PV|} = \dfrac{|VP|}{|BP|}$, or

$$(**) \quad |PF||BP| = |PV|^2 = r^2\,.$$

Beginning with equation $(*)$, we have:

$$\begin{aligned} \frac{|AB|}{|AD|} + 1 &= \frac{|AF|}{|PF|} + 1 \\ \frac{|AB| + |AD|}{|AD|} &= \frac{|AF| + |PF|}{|PF|} \\ \frac{s}{s-c} &= \frac{s-a}{|PF|} \\ \frac{s^2}{s(s-c)} &= \frac{(s-a)|BP|}{|PF||BP|} \\ &= \frac{(s-a)(s-b)}{r^2} \qquad \text{by equation } (**) \end{aligned}$$

so $r^2 s^2 = s(s-a)(s-b)(s-c)$, completing the proof. □

### Exercises

**2.49**. Find the area of a triangle with side lengths 4, 5, and 6.

**2.50**. Use Heron's Formula to find the height of a triangle with side lengths 1, 2, and $2\sqrt{2}$ with respect to the longest side.

**2.51**. Repeat Exercise 2.50 without using Heron's Formula.

**2.52**. Prove the following generalization of the fact (used in the proof of Heron's Formula – see p.82) that area$(ABC) = sr$: if a simple polygon $P_1P_2 \cdots P_n$ has an inscribed circle then the area of the polygon equals the product of its perimeter and the radius of the inscribed circle.

**2.53**. Show (using some algebra!) that Heron's formula can be expressed as

$$\text{area}(ABC) = \frac{1}{4}\sqrt{[(a+b)^2 - c^2][c^2 - (a-b)^2]}\,.$$

## G. Additional Exercises

We conclude this chapter with a set of exercises you may return to occasionally as you continue through the chapters that follow. The topics here are varied, and the problems range in difficulty from easy to challenging. But all of them can be done using only our basic toolbox theorems.

### Exercises

**2.54**. Prove that a 4-gon is a parallelogram if and only if its diagonals intersect at their midpoints.

**2.55**. Prove that $|AB| = |BC|$ and $|CD| = |DA|$ if and only if $AC \perp BD$. Use this to conclude that a parallelogram is a rhombus if and only if its diagonals are perpendicular.

**2.56**. Prove that $ABCD$ is a parallelogram if and only if $AB \cong CD$ and $BC \cong AD$.

**2.57**. Let $ABCD$ be any simple 4-gon, and let the midpoints of $AB$, $BC$, $CD$, and $DA$ be $P$, $Q$, $R$, and $S$ respectively. Prove that $PQRS$ is a parallelogram.

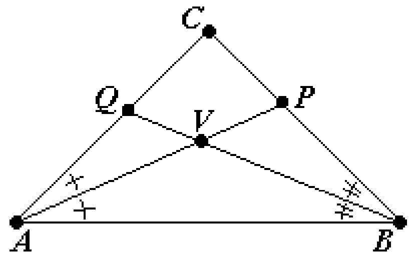
**2.58**. Let $\triangle ABC$ be an isosceles triangle with top vertex $A$ and let $D$ be a point on $AB$ such that $\triangle ADC$ is isosceles with top vertex $D$ and $\triangle BCD$ is isosceles with top vertex $C$. Find $m\angle A$.

**2.59**. Suppose $A$, $B$ and $C$ are distinct noncollinear points. Let $D$ be the midpoint of $AB$ and let $E$ be the midpoint of $BC$. Let $\Gamma_1$ be the circle through $A$ with center $D$ and let $\Gamma_2$ be the circle through $B$ with center $E$. Prove that $\Gamma_1$ and $\Gamma_2$ intersect at a point on $AC$.

**2.60**. Prove that a circle may be inscribed in the convex quadrilateral $ABCD$ if and only if $|AB| + |CD| = |BC| + |AD|$.

**2.61**. Let $A$, $B$, and $C$ be noncollinear points. Prove that the perpendicular bisectors of the three sides of triangle $ABC$ meet at a common point (called the *circumcenter* of $ABC$), and that the circle through $A$, $B$, and $C$ has its center at this point.

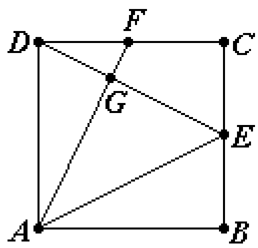
**2.62**. Suppose in triangle $\triangle ABC$ that $AP$ and $BQ$ are angle bisectors at $A$ and $B$ meeting at the incenter $V$. Prove that $m\angle AVB = 90° + \frac{1}{2}m\angle C$.




**2.63**. Let $ABC$ be a triangle and $P$ be the point on $BC$ so that $AP$ is the angle bisector of $\angle A$. Prove: $\dfrac{|AB|}{|BP|} = \dfrac{|CA|}{|PC|}$.

**2.64**. Let $ABC$ be a right triangle with right angle at $C$ and let $ABDE$ be a square with $D$ and $E$ on the side of $\overleftrightarrow{AB}$ opposite $C$. Let $F$ be the center of this square. Find $m\angle ACF$ and prove your answer.

**2.65**. In the figure at right, $ABCD$ is a square with $E$ and $F$ midpoints of the sides $BC$ and $CD$, and $G$ is the point of intersection of $AF$ and $DE$. Prove that $AEG$ is a right triangle with side lengths having ratio 3:4:5.



**2.66**. Suppose $ABCD$ is a rectangle of area 12. Let $E$ be the midpoint of $AB$ and $F$ the intersection of $CE$ and $BD$. Find the area of quadrilateral $AEFD$.

**2.67**. Let $ABCD$ be a square and let $P$ be its center. Let $\Gamma$ be the circle with center at $P$ and radius $|AB|$. Let $PQRS$ be a square with $Q$ on $\Gamma$. Prove that the area of the overlap of the two squares does not depend on where $Q$ is placed on $\Gamma$.

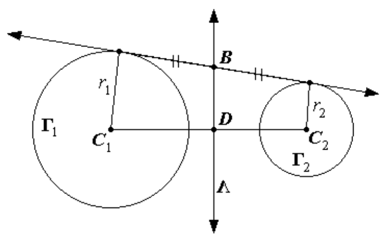
**2.68**. Given a circle $\Gamma$ and a point $P$ outside of $\Gamma$, give a straightedge and compass construction for a line through $P$ that is tangent to $\Gamma$. (*Hint:* use the Inscribed Angle Theorem on a circle through both $P$ and the center of $\Gamma$.)

**2.69**. Given two circles $\Gamma_1$ and $\Gamma_2$ not intersecting and neither inside the other, give a straightedge and compass construction for a line tangent to both circles. (See the previous exercise.)

**2.70**. As in the figure below, let $\Gamma_1$ and $\Gamma_2$ be disjoint circles, neither inside the other, and let their centers and radii be $C_1$ and $r_1$, $C_2$ and $r_2$ respectively. Let $A_1$ be a point of $\Gamma_1$ and $A_2$ a point of $\Gamma_2$ such that $A_1A_2$ is tangent to both circles. Let $B$ be the midpoint of $A_1A_2$. Let $\Lambda$ be the line through $B$ perpendicular to $\overleftrightarrow{C_1C_2}$ and let $D$ be the point at which $\Lambda$ intersects $\overleftrightarrow{C_1C_2}$.

**(a)** Prove that $|C_1D|^2 - |C_2D|^2 = r_1^2 - r_2^2$.

**(b)** Let $P$ be any point on $\Lambda$ and let $E_1$ and $E_2$ be points on $\Gamma_1$ and $\Gamma_2$ respectively such that $PE_1$ is tangent to $\Gamma_1$ and $PE_2$ is tangent to $\Gamma_2$. Use part **(a)** to prove that $|PE_1| = |PE_2|$.

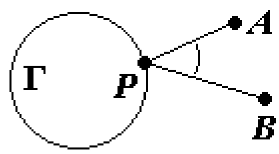



**2.71**. We say that two circles are *orthogonal* if they intersect, and if when $P$ is a point of their intersection then the lines tangent to the two circles at $P$ are perpendicular to each other. Let $\Gamma_1$, $\Gamma_2$, and $\Gamma_3$ be three circles not intersecting each other, no one of them inside another one, and with the three centers being noncollinear. Give a straightedge and compass construction for a circle that is orthogonal to $\Gamma_1$, $\Gamma_2$, and $\Gamma_3$. (*Hint:* see Exercises 2.68 and 2.70.)

**2.72**. Suppose $ABC$ is a triangle with medians $AP$, $BQ$, and $CR$ and centroid $V$ such that $AP \perp BQ$. Prove that $CV \cong AB$.

**2.73**. Given noncollinear points $A$, $B$, and $C$, give a straightedge and compass construction for points $D$, $E$, and $F$ such that $A$, $B$, and $C$ are the midpoints of the sides of $\triangle DEF$.

**2.74**. Let $A$ and $B$ be points outside the circle $\Gamma$. Give a straightedge and compass construction to find the point $P$ on $\Gamma$ for which $m\angle APB$ is maximum. (*Hint:* consider circles passing through both $A$ and $B$.)



# Chapter 3

# The Axiomatic Method

The material from Chapter 2 should have you convinced that the basic theorems of Euclidean geometry – those that you learned in high school and that we reviewed in Section 1B – have impressive power and usefulness. It should now seem worthwhile to spend some effort proving them from scratch. Before we start that project, though, we'll spend this chapter considering the method by which we'll work: the *axiomatic method*.

- Section A treats the nuts and bolts of the axiomatic approach. We'll discuss axiom systems and their workings, and introduce the concept of a geometry as a mathematical structure.
- In Section B we put the abstract concepts from Section A to practice with a brief excursion in finite geometries.
- Section C then traces the history of the axiomatic method relative to Euclidean geometry. We'll evaluate Euclid's *Elements* as an axiom system, and we'll outline the progress of more modern treatments. This will include a brief account of the discovery of "non-Euclidean" geometry and a discussion of the relationship between neutral, Euclidean, and hyperbolic geometries.

The material here marks a step upward in abstraction and formality (though not necessarily in difficulty). The ideas may take some getting used to, and careful reading and re-reading will be helpful. Our goal is that by the end of this chapter you will have an understanding of how the axiomatic method works and why it is important, as well as a working knowledge of the terminology used to employ it. This will be critical to understanding the overall goals and process in the chapters that follow.

# A. Axiom Systems and Geometries

The *axiomatic method* is a framework for setting forth mathematical knowledge in a systematic, organized, and logically sound way. Though we will be applying it to the study of geometry, it is not limited to that field. Any mathematical discipline can be set axiomatically, and in fact, since the early 20th century the axiomatic method has been *the* way that mathematics is done.

The rationale for the method is simple: statements can only be proved by using other statements which in turn would be proved by yet other statements. To avoid an infinite regression of proofs, we establish a small set of fundamental facts called *axioms* that are accepted as true without proof. We then set out to prove theorems as a consequence of these assumed facts. The ordering of the theorems is crucial, for the proof of Theorem 1 can call on nothing except the axioms, but once Theorem 1 is so proved, we can use its conclusion in proving Theorem 2, and so on.

The exact origins of this process are difficult to pin down. Certainly it could be said that elements of it were present in what Thales is reported to have done, for he is credited with being the first to order mathematical theorems so that their conclusions build on each other from the simple to the more complex. The axiomatic method made considerable progress among the Phythagoreans, perhaps motivated by the aftermath of incommensurability. Recall (see p.32) that when the Pythagorean belief in the commensurability of all quantities was disproved, all of the mathematical proofs they had constructed using that assumption collapsed. It suddenly became critical to know exactly when that assumption was used. The difficult fallout that ensued must certainly have demonstrated the need to specify and track the use of all assumptions in proofs.

These early examples highlight two advantages to the method that remain important today:

- It avoids logical circularities (using a statement to justify its own proof – see Exercise 3.2) by insisting on an orderly development upward from fundamental observations.
- It facilitates the tracking of assumptions. This may seem unimportant now, but by Chapters 5 and 6 it will be as important to us as it was to the Pythagoreans after incommensurability!

### Axiom systems

The axiomatic approach received its greatest boost from Euclid's *Elements*, for that work was so successful in its application of the method that all previous attempts were forgotten. We'll take a closer look at what Euclid did in Section C, but for now, consider the task that faced him. He had a body of knowledge before him – what we now call *Euclidean geometry* – that he wished to describe axiomatically. What would be involved in this task?

- First, he needed to set forth the exact meaning of the terms used in the subject matter.
- Second, he needed to somehow root out a small set of assumptions about those terms that were together capable of providing proofs for all of the theorems he wished to establish.
- Lastly, he needed to order the theorems in such a way that their proofs could be accomplished in turn.

Such a framework for a body of mathematical knowledge is called an *axiom system*. The actual definition is very much parallel to the informal steps outlined above.

DEFINITION. An *axiom system* consists of four things:

**1.** A set of *undefined terms* – words whose meaning we agree will be understood without definition.

**2.** A set of *defined terms* – words with accompanying definitions based on the undefined terms.

**3.** A set of axioms concerning the properties of the defined and undefined terms.

**4.** A listing of statements (called the *theorems* of the axiom system) about the defined and undefined terms such that Theorem 1 can be proved using only the axioms, and for $n \geq 2$, Theorem $n$ can be proved using only the axioms and Theorems 1 through $(n-1)$.

The undefined terms and defined terms together make up the *terms* of the system, and the axioms and theorems together make up its *valid statements*.

The undefined terms are to the terms of the axiom system what the axioms are to its valid statements. They form the foundation of words that are understood without definition, and on which the other needed terms can have their definitions constructed. In geometry, it is typical to leave terms such as *point*, *line*, and *plane* undefined, as also terms describing set membership such as *in*, *on*, or *through* (as in "a point *on* a line" or "a line *through* a point"). To illustrate the process, if we have undefined terms *point*, *line*, and *through*, we can then give a meaningful definition to the word *parallel*: two **lines** are *parallel* if they do not pass **through** a common **point**.

Note that we include the ordering of the theorems as part of the axiom system. Indeed, creating this ordering is one of the most crucial parts of applying the axiomatic method. For even with a well-chosen set of axioms, finding an ordering of the theorems in which they build naturally, efficiently, and elegantly on each other is truly an art!

But of course, axiom systems might differ in more than just the order their theorems are presented. If the terms are different for two axiom systems, not much can be said to compare them. For instance, an axiom system for Euclidean geometry would not be comparable in any way to an axiom system for the algebra of matrices – they deal with completely different objects. But supposing we keep the terms constant, how can we compare axiom systems built on different sets of axioms for those terms? This is a critical question in geometry where it is possible to give many different sets of axioms on the standard geometric terms.

The set of theorems we can prove in an axiom system is determined entirely by the axioms we choose. If we do not include enough axioms, we may not be able to prove all of the theorems we would like to have in our set of valid statements. But, of course, different sets of axioms *can* result in the same set of valid statements, for what is an axiom in one system might be proved as a theorem by a different set of axioms. It seems reasonable to consider two such axiom systems to be in some sense "the same" if they result in the same set of valid statements. This is the motivation for the next definition.

**Definition.** Two axiom systems are *equivalent* if they have the same terms and the same valid statements.

Let's emphasize this point: equivalent axiom systems need not have the same set of axioms! The axioms of system $\mathcal{A}_1$ may show up as theorems in system $\mathcal{A}_2$ and vice-versa. But if the union of axioms and theorems in $\mathcal{A}_1$ equals the union of axioms and theorems in $\mathcal{A}_2$ (and they share a common set of terms) then the two systems are equivalent.

If two axiom systems are equivalent, which should be preferred? While the ultimate answer probably depends on the purposes one has in mind, according to strict logic one would prefer an axiom system that included as few axioms as possible. As Aristotle stated a generation before Euclid, "All other things being equal, that proof is the better which proceeds from the fewer postulates".

A special case of this principle occurs if one of the axioms in an axiom system can be proved from the others. We can then improve the axiom system (by Aristotle's criterion) by changing the status of that one statement from "axiom" to "theorem". Repeating the process as many times as needed, we might hope to arrive at a set of axioms in which no axiom is a logical consequence of the others.

**Definition.** We say that an axiom in an axiom system is *independent* of the other axioms if that axiom cannot be proved as a consequence of the other axioms. If all axioms in the system are independent of each other then we say that the axiom system itself is *independent*.

Note that an axiom system has this property if and only if the removal of any of its axioms diminishes the set of valid statements.

In practice we do not always insist on independence, for it is a tremendous amount of work to build up significant results from such a sparse set of assumptions. Pedagogically, it is usually better to strike a balance by choosing axioms that are elementary enough to be believable as "self-evident" facts, yet powerful enough to prove the principal theorems without too much labor. This is the approach we will use.

One property of axiom systems that we *will* insist on (for obvious reasons) is *consistency*, defined here:

**Definition.** An axiom system is *inconsistent* if there is a statement X such that both X and ∼X may be proved from its axioms. A system that is not inconsistent is said to be *consistent*.

Clearly the conclusions drawn from an inconsistent axiom system are suspect! In fact, it can be shown that in an inconsistent system *any* statement about the terms can be proved both true *and* false.

The independence property says that an axiom system does not contain "too many" axioms. At the other extreme, we certainly want to avoid having "too few" axioms. The natural barometer is that we have too few axioms if we cannot prove the theorems we wish to prove. The ideal to which we could aspire is *completeness*, defined as follows:

**Definitions.** Suppose that $\mathcal{A}$ is an axiom system. If X is a statement about the terms of this system such that neither X nor ∼X can be proved by the axioms of $\mathcal{A}$ then we say that X is *undecidable* in $\mathcal{A}$. If $\mathcal{A}$ has no undecidable statements then we say that $\mathcal{A}$ is *complete*.

So an undecidable statement is one that can neither be proved nor disproved by the axioms of the system, and the system is complete if *no* statement about its terms is undecideable – that is, if *every* statement about its terms is either provably true or provably false. While this is surely a desirable quality, we'll see in Section C that it is usually too much to ask. But these are nevertheless important concepts for us to understand, for the undecideability of certain statements will play a crucial role in Chapters 4 through 6.

**Not one geometry, but many *geometries*.**

You are probably accustomed to the word *geometry* being used as a noun describing a subject of study. But it has another use as well: a noun that names a mathematical structure, much like *vector space* or *function*. And in this sense of the word, it makes sense to form the plural *geometries*, for just as there are many different varieties of function, the mathematical species of geometry comes in many varieties as well. We've already mentioned the two main geometries with which this book is concerned, namely Euclidean geometry and hyperbolic geometry (though we haven't said much about the latter one so far) – but there are many others, and we'll meet a few of them in this chapter.

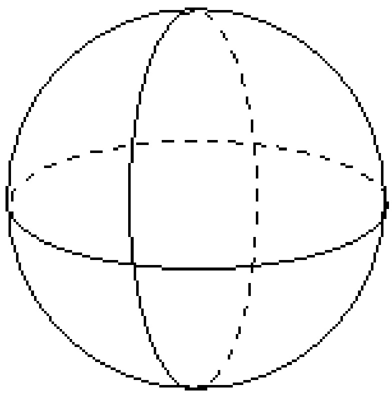


Figure 3.1: "Lines" in spherical geometry

But first, let's consider putting a meaning to the word *geometry* in the sense of a mathematical structure. What *is* a geometry? The answer lies in the framework of axiom systems. Recall that the set of *valid statements* in an axiom system consists of all the axioms of the system together with all of the theorems provable from those axioms – in short, it is the collection of statements that are "true" in the context of that axiom system. If the terms of the axiom system include the typical geometric objects of *point* and *line*, then the set of valid statements constitute a set of geometric statements that are "true" under some set of assumptions. We say then that the set of valid statments makes up a *geometry*. In short, a *geometry* is a collection of statements that constitutes the set of valid statements for some geometric axiom system.[1]

[1]More formally, we could consider a *geometry* to be an equivalence class of axiom systems under the notion of equivalence that we introduced on p.92.

This might all make more sense after an example. We'll now introduce one of the "other geometries" that is particularly easy to think about: *spherical geometry.* Think of a sphere (the surface of a ball) and suppose for a moment that our universe is this sphere. (That shouldn't be too hard, as you in fact live on such a sphere!) Now, what is a "line" in such a universe? Our intuition says that a line should always give the shortest path between two points. A moment's reflection, then, should convince you that a "line" in our spherical universe, if continued as far as possible, will actually be a *great circle* of the sphere – that is, a circle whose center is the center of the sphere – see Figure 3.1. On a globe, the equator is thus a "line", as are all the longitudinal circles that pass through both poles. The latitudinal circles however (other than the equator) are *not* "lines" because their centers are not the sphere's center. Indeed, the shortest distance between Salt Lake City and Madrid is not along a latitudinal circle (even though the two cities are at approximately the same latitude) but rather along a great circle that goes "over the pole" – the route an airplane would take.

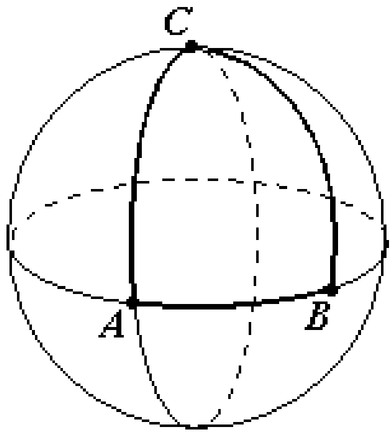


Figure 3.2: A triangle in spherical geometry

*Spherical geometry* is the set of facts that hold for this universe. It isn't hard to see that it's a very different geometry from Euclidean geometry, for many of the standard facts of Euclidean geometry do not hold here. For example, Figure 3.2 shows a triangle in spherical geometry that clearly has an angle sum greater than 180° – two of its angles are right angles! We conclude, then, that the 180° Sum Theorem is *not* part of spherical geometry. Even more simply, the statement "*Every pair of two distinct lines intersect in exactly two points*" is included in spherical geometry, but is clearly not true in Euclidean geometry.

Now, once we have a geometry in mind, how do we describe it? The answer, of course, is "*with an axiom system*!" In fact, there will likely be a choice of many axiom systems to describe a geometry, for any two equivalent axiom systems will give rise to the same geometry (since they will have the same set of valid statements). Choosing a good axiom system to describe a geometry is exactly the task that Euclid faced, as we described it at the beginning of this section.

**Models for a geometry**

We now have a formal description of what Euclid was trying to do: to produce an axiom system that would yield the geometry he had in mind (namely that of straight lines in a two-dimensional plane). Euclid was clearly using that mental image of what his lines and plane were like to motivate the task. In fact, the task was to make the abstract construction (the axiom system) capture the essential properties of the objects in the mental image.

In modern terminology for the axiomatic method, Euclid was using a *model* to motivate his geometry. We were doing the same thing above to motivate our concept of spherical geometry. A *model* for a geometry is formally a pair $(\Pi, \mathcal{L})$ where $\Pi$ is a set (that we refer to as the *underlying set*, or sometimes as the *plane*, even if it does not match our preconceived notion of what a "plane" is) and $\mathcal{L}$ is a collection of subsets of $\Pi$ that are designated as *lines*. To be a correct model for the gometry in question, the lines in $\mathcal{L}$ should behave according to the valid statements of the geometry. If the terms of the geometry include notions such as angle measure and distance, these must be accounted for in the model as well. The most familiar model for Euclidean plane geometry is to let $\Pi$ be the Cartesian ($xy$-coordinate) plane $\mathbb{R}^2 = \{(x, y) : x, y \in \mathbb{R}\}$ and $\mathcal{L}$ be the collection of solution sets to all linear equations $ax + by = c$, with distance and angle measure defined as they usually are in coordinate analytic geometry. This is the model that motivates Euclidean geometry, and an axiom system is said to describe Euclidean plane geometry only if the properties it describes are exactly the properties of the objects in this model.

But this is not the only model for Euclidean geometry! Let $\Sigma$ be the set of points on a sphere (the surface of a ball) and let $N$ be the "north pole" of this sphere. We can define a new model for Euclidean geometry, which we will call the *punctured sphere model*, by letting the underlying point set be $\Pi = \Sigma \setminus \{N\}$ (that is, all points of the sphere except for its north pole) and letting the "lines" be the intersections of this set with two-dimensional planes through the point $N$ (see Figure 3.3). It is possible to complete this model by describing how one would measure angles between intersecting "lines" from $\mathcal{L}$ and how one would measure the distance between points on the sphere. (As you might guess, distance in this model cannot just be the usual distance on the sphere – for otherwise our "lines" would not have infinite length. Instead, $N$ plays the role of a "point at infinity", and distances become increasingly "stretched" as we draw closer to $N$ on the sphere. See Exercise 3.5.) Suffice it to say that we can define all of the terms of this new model so that all the geometric properties of the Cartesian plane model are also valid on the punctured sphere. The new model associates the terms of Euclidean geometry with different-looking objects, but all of the theorems of Euclidean geometry remain valid when applied to those objects.

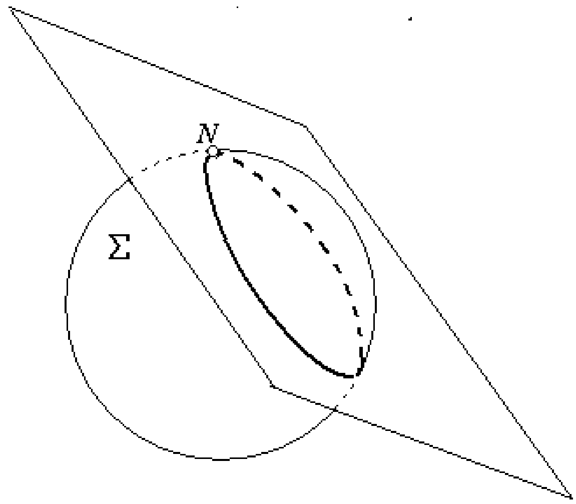



Figure 3.3: A "line" in the punctured sphere model

But surprisingly, there is not so much difference as it might seem between the Cartesian plane model and punctured sphere model. To see this, consider the function $f$ from the punctured sphere to the Cartesian plane illustrated in Figure 3.4. Here we have placed the sphere $\Sigma$ so that it is resting on the Cartesian $xy$-plane, with its south pole $S$ touching the plane at the origin $(0, 0)$. If $P$ is a point of $\Sigma \setminus \{N\}$ we define $f(P)$ to be the point where the ray $\overrightarrow{NP}$ intersects the Cartesian plane. Note that if $\Lambda$ is the "line" on the punctured sphere given by the intersection of the sphere and a plane $\Pi$ through $N$, then $f$ carries the points of $\Lambda$ to the line of intersection between the plane $\Pi$ and the Cartesian $xy$-plane. Thus, "lines" in the punctured sphere model are carried by $f$ to lines in the Cartesian plane model. We describe this by saying that $f$ *preserves lines.* In fact, the function $f$ is a *one-to-one* and *onto* correspondence between the underlying sets for the two models that preserves all aspects of the geometry.[2]

[2]This correspondence is actually the key to defining distance and angle measure in the punctured sphere model. We simply transfer the objects from the sphere to the Cartesian plane via this function $f$ and measure the distance or angle there. See Exercise 3.5.

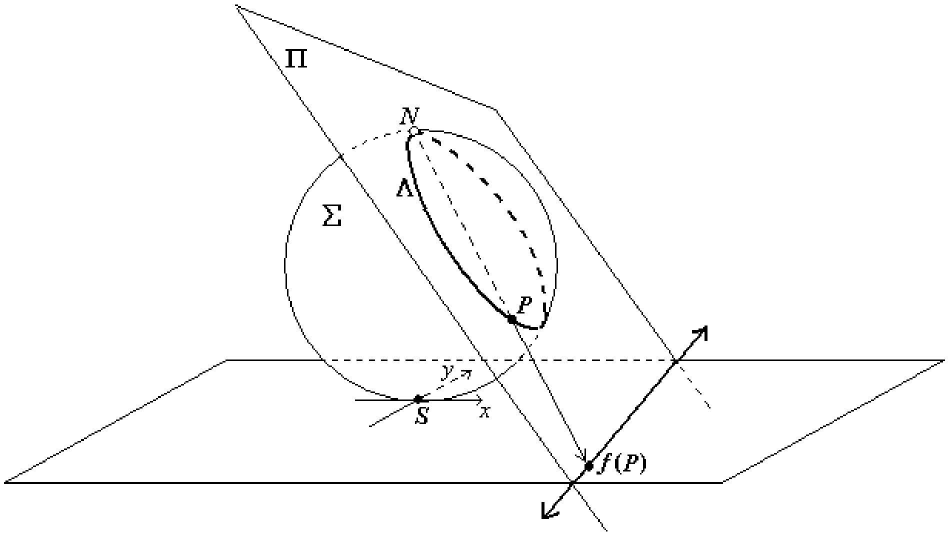



Figure 3.4: The correspondence $f$ between the punctured sphere model and the usual Cartesian plane model

This is an example of an *isomorphism* of models for a geometry. Formally, an isomorphism of models is a one-to-one and onto function between the underlying sets for the models that preserves the lines and other terms of the geometry. Any two models for a geometry that are related in this way are said to be *isomorphic*. Isomorphic models may look very different but they are really essentially the same – the points have just been renamed.

**DEFINITION.** An axiom system for a geometry is *categorical* if any two models for the geometry given by that axiom system are isomorphic.

In other words, a categorical axiom system is one that specifies enough things that it allows essentially only one model (because any two models are isomorphic). This seems to be closely related to "completeness" for an axiom system, because both properties somehow say that an axiom system contains enough information to specify the behavior of all its terms. We'll see shortly (in Section C of this chapter) though, that whereas the categorical property is often achieved, completeness for axiom systems is more-or-less a pipe dream.

### Uses of Models

The relationship between a geometry and a model for that geometry is best understood in terms of *general theory* versus *specific example*. The concept of "square" is general – it is a list of properties that constitute the defining characteristics of what it means to be a square. If I carefully construct a quadrilateral

$ABCD$ on a sheet of paper to meet those characteristics, I have before me a "model" of a square. It is a representation and realization of the abstract concept couched in the definition. As Plato wrote,

> ... I think you also know that although [those who deal with geometrics and calculations] use visible figures and argue about them, they are not thinking about these figures but of those things which the figures represent; thus it is the square in itself and the diameter in itself which are the matter of their arguments, not that which they draw; similarly, when they model or draw objects, which may themselves have images in shadows or water, they use them in turn as images, endeavoring to see those absolute objects which cannot be seen otherwise than by thought. [*Republic*, Book VI]

Though I may use my "model" square to help clarify an argument, I cannot pretend that something true with this particular model (such as the fact that its sides have length 3) will be true for all squares, so I cannot use it as the sole support of a general argument about squares. I certainly could not prove that *all* squares have diagonals of equal length by drawing a square and measuring its diagonals! In the same way, I cannot use a particular model of a geometry to prove theorems in the geometry. The model is merely a specific realization of the abstract concept that is the geometry itself, just as my picture of a square is but one specific realization of a general concept. An argument that relies on one model to prove a general statement about the geometry cannot be trusted, for how do we know that it holds for *all* models?

On the other hand, a specific example is perfectly able to *disprove* a universal statement. For example, a single Euclidean construction of a rectangle with sides of different lengths is enough to prove that the statement "*all rectangles are squares*" is false in Euclidean geometry. Similarly, I can disprove statements in a geometry by showing that their conclusions do not hold true in some valid model. After all, if the axioms of the geometry were really enough to prove statement X true, then any model satisfying the axioms should also satisfy statement X. So, if X fails in the model, then X must not be a correct theorem in the geometry. That is why we could say above (on p.95) that the 180° Sum Theorem is false in spherical geometry: we have a model in which a clear counterexample to the theorem exists, so it cannot be possible to prove the theorem from any set of axioms describing spherical geometry.

If, however, we wanted to prove a general theorem like

> *In spherical geometry the sum of the measures of the three angles in a triangle is greater than* $180°$

then no amount of work with the model would do. We could only prove this universal statement true by arguing from a set of axioms for spherical geometry.

While models are not capable of proving universal statements *within* the geometry, they are exactly the tool we need to prove things *about* the geometry. Two important instances of this are given in the following theorems.

THEOREM **3.1.** *If there exists a model for the geometry described by an axiom system* $\mathcal{A}$ *then* $\mathcal{A}$ *is consistent.*

*Proof:* If a model for the geometry exists, then that model must satisfy all of the axioms in $\mathcal{A}$. But then $\mathcal{A}$ cannot be inconsistent, for if its axioms were contradictory to one another then no model could satisfy them all simultaneously. □

THEOREM **3.2.** *Let* $\mathcal{A}$ *be an axiom system for a geometry, and let the axioms in* $\mathcal{A}$ *be* $\mathrm{X}_0, \mathrm{X}_1, \mathrm{X}_2, \ldots, \mathrm{X}_n$*. Let* $\mathcal{M}$ *be a model satisfying axioms* $\mathrm{X}_1, \mathrm{X}_2, \ldots \mathrm{X}_n$ *but not satisfying axiom* $\mathrm{X}_0$*. Then axiom* $\mathrm{X}_0$ *is independent of the other axioms.*

*Proof:* This is immediate. The model $\mathcal{M}$ shows that it is possible to have statements $\mathrm{X}_1, \mathrm{X}_2, \ldots, \mathrm{X}_n$ hold true without $\mathrm{X}_0$ being true. Thus, $\mathrm{X}_0$ is not a logical consequence of the other axioms. □

## Exercises

**3.1**. True or False? (Questions for discussion)

**(a)** If two axiom systems have the same set of axioms, they are equivalent.

**(b)** Equivalent axiom systems always have the same set of theorems.

**(c)** If the axiom system $\mathcal{A}$ has an undecidable statement, then $\mathcal{A}$ is not complete.

**(d)** A consistent axiom system will always be complete.

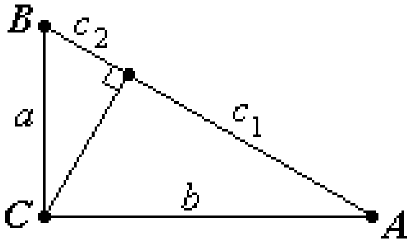
**(e)** An independent axiom system will never be complete.

**(f)** A model for a geometry can never prove or disprove a statement within the geometry.

**(g)** A model can prove that an axiom system is consistent.

**(h)** Models can never prove that an axiom system is independent.

**3.2**. In Chapter 5 we will prove the Pythagorean Theorem via the proof that Euclid gave in the *Elements*. We will then use this to establish the Similar Triangles Theorem. Once this fact is established, the sine and cosine functions can be defined as in Section 2E.




**(a)** Use the cosine function to prove that in the diagram above we have $a^2+b^2 = (c_1+c_2)^2$.

**(b)** The proof in part (a) is considerably simpler than Euclid's proof of the Pythagorean Theorem. Can we save effort by using this proof instead?

**3.3**. Let $\mathcal{A}$ be an axiom system and let $\mathcal{A}'$ be obtained by removing one or more of the axioms from $\mathcal{A}$. Prove or disprove each of the following statements:

**(a)** If $\mathcal{A}$ is consistent then so is $\mathcal{A}'$.

**(b)** If $\mathcal{A}$ is independent then so is $\mathcal{A}'$.

**3.4**. You may have seen (either in calculus or linear algebra) the formula

$$\cos(\theta) = \frac{\mathbf{u} \cdot \mathbf{v}}{||\mathbf{u}||\,||\mathbf{v}||}$$

(where $\mathbf{u}$ and $\mathbf{v}$ are vectors, $\mathbf{u} \cdot \mathbf{v}$ is their dot product, $||\mathbf{u}||$ and $||\mathbf{v}||$ are their lengths, and $\theta$ is the angle between them). Use this formula to derive a formula for the measure of angle $\angle BAC$ in the Cartesian plane model if $A = (a_1, a_2)$, $B = (b_1, b_2)$, and $C = (c_1, c_2)$.

**3.5**. Take $x^2 + y^2 + (z - \frac{1}{2})^2 = \frac{1}{4}$ as the sphere in the punctured sphere model (with "north pole" $N = (0, 0, 1)$ and define the function $f$ as outlined on p.97. We measure the distance between two points $(x_1, y_1, z_1)$ and $(x_2, y_2, z_2)$ on the punctured sphere by defining their distance to be the usual distance (in the $xy$-plane) between the points $f(x_1, y_1, z_1)$ and $f(x_2, y_2, z_2)$. Show that this distance is

$$\frac{\sqrt{[x_1(1-z_2) - x_2(1-z_1)]^2 + [y_1(1-z_2) - y_2(1-z_1)]^2}}{(1-z_1)(1-z_2)}$$

What happens as one of the points approaches $N$?

**3.6**. Let $P$ be a point in the model for spherical geometry and let $d > 0$ be given. Depending on the value of $d$, what might the "circle" with center $P$ and radius $d$ look like in this model?

**3.7**. Use the model for spherical geometry to show that one cannot always construct equilateral triangles in spherical geometry. That is, given points $A$ and $B$, there may not be a point $C$ so that $|AB| = |BC| = |CA|$. Where does the construction of an equilateral triangle in Euclidean geometry (Example 1.5) fail in spherical geometry?

**3.8**. Is the statement "*If a transversal of lines $\Lambda_1$ and $\Lambda_2$ by a third line has a pair of congruent alternate interior angles, then $\Lambda_1 \parallel \Lambda_2$*" true in spherical geometry? (Note: can you see why it isn't enough to just say that spherical geometry has no parallel lines?)

# B. Finite Geometries

If the concepts of axiom systems, geometries, and models seemed difficult and abstract, this section will present an opportunity to make them more familiar. We'll get some experience applying everything from the last section to some simple examples called *finite geometries*. We'll be able to experiment with axiom systems for these geometries and use models to show the independence and consistency of those systems. And, of course, we'll prove some theorems in these geometries!

### A first example

Consider first the following set of four axioms:

**Axiom $\mathbf{X}_1$.** There exists at least one point.
**Axiom $\mathbf{X}_2$.** Every pair of distinct points is a line.
**Axiom $\mathbf{X}_3$.** Every line contains exactly two points.
**Axiom $\mathbf{X}_4$.** Every point belongs to exactly four lines.

This axiom system (the above axioms together with the usual terms of *point*, *line*, and elementary set membership notions) describes a geometry which we can call *five point geometry* in light of the following theorem:

THEOREM **3.3.** *In five point geometry, there are exactly five points.*

*Proof:* We first prove that there are at least five points.

- First, by Axiom $X_1$ there exists at least one point $P$.
- By Axiom $X_4$ $P$ belongs to four lines.
- By Axiom $X_3$ each of these lines contains exactly two points, say $\Lambda_i = \{P, Q_i\}$ for $i = 1, 2, 3, 4$.
- Since these lines are distinct, the points $Q_1$, $Q_2$, $Q_3$, and $Q_4$ are distinct, so there are at least five points.

Now we complete the proof by showing that there cannot be more than five points. Assume to reach a contradiction that there are six points $\{R_1, R_2, R_3, R_4, R_5, R_6\}$.

- By Axiom $X_2$ each of the sets $\{R_1, R_2\}$, $\{R_1, R_3\}$, $\{R_1, R_4\}$, $\{R_1, R_5\}$, and $\{R_1, R_6\}$ is a line.
- This contradicts Axiom $X_4$ since $R_1$ belongs to five distinct lines.

The contradiction must have come from the assumption that there are six distinct points. Thus, we have proved that there are exactly five points. $\square$

We can give one model for this geometry by simply forming sets that satisfy the axioms. To this end, let $\Pi = \{A, B, C, D, E\}$, and let $\mathcal{L} = \{\{A, B\}, \{A, C\}, \{A, D\}, \{A, E\}, \{B, C\}, \{B, D\}, \{B, E\}, \{C, D\}, \{C, E\}, \{D, E\}\}$ be our collection of designated lines. You can check by hand that these sets satisfy all four axioms in $\mathcal{A}$.

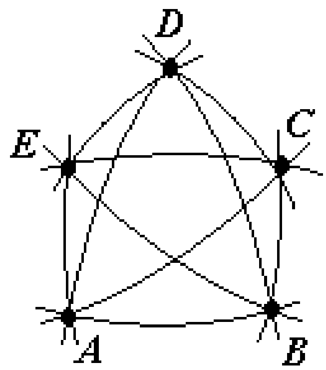



Figure 3.5:

If you're more comfortable with visual models, consider Figure 3.5. This is a graphical representation of the sets listed above. The lines in the geometry are associated with the arcs in the figure, with each arc passing through the two points that make up the line. (Don't be tempted to think of these arcs as containing intermediate points between the pairs. The model has only the five points $A$, $B$, $C$, $D$, and $E$. The arcs are simply visual cues that suggest pairs of points that constitute lines.)

As Theorem 3.1 indicates, the fact that there is a model for this geometry proves that its axiom system is consistent. Again, the rationale for this conclusion is that if the axioms were self-contradictory we would not be able to find any representation that satisfied them all simultaneously.

Are the axioms independent? They are! It takes four models to show this, however, using Theorem 3.2.

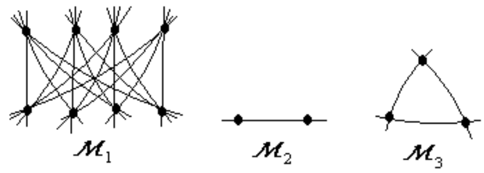


Figure 3.6:

- First, to show that Axiom $X_1$ is independent of the other three axioms consider the empty set as a model. This may seem strange, but each of Axioms $X_2$, $X_3$, and $X_4$ are true statements about a geometry that contains *no points* and *no lines*. (In the same way, the statement "*All living unicorns have green eyes*" is true. Any statement about the elements of the empty set is said to be *vacuously true*.) This shows that the other axioms do not imply Axiom $X_1$.

- The model $\mathcal{M}_1$ at the left of Figure 3.6 satisfies Axioms $X_1$, $X_3$, and $X_4$ but not Axiom $X_2$. So, $X_2$ is independent of the other axioms.

- To show that Axiom $X_3$ is independent of the other axioms, we use a model that is difficult to draw in a diagram. Let the underlying set be just four points $\{A, B, C, D\}$ and let there be seven lines: $\{\{A, B\}, \{A, C\}, \{A, D\}, \{B, C\}, \{B, D\}, \{C, D\}, \{A, B, C, D\}\}$. Note that one of the lines contains all four points, so Axiom $X_3$ is certainly not true in this model. However, it's easy to check that all three other axioms are satisfied.

- Both of the models $\mathcal{M}_2$ and $\mathcal{M}_3$ in Figure 3.6 satisfy Axioms $X_1$, $X_2$, and $X_3$, but not Axiom $X_4$. Thus, Axiom $X_4$ is independent of the other axioms.

Taken together, the above models show that the axiom system $\mathcal{A}$ is independent.

The models $\mathcal{M}_2$ and $\mathcal{M}_3$ mentioned above show one other interesting fact: the axiom system containing only Axioms $X_1$, $X_2$, and $X_3$ is not categorical. We can see this because we have two models[3] satisfying these axioms, and the two

[3] Actually, we have *three* distinct non-isomorphic models. The model in Figure 3.5 also satisfies Axioms $X_1$, $X_2$, and $X_3$.

models are clearly not isomorphic. (No one-to-one correspondence could exist between these models because they include different numbers of points.)

Now suppose we were to add to the four axioms in $\mathcal{A}$ the new axiom

**Axiom $\mathbf{X}_5$.** There exist three lines no two of which intersect.

In this new axiom system of five axioms, Axioms $X_3$ and $X_5$ easily prove that there are at least six points. But this shows that the new larger axiom system is neither independent nor consistent. Independence fails because Axiom $X_1$ is now proved by using $X_3$ and $X_5$. (The fact that there are at least six points certainly proves that there exists at least *one* point!) Consistency fails because Theorem 3.3 is still a valid statement (since Axioms $X_1$ through $X_4$ are still in force) and is in direct contradiction to the conclusion that six points exist.

In summary:

- $\{X_1, X_2, X_3\}$ is not categorical.
- $\{X_1, X_2, X_3, X_4\}$ is independent and consistent (and is in fact also categorical).
- $\{X_1, X_2, X_3, X_4, X_5\}$ is neither independent nor consistent.

Exercise 3.9 at the end of this section will give you the opportunity to carry out similar reasoning to the above on a similarly simple finite geometry.

**Fano's Geometry**

We'll now meet a more complex instance of a finite geometry. *Fano's geometry* has five axioms, as follows:

**Axiom $\mathbf{F}_1$.** There exists at least one line.
**Axiom $\mathbf{F}_2$.** Every line contains exactly three points.
**Axiom $\mathbf{F}_3$.** There is no line containing all of the points.
**Axiom $\mathbf{F}_4$.** Given any two distinct points there is exactly one line containing both of them.
**Axiom $\mathbf{F}_5$.** Every two lines contain at least one point in common.

The first theorem we will prove with these axioms is a simple strengthening of Axiom $F_5$.

THEOREM **3.4.** *In Fano's geometry, every two distinct lines contain exactly one point in common.*

*Proof.* Let $\Lambda_1$ and $\Lambda_2$ be two distinct lines. We will prove that there is exactly one point in their intersection.

- Axiom $F_5$ states that there is at least one point in this intersection, so we need only show that there *cannot* be *more than one* point.
- So, assume to reach a contradiction that $A$ and $B$ are distinct points common to both $\Lambda_1$ and $\Lambda_2$.
- The contradiction is immediate: it is a violation of Axiom $F_4$ for two distinct lines to contain both $A$ and $B$. □

Now you might be wondering why, given the result of this theorem, did we only state in Axiom $F_5$ that two lines contain *at least one* point in common. Would it not have saved us effort to have made the statement of Theorem 3.4 our Axiom $F_5$ instead? The answer is that, yes, it would have saved work, because then the above proof would not be needed. However, it would have made for a poorer axiom system according to the criteria discussed in the previous section. Remember that we generally want our axioms to assume as little as possible – if we can get by with assuming "at least one point" instead of "exactly one point", so much the better!

Our next theorem shows that Fano's geometry is, in fact, finite. Notice that our theorems are in this order for a reason – we need Theorem 3.4 to justify a step in the following proof.

**THEOREM 3.5.** *In Fano's geometry there are exactly seven points.*

*Proof.* We first show that there are at least seven points.

- By Axiom $F_1$ there is at least one line $\Lambda$.
- By Axiom $F_2$ $\Lambda$ contains three points $A$, $B$, and $C$.
- By Axiom $F_3$ there is a point $D$ not on $\Lambda$.
- By Axiom $F_4$ there is a line $\Lambda_{AD}$ containing both $A$ and $D$.
- By Axiom $F_2$ $\Lambda_{AD}$ must contain a third point, and by Theorem 3.4 this point cannot be either $B$ or $C$ (since $\Lambda$ and $\Lambda_{AD}$ already share point $A$ in common). Call this third point $E$.
- Again by Axiom $F_4$ there is a line $\Lambda_{BD}$ containing both $B$ and $D$.

- By Axiom $F_2$ $\Lambda_{BD}$ must contain a third point, and again by Axiom $F_5$ this point cannot be $A$, $C$, or $E$. Call this point $F$.
- Finally, by Axiom $F_4$ there is line $\Lambda_{CD}$ containing both $C$ and $D$.
- As before, $\Lambda_{CD}$ must contain a third point $G$ distinct from any of the points named so far.
- We have now accounted for at least seven points.

Now it remains to show that there cannot be any more than seven points in this geometry. We do this by contradiction – assume to reach a contradiction that there is a point $P$ distinct from all seven points described so far.

- By Axiom $F_4$ there is a line $\Lambda_{PD}$ containing both $P$ and $D$.
- By Axiom $F_5$ $\Lambda_{PD}$ must intersect $\Lambda$ in at least one point.
- But if $\Lambda_{PD}$ contains either $A$, $B$, or $C$, then it shares two points in common with either $\Lambda_{AD}$, $\Lambda_{BD}$, or $\Lambda_{CD}$, contrary to Theorem 3.4.
- So, $\Lambda_{PD}$ intersects $\Lambda$ at some point $Q$ distinct from $A$, $B$, and $C$.
- This means $\Lambda$ contains at least four points, contradicting Axiom $F_2$.

This contradiction shows that no eighth point can exist. □

You most likely made a sketch while reading the above proof similar to the solid lines in Figure 3.7. If we add the dashed lines, this figure becomes a model for Fano's geometry. You can check directly that all five axioms are satisfied by this arrangement.

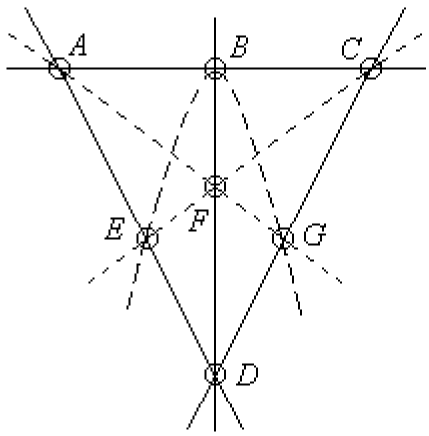


Figure 3.7: A model for Fano's geometry

From the model it appears that every point in Fano's geometry lies on exactly three lines. If we are to make this a theorem, however, we will need to prove it from the axioms, not from a reference to any picture. (Again, because we can't say that just because *this* particular model for Fano's geometry has a given property that *every* model for Fano's geometry would have that property.) In this case, it isn't difficult to supply the formal proof from the axioms. Again notice that we are free to make use of the theorems we have already established.

**THEOREM 3.6.** *In Fano's geometry, every point lies on exactly three lines.*

*Proof:* Let $P$ be a point in Fano's geometry. We will prove directly that $P$ is on exactly three lines.

- By Theorem 3.5 there are exactly six other points.
- If $A$ is one of these other points, there is a line $\Lambda_{PA}$ containing both $P$ and $A$ by Axiom $F_4$.
- By Axiom $F_2$ $\Lambda_{PA}$ must contain a third point $A'$.
- We have so far accounted for only three points, so let $B$ be another point in the geometry.
- By Axiom $F_4$ there is a line $\Lambda_{PB}$ containing both $P$ and $B$, and by Axiom $F_2$ it must contain a third point $B'$.
- Now we have accounted for at most five points, so let $C$ be a point not yet used.
- By Axiom $F_4$ there is a line $\Lambda_{PC}$ containing both $P$ and $C$, and by Axiom $F_2$ it must contain a third point $C'$.
- We now have at least three lines ($\Lambda_{PA}$, $\Lambda_{PB}$, and $\Lambda_{PC}$) containing $P$.
- But all of the points $A$, $A'$, $B$, $B'$, $C$, and $C'$ are distinct, for if any two coincide then two of the three lines mentioned would contain two points in common, contrary to Theorem 3.4. So, the three lines mentioned so far together contain all seven points of the geometry.
- But this means there can be no fourth line containing $P$, else this new line would share two points with one of $\Lambda_{PA}$, $\Lambda_{PB}$, $\Lambda_{PC}$, contrary to Theorem 3.4. □

There are other facts suggested by the model in Figure 3.7, but we will leave them for exercises 3.10 to 3.12.

The fact that we were able to draw a model for Fano's geometry means that the axiom system we gave is consistent. Is it also independent? The "empty model" shows that Axiom $F_1$ is independent of the other axioms. Axiom $F_2$ is more difficult. Consider the model in which the "points" are pairs of antipodal points on a sphere (that is, each "point" in our model is the intersection of the sphere with some line through its center). Let the "lines" of the model be the great circles of the sphere. A little thought should convince you that this model satisfies all of Fano's axioms except Axiom $F_2$. (It satisfies Axiom $F_4$ because any two pairs of antipodal points on the sphere together determine a unique plane through the sphere's center and thus a unique great circle on the sphere. It satisfies Axiom $F_5$ because every pair of great circles share a pair of antipodal points and thus each pair of "lines" in the model meet in a "point".) Accordingly, Axiom $F_2$ cannot possibly be a consequence of the other four axioms, or this model (an extremely important example called the *projective plane*) would not exist.

We leave to you the task of finding models to show the independence of the remaining axioms in Fano's geometry.

## Exercises

**3.9**. Consider the axiom system $\mathcal{A}'$ containing the following four axioms:

> **Axiom $\mathbf{Y_1}$.** There exists at least one point.
> **Axiom $\mathbf{Y_2}$.** Every line contains exactly two points.
> **Axiom $\mathbf{Y_3}$.** If $P$ is a point and $\Lambda$ is a line not containing $P$ then there is exactly one point on $\Lambda$ that is not on a common line with $P$.
> **Axiom $\mathbf{Y_4}$.** Every point is on exactly three lines.

Let $\mathcal{G}'$ be the geometry generated by this axiom system.

**(a)** Prove the following theorem: *In the geometry $\mathcal{G}'$, there exist exactly six points.*

**(b)** Prove that $\mathcal{A}'$ is consistent by exhibiting a model.

**(c)** Show that $\mathcal{A}'$ is independent.

**(d)** Show that if Axiom $Y_4$ is omitted then the axiom system that remains is not categorical.

**3.10**. Prove that in Fano's geometry there are exactly seven lines.

**3.11**. Prove that in Fano's geometry if $A$ and $B$ are distinct points then there are exactly two lines containing neither of them.

**3.12**. Prove that in Fano's geometry that if $A$, $B$, and $C$ are three points not all on one line then there is exactly one line containing none of them.

**3.13**. Give models to complete the verification that the axioms of Fano's geometry are independent.

Exercises 3.14 through 3.19 deal with the *finite geometry of Pappus*, which has the following six axioms.

> **Axiom $\mathbf{P}_1$.** There exists at least one line.
> **Axiom $\mathbf{P}_2$.** No two lines intersect in more than one point.
> **Axiom $\mathbf{P}_3$.** Every line contains exactly three points.
> **Axiom $\mathbf{P}_4$.** No line contains all points.
> **Axiom $\mathbf{P}_5$.** If $\Lambda$ is a line and $P$ is a point not on $\Lambda$ then there is exactly one line containing $P$ which does not intersect $\Lambda$.
> **Axiom $\mathbf{P}_6$.** Given a point $P$ not on a line $\Lambda$, there is exactly one point $A$ on $\Lambda$ such that no line contains both $P$ and $A$.

**3.14**. Prove in Pappus' geometry that given any point $P$ there is a line not containing $P$. (Your proof should begin "*Let $P$ be a point. We will prove that there is a line not containing $P$.*" Note that you cannot use Axiom $\mathbf{P}_4$ directly.)

**3.15**. Prove in Pappus' geometry that every point belongs to exactly three lines. (Use the theorem in Exercise 3.14.)

**3.16**. Prove in Pappus' geometry that there are exactly nine points.

**3.17**. Prove in Pappus' geometry that there are exactly nine lines.

**3.18**. Prove in Pappus' geometry that if $\Lambda$ is a line then there are exactly two lines not intersecting $\Lambda$.

**3.19**. The geometry of Pappus is named for the 4th century mathematician Pappus of Alexandria who proved the following theorem in Euclidean geometry: *Let $A$, $B$, and $C$ be distinct points on one line and let $A'$, $B'$, and $C'$ be distinct points on a different line. Let $P$ be the intersection of $\overleftrightarrow{AB'}$ and $\overleftrightarrow{A'B}$, $Q$ the intersection of $\overleftrightarrow{BC'}$ and $\overleftrightarrow{B'C}$, and $R$ be the intersection of $\overleftrightarrow{AC'}$ and $\overleftrightarrow{A'C}$. Then $P$, $Q$, and $R$ are collinear.* Find a model for the geometry of Pappus by examining the statement of this theorem.

# C. Axiom Systems for Euclidean Geometry

This section tells, at least in outline, a long and complex story. It is the story of the development of geometry from Euclid to the present as seen through the lens of the axiomatic method. We will spend the next three chapters working out the mathematical details of what we sketch here, so reading this section carefully will give you both a preview and a perspective of what is to come.

### Euclid's Axiom System

It would be difficult to overstate the importance of Euclid's *Elements*. It is not just a great work of mathematics – it is a classic of western civilization. It is no exaggeration to say that for two millennia, the greatest minds in the western world read and studied Euclid's text. So universal is its admiration and so dominating its influence on the subject that one sometimes sees theorems cited by their location in the *Elements* without explanation: saying that the 180° Sum Theorem is I.32 (Theorem 32 in Book I) need not be followed by the explanation "in the *Elements*" any more than we would need to explain that Deuteronomy 6:7 is "in the Bible". Euclid's organization of material is masterful and is the greatest reason for the work's fame. Its 465 theorems cover more than just geometry. In fact, while Books I, III, IV, and VI cover plane geometry and Books XI, XII, and XIII cover solid geometry, the remaining six books are devoted to other topics. An outline of its contents is as follows.

- Book I gives an efficient and elegant tour of the geometry of simple plane figures, concluding with the Pythagorean Theorem (I.47) and its converse (I.48).
- Book II covers what we might call "*geometric algebra*" – relationships we would express through algebraic notation were handled geometrically by the early Greeks.
- Book III covers the geometry of circles.
- Book IV covers regular polygons.
- Book V presents Eudoxus' theory of proportions (the mathematics that finally resolved the crisis of incommensurability).
- Book VI applies the ideas of Book V to develop the geometry of similar figures.

- Books VII through IX depart from geometry altogether and treat the subject of number theory. Euclid's algorithm for computing the greatest common divisor of two integers (VII.1 and VII.2) and his proof that there exist infinitely many primes (IX.20) are as famous as any of the *Elements*' geometry results!

- Book X concerns square roots and incommensurable magnitudes. The techniques here have been made obsolete by algebra, but were much acclaimed in Euclid's time.

- Books XI, XII, and XIII then give a development of solid (three-dimensional) geometry.

We'll now examine the *Elements* as *an axiom system.* But first, a warning: though we will point out a few "flaws" in Euclid's work, they are flaws only in our modern viewpoint. The *Elements* stood as *the* example of mathematical rigor and exactness for many centuries. That we have refined the idea of rigor even further (to the point that some features in the *Elements* now appear inadequate) should not detract from our admiration of what Euclid accomplished.

Euclid never specified a set of undefined terms. This is, in fact, one of the weaknesses of the *Elements* when seen from a modern perspective. The work opens with a list of definitions that are for the most part not really proper definitions at all, but rather *descriptions* meant to shape the reader's mental image of what each term designates. Thus, he attempts to define a *point* as "that which has no part", and an *angle* is the inclination between two crossing lines. *Point*, of course, would be better left undefined, and this definition of *angle* is flawed in that "inclination" is certainly no better understood than the word "angle" itself.

With the terminology set forth, Euclid proceeds to state his axioms. A note on word usage is in order here. We have been employing the term *axiom* to designate a basic assumption that is taken as true without proof. Euclid has two distinct terms for such assumptions, calling some *axioms* and some *postulates.* While it is not entirely clear how the distinction between axiom and postulate was understood in Euclid's day, it appears that *axiom* (or *common notion*, as it is sometimes rendered) is reserved for assumptions common to all mathematical inquiries. Thus, Euclid states as an axiom that "*things equal to the same thing are equal also to each other*" – what we would call the *transitive law* for equality. For assumptions that deal with geometric relationships or objects he uses the term *postulate.* We will respect Euclid's choice of terminology and refer to the fundamental geometric assumptions in his axiom system as *postulates.* The (slightly modernized) statements of Euclid's five postulates are:

**Postulate 1.** It is possible to draw a straight line from any point to any other point.

**Postulate 2.** It is possible to produce a line segment of any length within a straight line.

**Postulate 3.** It is possible to draw a circle with center at any point and passing through any other point.

**Postulate 4.** All right angles are congruent.

**Postulate 5. (Euclid's "parallel postulate")** If a line $\Lambda_3$ transverses two lines $\Lambda_1$ and $\Lambda_2$, and if the interior angles on one side of $\Lambda_3$ formed by these intersections (the angles marked $\alpha$ and $\beta$ in Figure 3.8) sum to less than 180°, then $\Lambda_1$ and $\Lambda_2$ intersect on that side of $L$.

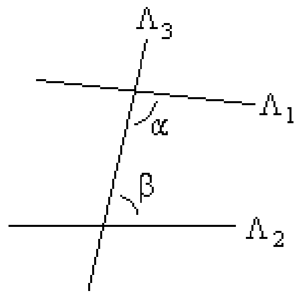


Figure 3.8: A picture for Euclid's parallel postulate

You should note from the statements of these postulates that Euclid's approach to geometry was greatly influenced by straightedge and compass constructions. Indeed, the first three of his postulates merely clarify what can be done by using these tools. Postulate 4, meanwhile, establishes the right angle (one of the terms defined in his prelude list of definitions) as the standard unit in measuring angles.

Postulate 5 is very much different, and we will have a good deal to say about it throughout the next several chapters of this text. For now, note that a consequence of this assumption is that $\Lambda_1$ and $\Lambda_2$ in Figure 3.8 cannot be parallel unless any transversal of these lines has congruent corresponding angles. It is thus really an assumption about the nature of parallel lines, and is for that reason called "*Euclid's parallel postulate*".

The bulk of the *Elements* consists of proofs of theorems built from these axioms. For instance, the first theorem asserts that it is possible to construct on any segment $AB$ an equilateral triangle $\triangle ABC$. Its proof is exactly the construction given in Example 1.5, with each step of the construction justified by one of the five axioms. But there is (by modern standards) a problem with the argument. It relies on finding the intersection of two circles – but how do we know that the two circles will in fact intersect? If you answered that it is "clear from the picture" that they intersect, you have made the error of relying on the *model* rather than on the *axioms*. Euclid's postulates are not strong enough to tell us anything about how circles behave relative to other objects, yet he clearly intended an intuitive notion of what a circle is like. (If you think this is a trivial point, consider that this particular step may *not* work in spherical geometry – see Exercise 3.7!) In other words, Euclid uses the model for his axiom system in a way that is inappropriate by modern standards. Since the five postulates of Euclid are not sufficient to justify this step, we would need to augment the axiom system, perhaps by adding a new postulate something like this:

> *If circle* $\Gamma_1$ *contains a point inside circle* $\Gamma_2$ *and if* $\Gamma_2$ *contains a point inside* $\Gamma_1$*, then the two circles* $\Gamma_1$ *and* $\Gamma_2$ *intersect in exactly two points.*

(Of course, this could not be adopted without first defining what is meant by the *inside* of a circle!)

This is no small matter, for remember that the goal of the axiomatic approach is to derive as much information as possible from as few assumptions as possible. In building axioms for plane geometry we try to discover the assumptions which lie buried at the roots of all geometry theorems. We ruin our chances at this endeavor if we make any use of assumptions not strictly implied by our axioms.

So already in the proof of the first theorem of the *Elements* we see the difficulties of the axiomatic approach. The axioms we adopt should in theory allow us to prove geometric facts without recourse to a model, yet so much of our geometric understanding is tied up in reasoning with diagrams that it is truly difficult to account for *all* of the needed assumptions to accomplish this.

Another example of this arises in establishing the SAS congruence criterion for triangles. Euclid here employs a proof based on a "superposition" argument: if two triangles have two pairs of congruent sides with the angles between these sides congruent as well, then by "moving" one triangle to overlay the other, we see that they must be congruent (see Figure 3.9). While this is certainly in line with our intuitive notion of "congruence" from the usual model, it is not a technique allowable by Euclid's axioms. Indeed, outside of the usual model, the

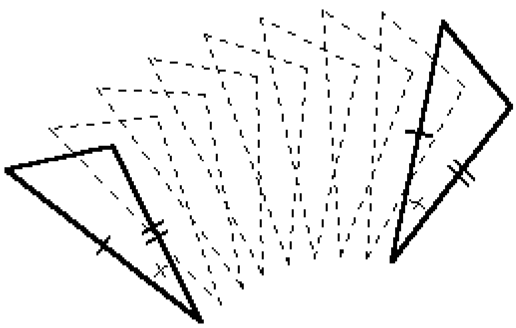


Figure 3.9:

notion of "moving" a triangle is meaningless. Euclid himself seemed reluctant to use the technique of superposition, but found no better way of justifying the SAS theorem.[4]

Note that both of these flaws we have pointed out amount to an improper use of the mental image or model – assuming properties evident from the picture but yet not supported by the axioms themselves. Again, we should not fault Euclid, for this is a "flaw" only through the lens of our modern understanding of the axiomatic method. As that modern understanding developed, mathematicians were able to adjust Euclid's axiom system accordingly. There *are* indeed modern axiom systems that properly establish Euclidean geometry. But before we give the story of their development, we have a different tale to tell.

**The parallel postulate controversy**

Look again at Euclid's five geometric postulates and you will no-doubt agree that the parallel postulate is of a different character than the others. Almost from the time of Euclid this one assumption was the subject of criticism and scrutiny. Many objected that the truth of Postulate 5 is not sufficiently self-evident for it to stand as a basic assumption. There were two suggested cures for this, springing from different views on whether or not the postulate was even needed.

**(1)** If we don't believe that Postulate 5 is independent of the other postulates, then we might hope to do away with it altogether by supplying a proof for the parallel postulate based only on the other four postulates. This would remove it from the list of assumptions in the axiom system and change its status to that of a *theorem*. (Advocates of this solution noted that Postulate 5 *sounds* more like a theorem than an axiom, taking as it does the form "*If* X, *then* Y".)

**(2)** If we agree with Euclid that some sort of assumption about properties of parallel lines is necessary in our axiom system, we might still hope to replace Euclid's awkwardly worded postulate with one that is more simply worded and believable.

Consider first the more ambitious option **(1)**. "Proving the parallel postulate" was tried many times over many centuries, and the independence of the parallel postulate remained an unsettled question for over 2000 years. The usual technique was to attempt a proof of the parallel postulate by contradiction. One

[4]Modern axiom systems for geometry, including the one we will construct in Chapters 4 and 5, do not try to prove the SAS criterion, but rather accept it as one of the axioms.

would assume (in hopes of finding a contradiction) that the statement of Postulate 5 is false but that the other Postulates are valid. Some of the would-be postulate provers even convinced themselves that they had found a contradiction, but none of their proofs held up under careful scrutiny. It took centuries for the realization to set in (fostered by a sufficiently mature view of the axiomatic method) that what they were doing was developing a new geometry – one in which Euclid's fifth postulate is replaced by its negation. Thus, if $\mathrm{E}_1$ through $\mathrm{E}_5$ represent Euclid's five postulates, then the new geometry was built on the axioms $\{\mathrm{E}_1, \mathrm{E}_2, \mathrm{E}_3, \mathrm{E}_4, {\sim}\mathrm{E}_5\}$. The attempted proofs were leading not to contradictions, but rather to theorems in this new "non-Euclidean" geometry. Once a model for the new geometry (now called *hyperbolic geometry*) was found in the second half of the 19th century, the principle of Theorem 3.2 established that the parallel postulate was indeed independent. (We'll give a more detailed account of this discovery in Chapter 6.)

Actually, our description of hyperbolic geometry here is quite a bit oversimplified. For, as we have already noted, Euclid's postulates are really quite inadequate for the job of describing any meaningful geometry. The gaps between what Euclid stated in his postulates and what he clearly assumed to be true of his points and lines had to be filled in with additional axioms, and there were several such attempted repairs. It is to these improved axiom systems that we must turn for a real definition of hyperbolic geometry. As it turns out, all sets of axioms for Euclid's geometry share one feature with Euclid's set of postulates – they consist of a set of axioms accomplishing what Euclid had in mind for $\{\mathrm{E}_1, \mathrm{E}_2, \mathrm{E}_3, \mathrm{E}_4\}$ together with some axiom on the properties of parallel lines. The former group of axioms are called *neutral axioms*. The situation can be illustrated in shorthand as

**Euclid's axiom system:** $\{\mathrm{E}_1, \mathrm{E}_2, \mathrm{E}_3, \mathrm{E}_4\} \cup \{\mathrm{E}_5\}$

**Modern axiom system for Euclidean geometry:** $\{\text{neutral axioms}\} \cup \{\mathrm{P}\}$

where P is some axiom on parallel lines. Hyperbolic geometry is then given by

**Axiom system for hyperbolic geometry:** $\{\text{neutral axioms}\} \cup \{{\sim}\mathrm{P}\}$

In the centuries leading up to the discovery of hyperbolic geometry, mathematicians trying to remove the parallel postulate from Euclid's set of assumptions attempted to remove it also from the proofs of Euclidean geometry. If the neutral axioms were to stand alone, after all, all proofs in Euclidean geometry would have to be constructed from just them. And many of Euclid's theorems *can* be proved using only neutral axioms.[5] The collection of such theorems constitute what is called *neutral geometry* – the geometry arising from the axiom system of only the neutral axioms.

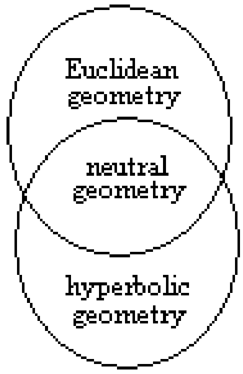



Figure 3.10:

Neutral geometry is worthy of study (we will devote Chapter 4 to its development before introducing a parallel postulate for use in Chapter 5) if for no other reason than this: the theorems of neutral geometry are valid in both Euclidean and hyperbolic geometries. After all, a proof of a neutral geometry theorem can rely on only the neutral axioms, and those axioms are valid in both Euclidean and hyperbolic geometries. The relationship is illustrated as a Venn diagram in Figure 3.10: the axioms (and thus also the theorems) of Euclidean geometry and hyperbolic geometry overlap in neutral geometry. Though the overlap is significant (it contains, for example, all of the congruence criteria for triangles, the Isosceles Triangle Theorem, the Vertical Angles Theorem, and half of the Alternate Interior Angles Theorem), there are theorems in Euclidean geometry that cannot be proved without the parallel postulate and thus lie in the top part of this Venn diagram. These are the theorems we will prove in Chapter 5. Likewise, there are theorems (we'll meet several in Chapter 6) strictly in the lowest portion of this Venn diagram – theorems provable in hyperbolic geometry that are certainly *not* true in Euclidean geometry.

Now recall that we began our discussion of the parallel postulate controversy (see p.115) by mentioning two possible remedies for the unsatisfactory nature of Euclid's parallel postulate. We now know that remedy **(1)** is not possible, because the parallel postulate is independent of the neutral axioms. What about remedy **(2)**? Can we at least replace the parallel postulate with a simpler, more believable axiom? The answer is yes. But before we get too far, let's consider what it means to "replace" one axiom by another.

**Definition.** Let $\mathcal{N}$ denote a set of axioms for neutral geometry. We say that two statements $P_1$ and $P_2$ are *equivalent relative to neutral geometry*

[5]Euclid himself seems to have been sensitive to the special character of the parallel postulate and to have delayed its use as long as possible in the *Elements*. It is first used in the proof of I.29, which states that *alternate interior angles in a transversal of parallel lines are congruent.* We will see in Chapters 4 and 5 that Euclid was remarkably correct: though its converse can be proved using only the neutral axioms, this half of the Alternate Interior Angles Theorem cannot be proved without reference to an assumption like the parallel postulate.

if $P_1$ can be proved as a theorem using axioms $\mathcal{N} \cup \{P_2\}$ and $P_2$ can be proved as a theorem using axioms $\mathcal{N} \cup \{P_1\}$.

So, if a statement P is equivalent relative to neutral geometry to Euclid's Postulate 5, then $\mathcal{N} \cup \{P\}$ would serve as an axiom set for Euclidean geometry. It makes sense, then, to consider P as a valid "parallel postulate".

Through the history of the parallel postulate controversy, there were many statements found to have this property. For example,

- *There exists a triangle with angle measures summing to* 180°.
- *There exists a rectangle.*
- *If* $\Lambda$ *is a line and* $P$ *is a point not on* $\Lambda$ *then there is exactly one line through* $P$ *that is parallel to* $\Lambda$.

are all equivalent relative to neutral geometry to Euclid's Postulate 5. So, any of these *could* serve as the "parallel postulate" whose addition to the neutral axioms gives us all of Euclidean geometry.

One could argue that the first of these is not "obvious enough" to be an axiom. Is it really clear that the mental image we have of the Euclidean plane *necessitates* angle sums of 180° for triangles? The second would be a reasonable choice, for it is believable and elementary and easy to state. It has the disadvantage, however, that it is not at all clear that this is an assumption about the behavior of parallel lines. History has instead opted for the third statement as the best alternative to Euclid's Postulate 5. This statement, known as *Playfair's Postulate*,[6] has its drawbacks for certain. It, like Euclid's parallel postulate, has the form of an "*If* X *then* Y" statement – more the form for a theorem than an axiom. Still, it is far easier to comprehend than is Euclid's, and captures well the fact that what we are doing is specifying the behavior of parallels. In Euclidean geometry, there is *one and only one* line through $P$ parallel to $\Lambda$. In hyperbolic geometry, there is *more than one* such parallel. (And in spherical geometry, there are *no* such parallels!) For an outline of the proof that Euclid's Postulate 5 and Playfair's Postulate are equivalent relative to neutral geometry, see Exercise 3.22.

**New axiom systems for Euclidean geometry**

The discovery of hyperbolic geometry emphasized the critical role played by axioms, and thus brought a renewed interest in the axiomatic approach. Mathematicians of that era set to the task of once and for all establishing Euclidean

[6]This axiom appeared in John Playfair's 1795 treatment of geometry, though it was used as early as the 5th century AD by Proclus (of whom we will say more in Chapter 6).

geometry on a solid axiomatic footing. In 1882 Moritz Pasch (1862-1943) produced such a foundation for Euclidean geometry by patching the flaws in the axioms of the *Elements*. Giuseppe Peano (1858-1932) carried the effort further by setting the geometric axioms purely in formal logic. His 1889 work took the thesis that the best way to ensure against making unaccounted assumptions in reasoning from a model (as Euclid had done) is to strip the subject matter of any reference to pictures. One is not likely to let slip into the argument some assumption based on a mental image or diagram if the statements with which one works are as austere as "*through each two x's there is exactly one y*".

But the most successful of these new axiomatic treatments for geometry came in 1899 with the publication of *Grundlagen der Geometrie* (Foundations of Geometry) by David Hilbert (1862-1943 AD). Hilbert gave a development based on 14 axioms for plane Euclidean geometry (21 for plane and solid geometry) that, like Peano's, is careful to avoid the pitfalls of murky boundaries between proof and picture. Hilbert's often-quoted test is that the proofs should be equally convincing if the words "point", "line", and "plane" are replaced by "table", "knife", and "beermug". Yet Hilbert's axioms preserve much of the spirit of Euclid, and can be understood by anyone. Perhaps most importantly, Hilbert carefully demonstrated the independence, consistency (though see the discussion at the end of this section!), and categorical properties for his axiom system. He did this through application of the principles in Theorems 3.2 and 3.1 by creating many models satisfying various subsets of his axioms.

**Philosophies on the nature of mathematics and axioms**

The late 19th century also saw the first attempts to extend the axiomatic method beyond the realm of geometry. Mathematics since the dawn of history has been divided into two loose conglomerations – geometry and arithmetic. And while geometry has been the subject of axiomatic study since early Greece, the world of numbers had not been so treated up to that time. But in the 1880's and 1890's several attempts were made to establish axiom systems for the real numbers and the arithmetic on them. Richard Dedekind (1831-1916), Gottlob Frege (1849-1925), and Peano (mentioned above) all made noteworthy contributions to this effort. Peano's axioms, in particular, were widely accepted and are still used today as the standard axiom system for arithmetic.

This trend of applying the axiomatic method to all of mathematics culminated with the publication from 1910 to 1913 of the monumental three-volume work *Principia Mathematica* by Bertrand Russell (1872-1970) and Alfred North Whitehead (1861-1947). This was truly the axiomatic approach on a grand scale!

It presented a plan to derive all of mathematics from a few fundamental axioms of logic, and carried out such derivations for enough topics to make the entire goal seem plausible. This idea, that mathematics can be viewed as an extension of logic, became known as the *logicist theory*, and was much debated in the first part of the 20th century.

Hilbert, meanwhile, initiated his own plan for setting mathematics on axiomatic footing. But unlike the Russell/Whitehead model, in Hilbert's vision each subdiscipline within mathematics would have its own set of axioms. The creation and development of new areas within mathematics would then be reduced to the investigation of different sets of axioms and their consequences. To oversimplify a bit, we could state this philosophy (known as the *formalist theory*) thus: a mathematical investigation is a game – we choose the rules (axioms) and then we see where the game leads based on those rules.

**Gödel's incompleteness theorems**

Both the logicist and formalist theories viewed mathematics entirely through the lens of the axiomatic method: every mathematical endeavor was to be accomplished through the manipulation of some axiom system. Both theories assumed that a complete understanding of the subject matter (whether it be the entirety of mathematics in the logicist view or merely one of the many possible subdisciplines resulting from some choice of axioms in the formalist view) was possible through choosing axioms wisely. In other words, they presumed the possibility of finding complete axiom systems for the given areas of study.

Thus, it was a severe blow to both of these theories when in 1931 the young Austrian mathematician Kurt Gödel (1903-1978) proved that the goal of a complete axiomatic description of mathematical facts is (in a very definite sense) unattainable. Gödel proved two remarkable theorems, now referred to as the "*incompleteness theorems*". Their actual statements would involve more detail and formality than we wish to go into here, but the general idea is summarized as follows.

## Gödel's Incompleteness Theorems.

- *No consistent axiom system that includes arithmetic on the natural numbers can be complete.*
- *Specifically, if $\mathcal{A}$ is an axiom system for arithmetic then the statement "$\mathcal{A}$* is consistent*" is either false or undecideable in $\mathcal{A}$.*

Pause for a moment to think about these two assertions. The first claims that we cannot construct any consistent and complete axiom system supporting standard arithmetic. That is, no matter what axioms we suggest for arithmetic, there will (unless our axioms are contradictory) be statements about arithmetic that can be neither proved nor disproved by those axioms. And you probably thought arithmetic was elementary!

The second statement expands on this by specifying one statement that will always be undecideable in a consistent axiom system for arithmetic – namely the consistency of the axioms in that system. According to this second incompleteness theorem, *if* our axioms for arithmetic are in fact consistent, we'll never be able to prove it!

Now we *do* have axiom systems for the real numbers (which, of course, include the natural numbers) and their arithmetic – we have already mentioned Peano's axioms. The second incompleteness theorem tell us we can never know for sure that they are consistent. It is possible – though no one believes it likely – that we might someday discover a contradiction within our axioms for this number system. The situation is this: the only answer we can ever prove to the question "*Is our axiom system consistent?*" is "*No*", and that would be by finding a contradiction in our axioms. The answer might well be "*Yes, they are consistent*", but we'll never know it. We may not have found a contradiction yet, but we can't be sure there isn't one somewhere waiting to be eventually found.

**The consistency of Euclidean geometry**

So how is this relevant to geometry? Recall that there is a model for Euclidean geometry – namely the Cartesian plane model (see p.96). This model obeys all of the axioms of Hilbert's axiom system (and all of the axioms we will introduce over the next two chapters). By Theorem 3.1, if a model exists that satisfies the axioms in an axiom system, then that axiom system is consistent. So, can we conclude that Hilbert's axioms for Euclidean geometry (and the system we will develop in Chapters 4 and 5) is consistent?

No. The problem is that our model is built upon the real numbers. The coordinate plane is certainly only as "real" as the number system from which it is defined. If the real numbers are riddled with contradictions (a possibility we cannot rule out, according to the second incompleteness theorem) then our model doesn't really exist. But, on the other hand, if the real number system (as defined by our axioms for that system) *is* consistent, then we *do* in fact have a model for our geometric axiom system. So, what we can conclude is that *our geometric axiom system is at least as consistent as the real numbers themselves.*

The incompleteness theorems are in some ways comparable to the discovery of incommensurability by early Greet mathematicians. Both revelations shook to the core a prevailing belief of the time: in the case of incommensurability it was the belief in commensurability of all numbers, while in the case of Gödel's theorems it was faith in the ability of the axiomatic approach to encompass all of mathematics. But the discovery of incommensurability was in its time both a setback and a catalyst to progress. It led to the development of a satisfactory theory of ratios, and a more complete understanding of the real numbers. So it will likely be with Gödel's theorems. Time will tell what mathematical achievements might eventually flow from the shake-up they caused.

## Exercises

**3.20**. True or False? (Questions for discussion)

- **(a)** Euclid's axioms are not independent.
- **(b)** The parallel postulate controversy was a debate over the consistency of Euclid's axioms.
- **(c)** Hilbert's axiom system for geometry is independent and categorical.
- **(d)** Every theorem in Euclidean geometry is also true in neutral geometry.
- **(e)** The parallel postulate is independent of the axioms of neutral geometry.
- **(f)** Gödel's incompleteness theorems imply that any axiom system for the real numbers is either inconsistent or incomplete.
- **(g)** Gödel's incompleteness theorems say that we can never prove an axiom system for the real numbers is inconsistent.

**3.21**. Euclid's chose his axioms to allow the construction of a circle only from a given center and point on the circle instead of from a given center and radius (see p.40 and Exercise 1.46). Why do you suppose he did this?

**3.22**. In this problem we prove the equivalence relative to neutral geometry of Euclid's Postulate 5 and Playfair's Postulate. We will need to assume for now that the Vertical Angles Theorem can be proved in neutral geometry and that the 180° Sum Theorem can be proved using neutral geometry together with Playfair's Postulate. (We will accomplish these proofs for ourselves in Chapters 4 and 5.)

**(a)** Using only Euclid's Axiom 5 and the Vertical Angles Theorem, prove that the lines $\Lambda_1$ and $\Lambda_2$ in Figure 3.8 cannot be parallel unless the corresponding angles $\alpha$ and $\gamma$ are congruent.

**(b)** Assuming Playfair's Postulate and neutral geometry, prove Euclid's parallel postulate. (Do this as a proof by contradiction: assume a counterexample to Euclid's parallel postulate exists and reach a contradiction using Playfair's Postulate and the 180° Sum Theorem.)

**(c)** Explain why parts **(a)** and **(b)** prove the equivalence of the two postulates relative to neutral geometry.

# Chapter 4

# Neutral Geometry

In this chapter and the next we'll attempt our own version of what Euclid did in the *Elements* – we'll set up an axiom system for Euclidean geometry. We've seen in Chapter 2 that the "toolbox theorems" of Section 1B provide a useful set of working facts, so proving those theorems will be our primary goal.

This chapter will take us only as far as introducing the "neutral axioms" and proving the theorems of neutral geometry. These will include the triangle congruence criteria, the Isosceles Triangle Theorem, the Perpendicular Bisector Theorem, the Vertical Angles Theorem, and one-half of the Alternate Interior Angles and Corresponding Angles Theorems. In Chapter 5 we will add a parallel postulate and with it prove the rest of the basic Euclidean theorems. But because the theorems in this chapter are proved without reference to a parallel postulate, they will remain in our arsenal for Chapter 6 when we exchange the parallel postulate of Chapter 5 for its negation.

## A. Neutral Axioms, Part I

We have a big job to do to begin this chapter! In this section and the next we will lay the foundation of the axiom system we will use to prove the theorems of neutral geometry. Working carefully now is important, for what we do here will be the basis for our work all the way through Chapter 6. If our approach seems a bit fussier than it has been until now, it is because we want to emphasize the exactness of the axiomatic method in practice. The wording of the axioms must be precise, and any terms used must be either defined or specified as undefined.

Nothing we prove in this section or the next really deserves the title "theorem", for our results now will be much more elementary than the true theorems that come later. But the axioms can't (or shouldn't!) specify everything, and there will be a fair amount of preliminary work to do in building up the basic facts from the axioms. For example, we will want to prove (after we have formally defined what a *ray* is!) that if $C$ is a point on ray $\overrightarrow{AB}$ other than $A$ then ray $\overrightarrow{AC}$ is the same as $\overrightarrow{AB}$. This may seem "obvious", but remember that we are insisting that all conclusions must be supported only by the axioms. If the axioms don't say it, we have to prove it! We will leave some of the verifications of these basic facts as exercises, and you should complete as many of these as possible before moving on to the "real theorems" of neutral geometry given in Sections C through F.

For obvious reasons, the less your axioms specify, the more initial work there is to do. To formulate axioms for Euclidean geometry that are truly minimal (in the sense of being independent, as discussed in Section 3C) would create more work than is appropriate for our purposes here. Instead, we will strike a balance. We will make our axioms simple enough that proving the toolbox theorems from them should strike you as an impressive feat. But at the same time, we will include enough in our axioms to stave off some of the early drudgery that would result from more sparse assumptions.

We begin, appropriately, with undefined terms. As is typical, we take the fundamental notions of set membership ("*element*", "*subset*", etc.) to be undefined. We also take "*point*", "*line*", and "*plane*" to be undefined, with the understanding that the *plane* is a set whose *elements* are *points*, and that *lines* are special *subsets* of the plane. Our first defined term is a simple and familiar concept, and will be needed in our first axiom.

**Definition.** A set of points is *collinear* if there is a line that contains every point in the set. A set that is not collinear is said to be *noncollinear*.

One more notion is needed before we state our first set of axioms. You may have encountered the concept of a *distance function* or *metric* before – certainly you've seen examples, such as the usual "Euclidean distance function" $d((x_1, y_1), (x_2, y_2)) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$ for $\mathbb{R}^2$. Now the Cartesian plane $\mathbb{R}^2$ is in fact the model we have in mind for our geometry, so we want our axioms to describe its properties. But, of course, we can't use this distance formula itself, since it only exists within that one model. (It relies on $(x, y)$ coordinates for points!) To introduce distances to our axiom system, we'll need to use only the abstract properties possessed by a distance function. Thus, recall that a *distance function* on a set $\Pi$ is a function $d : \Pi \times \Pi \longrightarrow [0, \infty)$ that satisfies the following three properties:

**(i)** $d(A, B) = 0$ if and only if $A = B$,

**(ii)** $d(A, B) = d(B, A)$ for all $A$ and $B$, and

**(iii)** for all $A$, $B$, and $C$, we have $d(A, C) + d(C, B) \geq d(A, B)$.

Properties **(i)** and **(ii)** are obvious enough. Property **(iii)**) is called the "triangle inequality" since it can be interpreted as saying that one side length in a triagle is never greater than the sum of the other two side lengths.

We're now ready to state ...

## Axioms on the nature of the plane.

**Axiom P1:** *The plane is a nonempty and noncollinear set.*

**Axiom P2:** *There is a distance function $d$ defined on the plane.*

These meager assumptions don't provide much to go on! The real interest in geometry starts with the behavior of lines, and our next set of axioms will specify the behaviors we assume to be true for our lines.

## Axioms on the nature of lines, part 1.

**Axiom L1:** *Through any two points $A$ and $B$ there is a unique line $\overleftrightarrow{AB}$.*

**Axiom L2:** *Each line can be put in a natural one-to-one correspondence with the real number line relative to the distance function. Specifically, given any two points $A$ and $B$ with $d(A, B) = \delta$, there is a unique one-to-one and onto "*ruler function*" $r : \mathbb{R} \longrightarrow \overleftrightarrow{AB}$ (where we denote $r(t)$ by the point $R_t$) so that*

- *$R_0 = A$,*
- *$R_\delta = B$, and*
- *$d(R_s, R_t) = |s - t|$ for all $s$ and $t$.*

Axiom L1 is, of course, the familiar "two points determine a uniqu line" rule. Surprisingly, it is (along with a simple definition) all we need to prove our first two elementary facts.

**Definition.** Two lines that share no points in common are said to be *parallel*.

**FACT 4.1.** *Any two distinct lines that are not parallel meet in exactly one point.*

*Proof:* Our proof is by contradiction. Let $\Lambda_1$ and $\Lambda_2$ be distinct lines, and assume to reach a contradiction that they share two distinct points $A$ and $B$ in common. The contradiction is immediate: there are two lines ($\Lambda_1$ and $\Lambda_2$) through $A$ and $B$, contrary to Axiom L1. $\square$

**FACT 4.2.** *If $C$ is a point on line $\overleftrightarrow{AB}$ other than $A$ then $\overleftrightarrow{AC} = \overleftrightarrow{AB}$.*

*Proof:* Both $\overleftrightarrow{AC}$ and $\overleftrightarrow{AB}$ contain the points $A$ and $C$, so they must be the same line by Axiom L1. $\square$

Axiom L2 is extremely powerful. It allows us to think of any line as a ruler with respect to the distance function by giving an exact (one-to-one and onto) correspondence between real numbers and points along the ruler. Furthermore, we have the freedom to place the "zero" of the ruler at whatever point of the line we want to. All of the properties of the real numbers are in this way brought to bear on our geometry.

We can now use Axiom L2 to give definitions for some fundamental geometric objects.

**DEFINITIONS.** Let $A$ and $B$ be distinct points, and let $r(t) = R_t$ be a ruler function for $\overleftrightarrow{AB}$ as in Axiom L2 (with $d(A, B) = \delta$, $A = R_0$ and $B = R_\delta$).

- The set of points $\{R_t : 0 \leq t \leq \delta\}$ $(= r([0, \delta]))$ is called the *segment* with *endpoints* $A$ and $B$, and is denoted by $AB$. If $C$ is a point of $AB$ other than one of the endpoints, then we say that $C$ is *between $A$ and $B$*. We denote this in shorthand notation by $A * C * B$.
- The *length* of the segment $AB$ is the number $d(A, B)$ and is denoted by $|AB|$.
- We say that two segments $AB$ and $CD$ are *congruent* (written $AB \cong CD$) if they have the same length.
- The *midpoint* of segment $AB$ is the point $R_{\delta/2}$.
- The set of points $\{R_t : t \geq 0\}$ $(= r([0, \infty)))$ is called the *ray* through $B$ with *endpoint* $A$, and is denoted by $\overrightarrow{AB}$.

To illustrate the care with which one must proceed when using the axiomatic method, let's consider a simple question: given two points $A$ and $B$, is segment $AB$ the same as segment $BA$? Of course it *should* be – but do our definitions and axioms really guarantee it to be so? We'll prove now that the answer is yes.

**FACT 4.3.** *If $A$ and $B$ are any points then $AB = BA$. Also, if $C$ is any point on $\overrightarrow{AB}$ other than $A$ then $\overrightarrow{AC} = \overrightarrow{AB}$.*

*Proof:* We will prove that $AB = BA$ and leave the proof that $\overrightarrow{AC} = \overrightarrow{AB}$ as Exercise 4.3.

We must use the definition of segment to show that the two point sets $AB$ and $BA$ as given by that definition are the same. The problem is that $AB$ and $BA$ are technically defined by different ruler functions – for $AB$ we have the ruler aligned with zero at point $A$ and $\delta = d(A, B)$ at point $B$, while for defining $BA$ we have zero at $B$ and $\delta$ at $A$. But we can discover a relationship between these ruler functions that will enable us to conclude that the two segments contain the same points.

Let $r(t) = R_t$ be the ruler function used in defining segment $AB$ (with $R_0 = A$ and $R_\delta = B$). Define a function $r'$ by $r'(t) = r(\delta - t)$. If we denote $r'(t)$ by $R'_t$ then we can see:

- $R'_0 = R_{\delta-0} = R_\delta = B$,
- $R'_\delta = R_{\delta-\delta} = R_0 = A$, and
- for any $s$ and $t$ we have $d(R'_s, R'_t) = d(R_{\delta-s}, R_{\delta-t})$ which by Axiom L2 is equal to $|(\delta - s) - (\delta - t)| = |s - t|$.

But by Axiom L2 there is a *only one* function having these properties, namely the ruler function for line $\overleftrightarrow{AB}$ in which zero is placed at $B$ while $\delta$ is placed at $A$. Since this is the ruler function that *by definition* is used to describe segment $BA$, we conclude that

$$\begin{aligned} BA &= \{R'_t : 0 \le t \le \delta\} \\ &= \{R_{\delta-t} : 0 \le t \le \delta\} \\ &= \{R_s : \delta \ge s \ge 0\} \\ &= AB \end{aligned}$$

□

We need one more definition before introducing our remaining axioms on lines.

**Definition.** A set $\Sigma$ of points is said to be *convex* if whenever $A$ and $B$ are points of $\Sigma$, the segment $AB$ lies entirely in $\Sigma$.

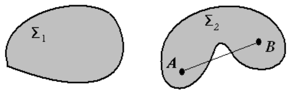



Figure 4.1:

Figure 4.1 illustrates this definition. The set $\Sigma_1$ on the left is convex because no matter which two points of $\Sigma_1$ are chosen, the entire segment between them will lie in $\Sigma_1$. However, the set $\Sigma_2$ on the right is *not* convex since there exist two points $A$ and $B$ in $\Sigma_2$ so that the segment $AB$ is not contained entirely in $\Sigma_2$. A fundamental property of convex sets is given in the following fact. The proof is left as an easy exercise (Exercise 4.4).

**Fact 4.4.** *The intersection of any collection of convex sets is itself a convex set.*

We now introduce the remaining axioms covering the properties of lines. Axiom L3 encodes the property that shortest distances in the Euclidean plane occur along lines, while Axiom L4 covers the way in which lines are embedded in the plane.

## Axioms on the nature of lines, part 2.

**Axiom L3:** *For any points $A$ and $B$, if $C$ is a point such that $|AC| + |CB| = |AB|$, then $C$ is on segment $AB$.*

**Axiom L4:** *For each line $\Lambda$, the set of points not on $\Lambda$ is the union of two disjoint nonempty convex sets $\Sigma_1$ and $\Sigma_2$ called the* halfplanes determined by $\Lambda$. *If $A \in \Sigma_1$ and $B \in \Sigma_2$ then the segment $AB$ will intersect $\Lambda$.*

Suppose that $\Lambda$ is a line and that $\Sigma_1$ and $\Sigma_2$ are the halfplanes determined by $\Lambda$, as in Axiom L4. If points $A$ and $B$ are not on $\Lambda$ then we may refer to them as being on either the *same side of* $\Lambda$ (if both belong to the same halfplane determined by $\Lambda$) or on *different sides of* $\Lambda$ (if each halfplane contains one of the points). We leave it as Exercise 4.5 to prove the following simple characterization.

**Fact 4.5.** *If $A$ and $B$ are points not on the line $\Lambda$ then $A$ and $B$ are on opposite sides of $\Lambda$ if and only if the segment $AB$ intersects $\Lambda$.*

The halfplanes from Axiom L4 are what we might call "open" halfplanes since they do not include the points on the line $\Lambda$. Sometimes we will need to use the "closed" halfplane consisting of one of these open halfplanes together with the points on $\Lambda$. To be exact:

**Definition.** Let $\Lambda$ be a line and let $\Sigma$ be one of the halfplanes determined by $\Lambda$. We will use the symbol $\overline{\Sigma}$ to denote $\Sigma \cup \Lambda$ and will call this a *closed halfplane* determined by $\Lambda$.

Again, note the attention to detail and allegiance to the axioms in the following short proof.

**Fact 4.6.** *If $\Sigma$ is a halfplane determined by line $\Lambda$, $A$ is a point on $\Lambda$, and $B$ is any other point in the closed halplane $\overline{\Sigma}$, then the ray $\overrightarrow{AB}$ is contained entirely in $\overline{\Sigma}$.*

*Proof:* Let $\Lambda$, $\Sigma$, $A$, and $B$ be as in the statement above. First note that if $B$ is on $\Lambda$ then by the definition of a ray, $\overrightarrow{AB}$ is a subset of $\Lambda$ and is thus clearly contained in $\overline{\Sigma}$. So, we may assume that $B$ is in $\Sigma$.

Assume to reach a contradiction that $\overrightarrow{AB}$ contains a point $C$ that is not in $\overline{\Sigma}$.

- Then $C$ must be contained in the other halfplane determined by $\Lambda$.
- Then by Axiom L4 the segment $BC$ must contain a point $D$ of $\Lambda$.
- But from the definition of a ray it is impossible for $A$ to be between two points of $\overrightarrow{AB}$, so $D$ and $A$ are different points.
- Thus, $\overrightarrow{AB}$ contains two points of $\Lambda$, so $\Lambda$ and $\overleftrightarrow{AB}$ share at least two points.
- This contradicts Fact 4.1 because $\Lambda$ and $\overleftrightarrow{AB}$ are clearly distinct ($B$ is not on $\Lambda$) and not parallel ($A$ is on both). This contradiction shows that point $C$ cannot exist – that is, $\overrightarrow{AB}$ cannot contain a point not in $\overline{\Sigma}$. $\square$

Now suppose that $P_1P_2 \cdots P_n$ is a convex polygon (see definitions of polygons on p.1.6). We can define the *interior* of $P_1P_2 \cdots P_n$ to be the intersection of all the halfplanes determined by the lines through consecutive vertices that contain the remaining vertices. Fittingly enough, because of Fact 4.4 and our assumption in Axiom L4 that halfplanes are convex, the interior of a convex polygon is a convex set!

Remember that our goal in this section (and the next) is to set forth axioms that will capture the essence of our intuitive notion of what the two-dimensional plane is like. Towards this end, our next axiom will address properties of angles, angle measure, and circles. We will of course need to define angles and circles first – the definition of angle measure will come out of our axiom. The definitions of *circle*, *center*, *radius*, points *inside* a circle, and points *outside* a circle as given in Section 1B (see p.19) are certainly adequate for our use, so we will now consider them to be incorporated into our axiom system. For angles, we give the following definitions.

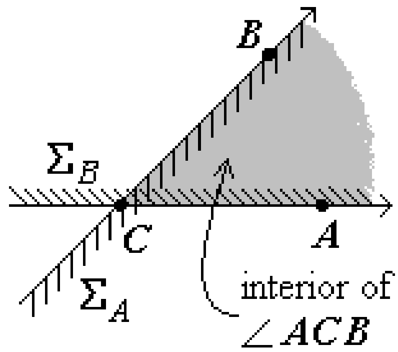



Figure 4.2:

**Definitions.** Let $A$, $B$, and $C$ be points with $A$ and $B$ both distinct from $C$. The *angle* $\angle ACB$ is the union of the two rays $\overrightarrow{CA}$ and $\overrightarrow{CB}$. The point $C$ is called the *vertex* of this angle. If $A$, $B$, and $C$ are noncollinear then the *interior* of the angle $\angle ACB$ is the set $\Sigma_B \cap \Sigma_A$ where $\Sigma_B$ is the side of $\overleftrightarrow{CA}$ containing $B$ and $\Sigma_A$ is the side of $\overleftrightarrow{CB}$ containing $A$. (See Figure 4.2.)

Note from the definition that $\angle ACB$ and $\angle BCA$ both denote the same angle. Furthermore, as noted in Fact 4.3, rays have many names. Thus, if $D$ is a point on $\overrightarrow{CA}$ (other than $C$) then $\angle DCB = \angle ACB$. Finally, as noted above for the interior of a convex polygon, Axiom L4 and Fact 4.4 together guarantee that the interior of an angle is a convex set.

Now consider a circle $\Gamma$ with center at point $C$ and radius $r > 0$. If $D$ is any point other than $C$ then we may define the ray $\overrightarrow{CD}$ as we did following Axiom L2, say $\overrightarrow{CD} = \{R_t : t \geq 0\}$ where $r(t) = R_t$ is a ruler function for $\overleftrightarrow{CD}$ with $R_0 = C$ and $R_{|CD|} = D$. By Axiom L2, this ray contains exactly one point, namely $R_r$, that is distance $r$ from $C$ and is thus on $\Gamma$ (see Figure 4.3). This shows that each ray with endpoint $C$ contains exactly one point of $\Gamma$, so there is a natural one-to-one correspondence between points of $\Gamma$ and rays with endpoint $C$. Since an angle with vertex $C$ is the union of two such rays, we can associate each such angle with a pair of points on $\Gamma$. We'll be able to use this association, along with the following powerful axiom, to define the measure of an angle. Just as Axiom L2 brought the properties of the real number line to our geometric lines, this axiom transfers some of those properties to circles.

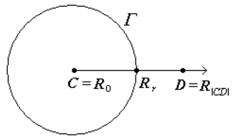



Figure 4.3:

## Axiom on angles and circles.

**Axiom AC:** *Let $\Gamma$ be a circle with center $C$ and radius $r > 0$. For each point $A$ on $\Gamma$ there is a designated "protractor function" $p : \mathbb{R} \longrightarrow \Gamma$ (where we denote $p(\alpha)$ by the point $P_\alpha$) with the following properties:*

- $P_0 = A$.
- *$p$ is one-to-one and onto from* $(-180, 180]$ *to* $\Gamma$.
- *$p$ is periodic, with $P_{\alpha+360} = P_\alpha$ for all $\alpha$.*
- *$P_\alpha * C * P_{\alpha+180}$ for all $\alpha$.*
- *The portions of $\Gamma$ contained in the two halfplanes determined by $\overleftrightarrow{CA}$ are $\{P_\alpha : -180 < \alpha < 0\}$ and $\{P_\alpha : 0 < \alpha < 180\}$. Also, if $0 < \beta < 180$ then the portion of $\Gamma$ in the interior of $\angle P_\beta CA$ is $\{P_\alpha : 0 < \alpha < \beta\}$.*
- *If $B = P_\beta$ then the protractor function for $B$ is given by $q(\alpha) = p(\alpha + \beta)$.*

This axiom (along with the definitions below) is pictured in Figure 4.4. The "protractor function" $p$ can be thought of as "wrapping $\mathbb{R}$ around $\Gamma$" in such a way that zero corresponds to point $A$ (that is, $P_0 = A$) and the circle repeats every 360 units on $\mathbb{R}$. Of course, we chose the range 360 because of its familiar association with degree measures, but this was really an arbitrary choice. We could just as easily define the protractor to give the range $[0, 100]$ or $[0, 2\pi]$ to a circle. When we define angle measures below, we will intentionally *not* use the symbol $^\circ$ just to emphasize that our angle measures are just numbers, and that the range for those numbers is a consequence of the choice made in this axiom rather than a pre-formed mental concept.

The last item in the axiom is worth mentioning. The axiom states that we can find a protractor function placing the zero of the protractor at any point of the circle we want to. The last item, however, says that any two such protractor functions are closely related – the protractor function that puts the zero at point $B$ is really just a shift (or "rotation" as it might be thought of) of the protractor function that puts the zero at point $A$. In fact, if (as in the axiom statement) we have a protractor function $p(\alpha) = P_\alpha$ with $P_0 = A$ and $P_\beta = B$, then the function $q(\alpha) = p(\alpha + \beta)$ clearly satisfies $q(0) = p(0 + \beta) = P_\beta = B$.

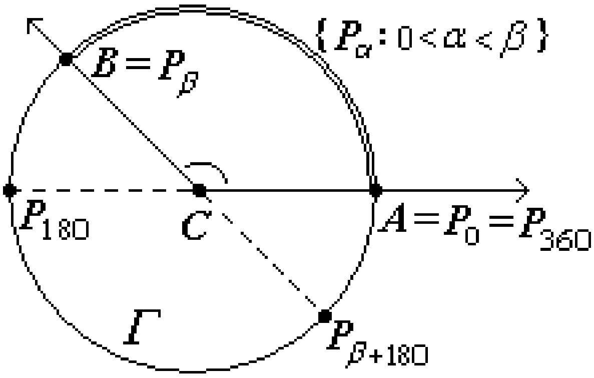



Figure 4.4:

Now, to define angle measure, recall that as we deduced before stating the axiom, each angle with vertex $C$ can be written $\angle ACB$ where $A$ and $B$ are points on $\Gamma$. Suppose as above that we have defined $p(\alpha) = P_\alpha$ to be the protractor function with $P_0 = A$ and that $B$ is $P_\beta$ on that protractor, with $-180 < \beta \leq 180$. (Note that we're guaranteed that $B$ will be equal to $P_\beta$ for exactly one $\beta$ in this range since (according to Axiom AC) the protractor function is one-to-one and onto from $(-180, 180]$ to $\Gamma$.

**Definition.** Under the situation described above, we define the *measure* of angle $\angle ACB$ to be $m\angle ACB = |\beta|$.

Note that with this definition, the measure of an angle will always be in the range $0 \leq m\angle ACB \leq 180$. This rather limited notion of angle measure will be sufficient for everything we need to prove. Arcs (which we define immediately below) will be allowed to have measures exceeding 180.

**Definition.** The angles $\angle ACB$ and $\angle DFE$ are *congruent* (written $\angle ACB \cong \angle DFE$) if they have equal angle measures.

**Definition.** We say that an angle is *acute* if its measure is less than 90, *obtuse* if its measure is greater than 90, and *right* if its measure is exactly 90. We say that lines $\overleftrightarrow{CA}$ and $\overleftrightarrow{CB}$ are *perpendicular* if $\angle ACB$ is a right angle. (We may also use the word *perpendicular* applied to segments and rays if the corresponding lines containing them are perpendicular by this definition.)

**Definition.** If $p(\alpha) = P_\alpha$ is a protractor function for circle $\Gamma$ then a set of the form $\{P_\alpha : \beta_1 < \alpha < \beta_2\}$ (where $-180 < \beta_1 < \beta_2 \leq 180$) is called an *arc* of $\Gamma$, and the measure of this arc is defined to be $\beta_2 - \beta_1$. If $Q$ is any point on this arc other than $P_{\beta_1}$ or $P_{\beta_2}$ then we will often denote the arc by the symbol $P_{\beta_1}\overset{\frown}{Q}P_{\beta_2}$.

Axiom AC should be thought of as establishing circles as "protractors". Its implications include the basic behaviors of angle measure that our experience with geometry dictates should be true. For example, the facts below can be proved from this axiom without too much trouble.

## Facts 4.7.

**(i)** *Given any line $\overleftrightarrow{AC}$, any halfplane $\Sigma$ determined by this line, and any number $\beta$ between 0 and 180, there is a unique ray $\overrightarrow{CB}$ in $\overline{\Sigma}$ with $m\angle ACB = \beta$.*

**(ii)** *If $D$ is a point in the interior of angle $\angle ACB$ then $m\angle ACB = m\angle ACD + m\angle DCB$.*

**(iii)** *If $\angle ACB$ is a right angle, $A'$ is any point of $\overleftrightarrow{AC}$ other than $C$, and $B'$ is any point of $\overleftrightarrow{BC}$ other than $C$, then $\angle A'CB'$ is also a right angle.*

*Proof:* Let $\Gamma$ be a circle centered at $C$ and passing through $A$ and let $p(t) = P_t$ be the protractor function for $\Gamma$ with $P_0 = A$. (That is, think of $\Gamma$ as a protractor for angles at vertex $C$, with the ray $\overrightarrow{CA}$ at the zero position.

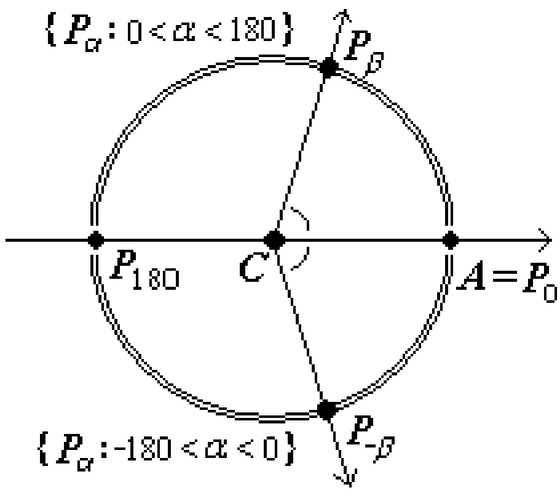



Figure 4.5:

We can now easily give a descriptive proof for part **(i)** by observing that (according to Axiom AC) each of the two halfplanes determined by $\overleftrightarrow{AC}$ contain one of the arcs $\{P_\alpha : -180 < \alpha < 0\}$ and $\{P_\alpha : 0 < \alpha < 180\}$ (see Figure 4.5). But then (by the definition of angle measure and the properties of our protractor as set forth in the axiom) for each of these arcs there is exactly one choice for $B$ (either $B = P_\beta$ or $B = P_{-\beta}$) which will make $m\angle ACB = \beta$.

We give a detailed proof of part **(ii)** in the following steps. Refer to Figure 4.6.

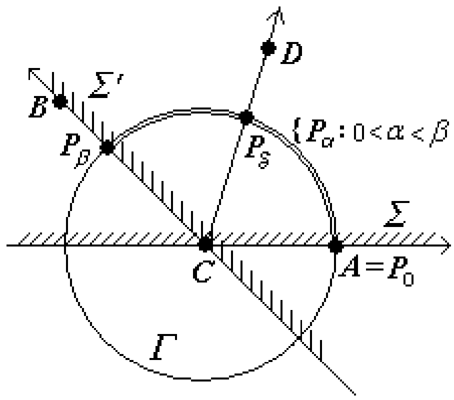



Figure 4.6:

- As we have observed, the ray $\overrightarrow{CB}$ meets $\Gamma$ in a single point, say $P_\beta$. (We will assume that $0 < \beta < 180$ – the case $-180 < \beta < 0$ is similar.)
- Thus $\overrightarrow{CB} = \overrightarrow{CP_\beta}$ by Fact 4.3, so $\angle ACB = \angle ACP_\beta$.
- Similarly, the ray $\overrightarrow{CD}$ meets $\Gamma$ in a point $P_\delta$.
- Let $\Sigma$ be the halfplane determined by $\overleftrightarrow{CA}$ that contains $P_\beta$ and let $\Sigma'$ be the halfplane determined by $\overleftrightarrow{CP_\beta}$ that contains $A$ (see Axiom L4). Now by definition, the interior of $\angle ACP_\beta$ is $\Sigma \cap \Sigma'$, so $D$ must be in both of these halfplanes.
- So, by Fact 4.6, the ray $\overrightarrow{CD}$ lies entirely in both $\overline{\Sigma}$ and $\overline{\Sigma'}$.
- Thus $\overrightarrow{CD}$ lies, except for the point $C$, in both $\Sigma$ and $\Sigma'$.
- It follows that $\overrightarrow{CD}$ is itself contained in the interior of $\angle ACP_\beta$ except for point $C$. In particular, point $P_\delta$ is in the interior of $\angle ACP_\beta$.
- But by Axiom AC, the portion of $\Gamma$ in the interior of $\angle ACP_\beta$ is the arc $\{P_\alpha : 0 < \alpha < \beta\}$, so we must have $0 < \delta < \beta$.

- But now by the definition of angle measure, we have $m\angle ACD = m\angle P_0CP_\delta = \delta - 0 = \delta$ and $m\angle DCB = m\angle P_\delta CP_\beta = \beta - \delta$, and $m\angle ACB = m\angle P_0CP_\beta = \beta - 0 = \beta$.

- So, $m\angle ACD + m\angle DCB = \delta + (\beta - \delta) = \beta = m\angle ACB$.

We leave the proof of part **(iii)** as Exercise 4.7. □

The reason for including the above facts is definitely *not* for their intrinsic interest, for neither their statements nor their proofs are very striking. Instead, we include them just as an indication of the kind of work that must be done when starting a formal axiomatic treatment. These facts and many more of the same caliber will be needed many times in the coming chapters. (See Exercise 4.8 for another example of the kind of situation that can arise.) But to continue writing detailed formal proofs of all these would slow our progress to a crawl and obscure the central ideas of our development. Too much attention to detail is stifling, and too much rigor can become *rigor mortis*! We will usually be content to check that we *could*, if asked to, produce a proof of such facts from the axioms. Such checks will usually be mental and will usually be left up to the reader. The rather finicky derivations in this section (that we hope were instructive to see at least once) should convince you that this compromise is a wise one.

## Exercises

**4.1**. Which of the seven axioms introduced in this section are valid in spherical geometry? Justify your answer with explanations and counterexamples.

**4.2**. Prove that if $C$ and $D$ are any two distinct points on $\overleftrightarrow{AB}$ then $\overleftrightarrow{CD} = \overleftrightarrow{AB}$.

**4.3**. Complete the proof of Fact 4.3 by showing that if $C$ is a point on $\overrightarrow{AB}$ other than the endpoint $A$ then $\overrightarrow{AC} = \overrightarrow{AB}$.

**4.4**. Prove Fact 4.4

**4.5**. Prove Fact 4.5.

**4.6**. Recall that Euclid's axioms (postulates) were motivated by straightedge and compass constructions. Given a ray $\overrightarrow{CA}$, a side $\Sigma$ of $\overleftrightarrow{CA}$, and an angle $\angle DFE$, part **(i)** of Fact 4.7 allows us to construct a point $B$ in $\Sigma$ with $m\angle ACB = \beta = m\angle DFE$. Euclid certainly did not include this as a fundamental construction step (as one of his axioms). Show, however, that such a point $B$ *can* be constructed with the usual Greek straightedge and compass rules.

**4.7**. Prove part **(iii)** of Fact 4.7.

**4.8**. Let points $A$ and $B$ lie on the same side of $\overleftrightarrow{CD}$, and let $E$ be a point on $\overleftrightarrow{CD}$ with $D * C * E$. Suppose also that $E$ and $B$ are on the same side of $\overleftrightarrow{AC}$. Prove that $A$ and $D$ are on the same side of $\overleftrightarrow{BC}$ and that $E$ is on the opposite side. (You will likely find Axiom AC to be necessary.)

# B. Neutral Axioms, Part II

Section A presented seven axioms of neutral geometry. In this section we'll complete the set of axioms by introducing five more.

Recall one of the flaws we discussed in the *Elements*: Euclid assumes that the two circles he uses to construct an equilateral triangle will intersect each other (see p.113) when in fact his axioms do not guarantee anything about intersections. Our next set of axioms will take care of this difficulty for us.

To understand Axiom C1 below, refer to Figure 4.7. The axiom states that if we move a point continuously along a line or a circle (using a ruler function $r(t) = R_t$ or a protractor function $p(\alpha) = P_\alpha$), that point's distance from a fixed point as well as the measures of the angles it forms with two fixed points will vary continously with the motion. "Continuity" is the appropriate word for Axioms C2 through C4 for a less direct reason – these axioms imply that our lines and circles are in some sense "without gaps".

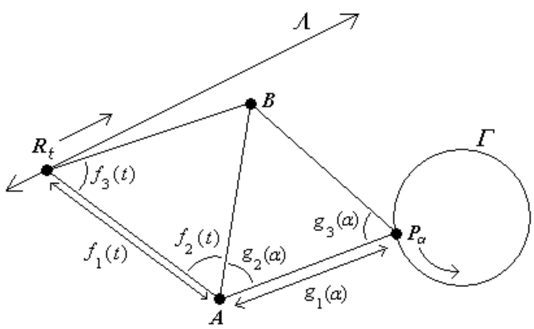



Figure 4.7:

## Axioms on continuity.

**Axiom C1:** *The "ruler functions" of Axiom L2 and the "protractor functions" of Axiom AC behave continuously with respect to lengths and angle measures. Specifically, if* $r(t) = R_t$ *is a ruler function for a line,* $p(\alpha) = P_\alpha$ *is a protractor function for a circle, and* $A$ *and* $B$ *are points, then the functions* $f_1(t) = |AR_t|$, $f_2(t) = m\angle BAR_t$, $f_3(t) = m\angle AR_tB$, $g_1(\alpha) = |AP_\alpha|$, $g_2(\alpha) = m\angle BAP_\alpha$, *and* $g_3(\alpha) = m\angle AP_\alpha B$ *are continuous wherever they are defined (see Figure 4.7).*

**Axiom C2:** *If a line* $\Lambda$ *contains a point inside a circle* $\Gamma$ *then* $\Lambda$ *intersects* $\Gamma$ *in at least two points.*

**Axiom C3:** *If* $\Gamma_1$ *and* $\Gamma_2$ *are circles, and* $\Gamma_1$ *contains a point inside* $\Gamma_2$ *and a point outside* $\Gamma_2$*, then the two circles intersect in at least two points.*

**Axiom C4:** *If* $D$ *is a point in the interior of angle* $\angle ACB$ *then ray* $\overrightarrow{CD}$ *intersects segment* $AB$*.*

Note that Axioms C2 and C3 say less than we actually expect to be true. Experience with the Euclidean model suggests that both of these statements should specify "*exactly two points*" of intersection rather than merely "*at least two*". In fact, later this chapter (once we have established a few theorems), we will be able to prove both of the "*exactly two points*" versions – see Exercises 4.20 and 4.23. We left the axioms this way simply because assuming less is always better, and assuming more in this case would not really save us any work.

However, it's worth pointing out here that we actually don't need Axioms C2 or C3 at all, if we're willing to do some extra work. For both of these can in fact be proved from Axiom P2, Axiom C1, and the Intermediate Value Theorem (from your calculus class) – see Exercises 4.10 and 4.11.

Axiom C4 is often called the "*Crossbar Axiom*" (or "*Crossbar Theorem*", depending on its status in the axiom system), and often plays a crucial role in axiom systems for geometry. As with Axioms C2 and C3, we have placed it as an axiom as a convenience – it can actually be proved from our previous axioms along with Axiom C1.

Another famous elementary statement in geometric axiom systems is *Pasch's*

*Theorem.*[1] While it is actually (in our approach) a fairly easy consequence of Axiom L4, we place it here because it has a similar feel to the continuity axioms. Its statement involves our first use of the term *triangle*, so this is a good time to adopt the definitions of *triangle, polygon*, and *quadrilateral* (along with their varieties and constituent parts) as given in Section 1B into our current axiom system. We also adopt the Chapter 1 definition of congruence for triangles (see p.14), including the notational distinction of $\triangle ABC$ as indicating an order to the vertices (so that corresponding parts in a congruence can be specified). We leave the easy proof of Pasch's Theorem as an exercise.

FACT **4.8.** *Let $\Lambda$ be a line not containing any of the vertices of triangle $ABC$. Then $\Lambda$ intersects either no sides of $ABC$ or exactly two sides of $ABC$.*

We have one more axiom to introduce. One of the flaws in the *Elements* we discussed in Section 3C was the Euclid's "proof" of the Side-Angle-Side congruence criterion. In fact, Euclid would have been better advised to leave SAS as an axiom, which is exactly what we will do.

## Side-Angle-Side Axiom.

**Axiom SAS:** *If $ABC$ and $DEF$ are triangles with $CA \cong FD$, $CB \cong FE$, and $\angle C \cong \angle F$ then $\triangle ABC \cong \triangle DEF$.*

We close this introduction to our axiom system with a very useful basic fact, whose proof is made possible by Axiom SAS.

FACT **4.9.** *Let $\overleftrightarrow{AC}$ be any line, $\Sigma$ one of the halfplanes it determines, and $\triangle DEF$ a triangle with $FD \cong CA$. There is a unique point $B$ in $\Sigma$ so that $\triangle ABC \cong \triangle DEF$.*

---

[1]It is also sometimes called *Pasch's Axiom.* Moritz Pasch, a prominent German mathematician of the late 19th century, published an axiomatic treatment of geometry in 1882 in which this appears as one of the axioms. Hilbert built on Pasch's work (and the work of a few of Pasch's contemporaries) to produce his axiom system for geometry that we discussed in Section 3C.

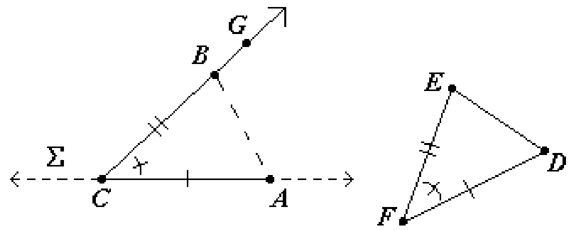



Figure 4.8:

*Proof:* There are two parts to the proof. First we must show that there is such a point $B$ in $\Sigma$. (Refer to Figure 4.8.)

- By Fact 4.7 **(i)** there is a ray $\overrightarrow{CG}$ in $\overline{\Sigma}$ so that $m\angle ACG = m\angle DFE$.
- Then by the Axiom L2 there is a point $B$ on $\overrightarrow{CG}$ so that $|CB| = |FE|$.
- But now by Axiom SAS we have $\triangle ABC \cong \triangle DEF$.

Now we must show that this point is unique. So let $B$ be the point as constructed above, and let $B'$ be any point in $\Sigma$ satisfying the condition $\triangle AB'C \cong \triangle DEF$. We will show that $B' = B$ (so that $B$ is in fact the *only* point of $\Sigma$ having this property).

- First note that we have $\triangle AB'C \cong \triangle DEF \cong \triangle ABC$, so $\triangle AB'C$ and $\triangle ABC$ are congruent to each other.
- From this congruence we know $\angle ACB' \cong \angle ACB$.
- But both $B$ and $B'$ are in $\Sigma$, so by Fact 4.7 **(i)** it must be that $\overrightarrow{CB'} = \overrightarrow{CB}$.
- Also from the congruence, $CB' \cong CB$, so $|CB'| = |CB|$.
- By Axiom L2 there is only one point on $\overrightarrow{CB}$ at distance $|CB|$ from $C$, so $B' = B$ as claimed. $\square$

## Exercises

**4.9**. Show that Axiom C3 would allow Euclid to prove the two circles in his construction of an equilateral triangle really do intersect each other.

**4.10**. Show how to use Axiom P2, Axiom C1, and the Intermediate Value Theorem to prove Axiom C2.

**4.11**. Show how to use Axiom P2, Axiom C1, and the Intermediate Value Theorem to prove Axiom C3.

**4.12**. Prove Fact 4.8.

## C. Angles

With our axioms in place and some basic facts established, we can now move to the real theorems of our axiom system. In this section we'll prove the first of the basic toolbox geometry theorems as set forth in Section 1B, namely the Vertical Angles Theorem and (half of) the Alternate Interior Angles and Corresponding Angles Theorems.

A quick reference back to Section 1B, however, should raise an objection: the definitions of these types of angles given there are inadequate for our resent purposes, for they rely on a diagram to carry their meaning. That isn't hard to remedy, and we invite you to check that the definitions given below do the job of formalizing the intuitive idea conveyed by the earlier figures.

DEFINITION. Suppose that the distinct lines $\overleftrightarrow{AB}$ and $\overleftrightarrow{DE}$ intersect at point $C$, with $A * C * B$ and $D * C * E$. The angles $\angle ACD$ and $\angle BCE$ are called *vertical angles* of this intersection.

DEFINITION. We say that line $\Lambda_0$ *transverses* the lines $\{\Lambda_1, \Lambda_2, \ldots, \Lambda_k\}$ if it intersects each one of these lines but is not equal to any of them. We refer to this as a *transversal* of $\Lambda_0$ with these lines.

DEFINITIONS. Let $C$ be a point on $\overleftrightarrow{AB}$ and $F$ a point on $\overleftrightarrow{DE}$ so that $A * C * B$ and $D * F * E$. Also suppose $A$ and $E$ are on different sides of $\overleftrightarrow{CF}$, and let $G$ be a point on $\overleftrightarrow{CF}$ with $C * F * G$. (See Figure 4.9.) Then in the transversal of $\overleftrightarrow{CF}$ over $\overleftrightarrow{AB}$ and $\overleftrightarrow{DE}$, the angles $\angle ACF$ and $\angle EFC$ are called *alternate interior angles*, and the angles $\angle ACF$ and $\angle DFG$ are called *corresponding angles*.

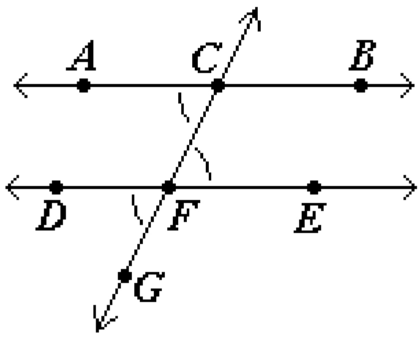



Figure 4.9:

Crucial to proving the Vertical Angles Theorem will be the notion of supplementary angles.

**DEFINITIONS..** Angles $\angle ACB$ and $\angle DFE$ are called *supplementary* (or *supplements* of one another) if and only if $m\angle ACB + m\angle DFE = 180$ and *complementary* (or *complements* of one another) if and only if $m\angle ACB + m\angle DFE = 90$.

Now if $A$, $B$, and $C$ are collinear with $A * C * B$ and if $D$ is any point in one of the halfplanes determined by that line then (from the way angle measure was defined in Section A) $m\angle ACD + m\angle DCB = 180$, so $\angle ACD$ and $\angle DCB$ are supplementary. That observation leads easily to the proof below.

**THEOREM 4.10.** *Suppose $\overleftrightarrow{AB}$ and $\overleftrightarrow{DE}$ meet at point $C$ where $A * C * B$ and $D * C * E$. Then the vertical angles $\angle ACD$ and $\angle BCE$ are congruent.*

*Proof:* The angles $\angle ACD$ and $\angle DCE$ are both (by the above observation) supplementary angles to $\angle DCB$, and thus both have measure $180 - m\angle DCB$. $\square$

We can also prove the other half of the Vertical Angles Theorem.

**THEOREM 4.11.** *Suppose $C$ is a point on $\overleftrightarrow{AB}$ between $A$ and $B$. Suppose also that $D$ and $E$ are points on opposite sides of $\overleftrightarrow{AB}$ and that $\angle ACD \cong \angle BCE$. Then $C$, $D$, and $E$ are collinear with $D * C * E$.*

*Proof.* Refer to Figure 4.10.

- Let $F$ be a point on $\overleftrightarrow{CE}$ so that $F * C * E$. Then we can apply Theorem 4.10 to the intersecting lines $\overleftrightarrow{AB}$ and $\overleftrightarrow{CE}$ to conclude $m\angle ACF = m\angle BCE$.

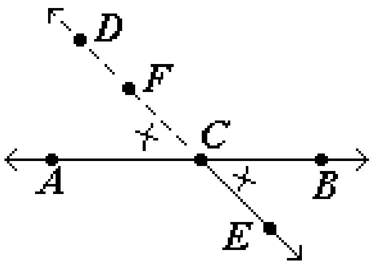



Figure 4.10:

- Then $D$ and $F$ are on the same side of $\overleftrightarrow{AB}$ (since they are both on the side opposite point $E$) and $m\angle ACD = m\angle ACF$.

- But there is only one ray on that side of $\overleftrightarrow{AB}$ making that angle with $\overrightarrow{CA}$ (see Fact 4.7 **(i)**), so $D$ must be on $\overrightarrow{CF}$.

- Since $\overrightarrow{DF}$ is a subset of $\overleftrightarrow{CE}$ this shows that $C$, $D$, and $E$ are collinear.

- Now $D$ and $E$ are given to be on opposite sides of $\overleftrightarrow{AB}$, so $DE$ must meet $\overleftrightarrow{AB}$ at some point. Since we know $\overleftrightarrow{AB}$ and $\overleftrightarrow{DE}$ share point $C$ in common (and cannot share any point other than $C$), it must be true that $C$ is on $DE$, so $D * C * E$. □

The Alternate Interior Angles and Corresponding Angles Theorems are vitally important tools in basic geometry. A high school course in geometry will often present these essentially as axioms, giving them no proofs. We intend to prove them, and will begin that process now. But both of these theorems are "if and only if" statements, and only half of each of them is a theorem of neutral geometry. The converses will not be proved until Chapter 5.

**THEOREM 4.12.** *Suppose that line $\Lambda_0$ transverses lines $\Lambda_1$ and $\Lambda_2$ and that this transversal has a pair of congruent alternate interior angles. Then $\Lambda_1 \parallel \Lambda_2$.*

*Proof.* Let $\Lambda_0$ intersect lines $\Lambda_1$ and $\Lambda_2$ in points $A$ and $B$, respectively. Let $C$ be a point on $\Lambda_1$ and $D$ be a point on $\Lambda_2$ so that $C$ and $D$ are on opposite sides of $\Lambda_0$ (see Figure 4.11 below). Assume (to reach a contradiction) that $\angle BAC \cong \angle ABD$ (a pair of congruent alternate interior angles!) but that $\Lambda_1$ and $\Lambda_2$ are *not* parallel. Then there must be a point $E$ on both $\Lambda_1$ and $\Lambda_2$, and we may assume (without loss of generality) that this point is on the same side of $\Lambda_0$ as $D$. See Figure 4.11 for reference in the following steps.

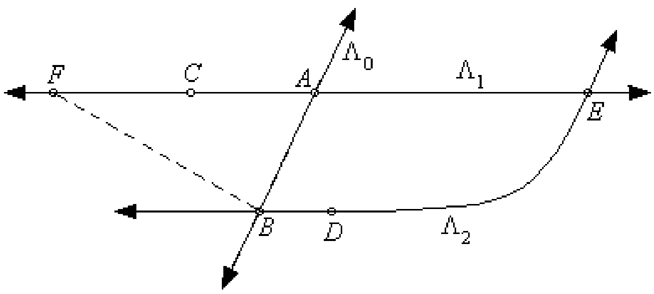



Figure 4.11:

- By Axiom L2 we may find a point $F$ on the ray $\overrightarrow{AC}$ so that $|AF| = |BE|$.
- We have $\overrightarrow{AF} = \overrightarrow{AC}$ and $\overrightarrow{BE} = \overrightarrow{BD}$, so therefore $\angle BAF = \angle BAC$ and $\angle ABE = \angle ABD$
- So, we have the following congruences:

$$\angle BAF \cong \angle ABE$$
$$BA \cong AB$$
$$AF \cong BE$$

- By Axiom SAS we have $\triangle BAF \cong \triangle ABE$.
- Then $\angle FBA \cong \angle EAB$ since these are corresponding angles in congruent triangles.
- Angle $\angle BAF$ is a supplement to $\angle EAB$, so we have:

$$m\angle FBA = m\angle EAB = 180 - m\angle BAF = 180 - m\angle ABE$$

- This means that point $F$ is on $\overleftrightarrow{BE} = \Lambda_2$.
- We now have our contradiction: points $E$ and $F$ are common to both of the distinct lines $\Lambda_1$ and $\Lambda_2$, contrary to Axiom L1.
- This contradiction must have arisen from our assumption that lines $\Lambda_1$ and $\Lambda_2$ are not parallel. So, these lines must in fact be parallel. □

**Corollary 4.13.** *If lines $\Lambda_1$ and $\Lambda_2$ are each perpendicular to the line $\Lambda_3$, then they must be parallel to each other.*

*Proof.* This follows immediately from the last theorem and observation that in the transversal of $\Lambda_3$ over $\Lambda_1$ and $\Lambda_2$ the alternate interior angles are all right angles and thus are congruent. □

The neutral half of the Corresponding Angles Theorem is an easy consequence of Theorems 4.10 and 4.12. We leave the proof to the reader.

THEOREM **4.14.** *Suppose that line $\Lambda_0$ transverses lines $\Lambda_1$ and $\Lambda_2$ and that this transversal has a pair of congruent corresponding angles. Then $\Lambda_1 \parallel \Lambda_2$.*

Playfair's Postulate (see p.118) will become one of our (Euclidean geometry) axioms in Chapter 5. While we can't prove in neutral geometry that through each point $P$ not on a line $\Lambda$ there is *exactly one* line parallel to $\Lambda$, we *can* prove that there is *at least one* such parallel through $P$.

COROLLARY **4.15.** *If $\Lambda$ is a line and $P$ is a point not on $\Lambda$ then there exists at least one line through $P$ which is parallel to $\Lambda$.*

*Proof:* To prove this statement we need only construct a line through $P$ which is parallel to $\Lambda$. See Figure 4.12 for reference in the following steps.

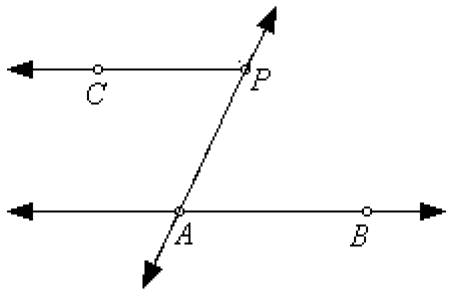



Figure 4.12:

- Let $A$ and $B$ be points on $\Lambda$ and construct the line $\overleftrightarrow{AP}$.
- There is (by Fact 4.7 **(i)**) a ray $\overrightarrow{PC}$ on the side of $\overleftrightarrow{AP}$ not containing point $B$ such that $m\angle CPA = m\angle BAP$.
- We now have line $\overleftrightarrow{AP}$ transversing lines $\Lambda = \overleftrightarrow{AB}$ and $\overleftrightarrow{PC}$ with congruent alternate interior angles, so line $\overleftrightarrow{PC} \parallel \Lambda$ by Theorem 4.12. $\square$

## Exercises

**4.13**. Discuss the following questions.

**(a)** Is Corollary 4.13 true in spherical geometry?

**(b)** Is the Vertical Angles Theorem true in spherical geometry?

**4.14**. Prove Theorem 4.14.

# D. Triangles, Part I

We have adopted the SAS congruence criterion for triangles as an axiom. We are now ready to prove the ASA and SAA criteria as theorems in our axiom system. The SSS criterion will wait until Theorem 4.23.

THEOREM **4.16.** *Suppose that $ABC$ and $DEF$ are triangles and that the following congruences are known: $AB \cong DE$, $\angle A \cong \angle D$, and $\angle B \cong \angle E$. Then $\triangle ABC \cong \triangle DEF$.*

*Proof:* Assume (to reach a contradiction) that triangles $\triangle ABC$ and $\triangle DEF$ are *not* congruent. Our strategy will be to build a triangle $\triangle DEF'$ on the segment $DE$ which *is* congruent to $\triangle ABC$, and to use this triangle to reach a contradiction. See Figure 4.13 for reference in the following steps.

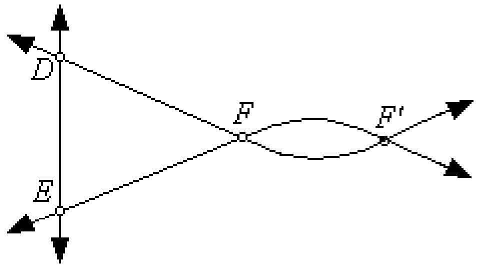



Figure 4.13:

- By Fact 4.9 we can find a point $F'$ on the same side of $\overleftrightarrow{DE}$ as $F$ so that $\triangle DEF'$ *is* congruent to $\triangle ABC$ (because the segments $AB$ and $DE$ are congruent).
- Note that points $F$ and $F'$ cannot be the same, since triangle $\triangle DEF$ is not congruent to $\triangle ABC$ and thus not congruent to $\triangle DEF'$.
- Now $\angle DEF' \cong \angle ABC$ and $\angle EDF' \cong \angle BAC$ since these are corresponding angles on congruent triangles.
- By transitivity of congruence, $\angle DEF' \cong \angle DEF$ and $\angle EDF' \cong \angle EDF$.
- By Fact 4.7 **(i)** there can be only one ray with endpoint $D$ on that side of $\overleftrightarrow{DE}$ which forms an angle of measure $m\angle EDF$ with ray $\overrightarrow{DE}$. So $\overrightarrow{DF} = \overrightarrow{DF'}$.
- Similarly, rays $\overrightarrow{EF}$ and $\overrightarrow{EF'}$ are equal.
- So, $D$, $F$, and $F'$ are collinear, as are $E$, $F$, and $F'$.
- But this means lines $\overleftrightarrow{DF}$ and $\overleftrightarrow{EF}$ share both points $F$ and $F'$. Since we have already noted that points $F$ and $F'$ are distinct, this gives us a contradiction to Axiom L1.

- This contradiction must have arisen from our assumption that $\triangle ABC$ and $\triangle DEF$ are not congruent. So, these triangles must in fact be congruent. $\square$

It isn't too difficult to construct a proof of the SAA criterion in the same spirit as the above proof. We leave it as an exercise to do so.

**THEOREM 4.17.** *Suppose that $ABC$ and $DEF$ are triangles and that the following congruences are known: $AB \cong DE$, $\angle B \cong \angle E$, and $\angle C \cong \angle F$. Then $\triangle ABC \cong \triangle DEF$.*

As mentioned above, we will delay the proof of the SSS criterion until later, as the results in the next section are needed for its proof. However, the congruence criteria we have so far established are enough to prove another basic toolbox theorem: the Isosceles Triangle Theorem.

**THEOREM 4.18.** *The triangle $ABC$ is isosceles with top vertex $C$ if and only if the angles $\angle A$ and $\angle B$ are congruent.*

*Proof:* As with all "if and only if" theorems, there are two parts to the proof. We first assume that $\angle A \cong \angle B$ and prove that $AC \cong BC$.

- First, note the following congruences (the first is our assumption, and the last two are trivial).

$$\angle ABC \cong \angle BAC$$
$$\angle ACB \cong \angle BCA$$
$$AB \cong BA$$

- By the SAA criterion we conclude that the triangles $\triangle ABC$ and $\triangle BAC$ are congruent. (Note the order of the vertices! These triangles are, of course, one-and-the-same as sets of points – but the congruence of these different *orders* of the vertices is of the essence here.)

- From this congruence we have $BC \cong AC$.

To complete the proof, we need to show that if $ABC$ is isosceles with top vertex $C$ then the angles $\angle A$ and $\angle B$ are congruent. So, assume $ABC$ is isosceles with top vertex $C$.

- We have the following congruences (the first and second follow from the definition of an isosceles triangle while the third is trivial):

$$AC \cong BC$$
$$BC \cong AC$$
$$\angle ACB \cong \angle BCA$$

- From Axiom SAS we conclude that $\triangle ACB \cong \triangle BCA$.
- Then $\angle ABC$ and $\angle BAC$ are corresponding angles in these two congruent triangles and are therefore congruent. □

We leave the proof of the following corollary as an exercise.

**Corollary 4.19.** *Let $ABC$ be an isosceles triangle with top vertex $C$, and let $D$ be the midpoint of the base $AB$. Then the line $\overleftrightarrow{CD}$ is the perpendicular bisector of the base $AB$.*

## Exercises

**4.15**. Prove the SAA congruence criterion (Theorem 4.17).

**4.16**. Prove Corollary 4.19.

**4.17**. Let $\triangle ABC$ be isosceles with top vertex $C$ and suppose that $D$ and $E$ are points on $AB$ such that $|AE| = |BD|$. Prove that $|CD| = |CE|$.

**4.18**. Suppose $A$, $B$, $C$, $D$, and $E$ are points not all on one line such that $B$ is the midpoint of $AC$, $D * B * E$, and $\angle DAB \cong \angle ECB$. Prove that $|AD| = |CE|$.

**4.19**. Prove that any triangle $ABC$ with $\angle A \cong \angle B \cong \angle C$ is equilateral.

**4.20**. The Isosceles Triangle Theorem (together with the Corresponding Angles Theorem) gives us what we need to improve on the statement of Axiom C2 by proving that a line and a circle can never share more than two points. Construct such a proof.

# E. Perpendiculars

This brief section will consider the topic of perpendicular lines. The main objective will be to prove the Perpendicular Bisector Theorem (stating that the perpendicular bisector of a segment is the set of points equidistant from its endpoints). But first, we use our just-proved results on isosceles triangles to verify two elementary facts about perpendiculars.

**FACT 4.20.** *Let $\Lambda$ be a line and let $P$ be any point. Then there is a unique line through $P$ that is perpendicular to $\Lambda$.*

*Proof:* If $P$ is a point on $\Lambda$ then this follows immediately from Fact 4.7 **(i)**. So, we may assume that $P$ is not on $\Lambda$. We will show that there is at least one line through $P$ perpendicular to $\Lambda$, leaving the proof that there cannot be more than one such line as an exercise.

- We first show that we can find points on $\Lambda$ equidistant from $P$.
  - Choose any point $A$ on $\Lambda$ and choose a number $r$ greater than the distance $|PA|$.
  - Let $\Gamma$ be the circle with center at $P$ and with radius $r$.
  - By definition, $A$ is inside $\Gamma$. So by Axiom C2, $\Gamma$ meets $\Lambda$ in at least two points.
  - If $B$ and $C$ are two such points then $|PB| = |PC|$.
- The triangle $PBC$ is now isosceles with top vertex $P$.
- So, if $D$ is the midpoint of $BC$, the line $\overleftrightarrow{PD}$ is perpendicular to $\overleftrightarrow{BC} = \Lambda$ by Corollary 4.19.
- This shows that there is at least one line through $P$ perpendicular to $\Lambda$. We leave the proof that there cannot be more than one such line as Exercise 4.21. $\square$

**FACT 4.21.** *Let $\Lambda$ be a line and let $P$ be a point not on $\Lambda$. Let $A$ be the (unique by Fact 4.20 above) point of $\Lambda$ such that $PA \perp \Lambda$. Then $|PQ| > |PA|$ for all points $Q \neq A$ on $\Lambda$.*

This last fact, the proof of which is outlined in Exercise 4.22, justifies the following definition.

**Definition.** Let $\Lambda$ be a line and $P$ a point not on $\Lambda$. The *distance* from $P$ to $\Lambda$ is $|PQ|$ where $\overleftrightarrow{PQ}$ is perpendicular to $\Lambda$ and $Q$ is on $\Lambda$.

We can now enlarge our basic toolbox by proving the Perpendicular Bisector Theorem.

**Theorem 4.22.** *The lengths $|CA|$ and $|CB|$ are equal if and only if $C$ is on the perpendicular bisector of $AB$.*

*Proof:* There are two parts to this proof. We first prove the "only if" part; that is, we prove that if $|CA| = |CB|$ then $C$ is on the perpendicular bisector of $AB$. So, assume that $|CA| = |CB|$. Then:

- $\triangle ABC$ is isosceles with top vertex $C$.
- So by Corollary 4.19, the top vertex $C$ lies on the perpendicular bisector of the base $AB$.

We now prove the "if" part; that is, we prove that if $C$ is on the perpendicular bisector of $AB$ then $|CA| = |CB|$. Assume, then, that $C$ is on the perpendicular bisector of $AB$.

- Let $D$ be the midpoint of $AB$.
- If $C = D$ then we are done since by definition of midpoint, $C$ would be equidistant from $A$ and $B$. So, we may assume that $C \neq D$.
- Then line $\overleftrightarrow{CD}$ is the perpendicular bisector of $AB$, so $\angle CDA$ and $\angle CDB$ are both right angles by definition.
- Also, $DA \cong DB$ by definition of midpoint.
- We now have $CD \cong CD$, $DA \cong DB$, and $\angle CDA \cong \angle CDB$.
- By Axiom SAS we conclude $\triangle CDA \cong \triangle CDB$.
- This means that $CA \cong CB$ since these are corresponding sides in congruent triangles.

- So, by definition of congruence for segments, $|CA| = |CB|$. This completes the proof. $\square$

## Exercises

**4.21**. Complete the proof of Fact 4.20 by showing that there cannot be more than one line through $P$ perpendicular to $\Lambda$.

**4.22**. Prove Fact 4.21 by filling in the details of the following argument.

- Assume to reach a contradiction that $B$ is a point of $\Lambda$ (other than $A$) so that $|PB| \leq |PA|$.
- Then there is a point $C \neq A$ on $\Lambda$ with $|PA| = |PC|$.
- Consider the triangle $PAC$ and conclude that $m\angle PAC = m\angle PCA = 90$.
- What is the contradiction? (Note: it is *not* that the angle measure sum in triangle $PAC$ exceeds 180, for we have not yet proved the 180° Sum Theorem, and won't prove it until Chapter 5!)

**4.23**. In Exercise 4.20 we sharpened the statement of Axiom C2. The Perpendicular Bisector Theorem allows us to similarly sharpen the statement of Axiom C3. Prove that if $A$ and $B$ are distinct points of intersection of two circles, then the centers of these circles are on the perpendicular bisector of $AB$. Use this to conclude that no two circles can share three distinct points in common.

# F. Triangles, Part II

In this section we will complete our brief treatment of neutral geometry by proving two more basic facts concerning triangles. First we will complete our collection of triangle congruence criteria by adding SSS. The proof is similar to that of of the ASA criterion (Theorem 4.16), but relies on the Perpendicular Bisector Theorem to reach the contradiction.

THEOREM **4.23.** **(The side-side-side congruence criterion for triangles)** *Suppose that all three pairs of corresponding sides in the triangles $\triangle ABC$ and $\triangle DEF$ are congruent. Then $\triangle ABC \cong \triangle DEF$.*

*Proof:* Assume (to reach a contradiction) that triangles $\triangle ABC$ and $\triangle DEF$ are *not* congruent.

- By Fact 4.9 we can find a point $F'$ on the same side of $\overleftrightarrow{DE}$ as $F$ so that $\triangle DEF'$ *is* congruent to $\triangle ABC$ (because the segments $AB$ and $DE$ are congruent).
- Note that points $F$ and $F'$ cannot be the same, since triangle $\triangle DEF$ is not congruent to $\triangle ABC$ and thus not congruent to $\triangle DEF'$.
- Then $DF' \cong AC$ since these are corresponding sides in congruent triangles.
- Since $DF \cong AC$, we have $DF' \cong DF$ by transitivity.
- But then $|DF'| = |DF|$, so by Theorem 4.22, $D$ is on the perpendicular bisector of $FF'$. perpendicular bisector of $FF'$.
- A similar sequence of steps shows that $|EF'| = |EF|$, so that $E$ is also on the perpendicular bisector of $FF'$.
- But then the line $\overleftrightarrow{DE}$ must *be* the perpendicular bisector of $FF'$.
- This is a contradiction: the line $\overleftrightarrow{DE}$ cannot intersect $FF'$ because $F'$ was chosen to be on the same side of $\overleftrightarrow{DE}$ as was $F$.
- This contradiction must have arisen from our assumption that $\triangle ABC$ and $\triangle DEF$ are not congruent. So, these triangles must in fact be congruent. $\square$

We should note that there are some very important neutral geometry theorems that are not included in this chapter, mostly because we do not need them for what we will do in Chapter 5. The most prominent example is the *Saccheri-Legendre Theorem*, which states that (in neutral geometry) the sum of the degree measures in a triangle will be no greater than 180. In Euclidean geometry we know that sum to be *exactly* 180, and remarkably, we will not need the Saccheri-Legendre Theorem when we prove that stronger result next chapter. However, the Saccheri-Legendre Theorem *will* be important to our work in hyperbolic geometry, and so we will prove it in Chapter 6.

But there is yet one fact from neutral geometry that will be needed in both Chapters 5 and 6, and we will close this chapter by proving it. You are no-doubt familiar with the fact that in any triangle it is possible to "drop a perpendicular" from one of the vertices to the opposite side (see Figure 4.14). To prove this we will need a lemma that amounts to a very weak version of the Saccheri-Legendre Theorem.

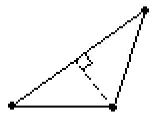


Figure 4.14:

**LEMMA 4.24.** *Every triangle has at least two acute angles.*

*Proof:* Let $ABC$ be a triangle and assume that $m\angle C \geq 90$. We will prove that $\angle A$ is acute – a similar argument would show that $\angle B$ is also acute. If we move a point $D$ from $A$ to $B$ along segment $AB$, the measure of the angle $\angle DCA$ will vary continuously from 0 to $m\angle C \geq 90$ by Axiom C1. So (by the Intermediate Value Theorem) there must be a point $D$ on $AB$ such that $m\angle DCA = 90$. Let $\Sigma$ be the side of $\overleftrightarrow{CD}$ containing $A$ and let $\Sigma'$ be the other side of $\overleftrightarrow{CD}$ (refer to Figure 4.15).

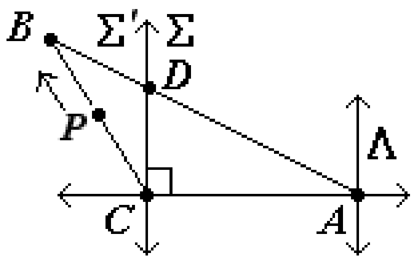



Figure 4.15:

- Note that point $B$ is in $\overline{\Sigma'}$ since segment $AB$ meets line $\overleftrightarrow{CD}$ at $D$. So by Fact 4.6 we see that $BC$ lies entirely in $\overline{\Sigma'}$ and so contains no point of $\Sigma$.
- Let $\Lambda$ be the line through $A$ perpendicular to $\overleftrightarrow{AC}$. Then $\Lambda$ contains all points $P$ with $m\angle PAC = 90$.
- By Corollary 4.13 we see that $\Lambda \parallel \overleftrightarrow{CD}$ (since both are perpendicular to $\overleftrightarrow{AC}$).
- So, $\Lambda$ must lie entirely in $\Sigma$ and so must miss segment $BC$.
- Thus $BC$ contains no point $P$ with $m\angle CAP = 90$. But again using Axiom C1, the value of $m\angle CAP$ varies continuously as $P$ moves from $C$ to $B$ along segment $CB$. We know this value starts at zero (when $P = C$) and is never equal to 90, so it must remain less than 90 at all times. In particular, $m\angle CAB < 90$. □

FACT **4.25.** *If $ABC$ is any triangle and both $\angle A$ and $\angle B$ are acute (see Lemma 4.24) then there is a point $D$ on $AB$ so that $CD \perp AB$.*

*Proof:* Let $r(t) = R_t$ be a ruler function for $\overleftrightarrow{AB}$ so that $A = R_0$ and $B = R_{|AB|}$ (see Axiom L2), and let $E = R_{|AB|+1}$ (so that $A * B * E$). Consider the function $\theta(t) = m\angle CR_tE$ (continuous by Axiom C1 – see Figure 4.16).

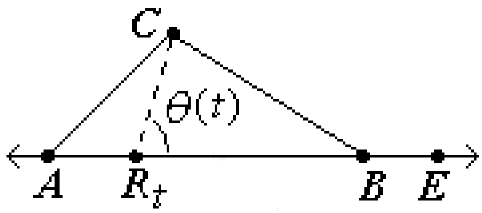



Figure 4.16:

- We have $\theta(0) = m\angle A < 90$.
- But also, $\theta(|AB|) = m\angle CBD = 180 - m\angle B > 90$.
- So (by the Intermediate Value Theorem) there must be a number $t_0$ with $0 < t_0 < |AB|$ so that $\theta(t_0) = 90$.
- Setting $D = R_{t_0}$ we see that $D$ is a point of $AB$ and that $m\angle CDB = 90$, so $CD \perp AB$. $\square$

## Exercises

**4.24**. True or False? (Questions for discussion)

**(a)** Only half of the isosceles triangle theorem is true in neutral geometry.

**(b)** In neutral geometry, if two lines are each perpendicular to a third line then they are parallel to each other.

**(c)** In neutral geometry, if $\Lambda_1 \parallel \Lambda_2$ and $\Lambda_2 \parallel \Lambda_3$, then $\Lambda_1$ must be parallel to $\Lambda_3$.

**(d)** In neutral geometry the alternate interior angles in a transversal of two parallel lines cannot be congruent.

**(e)** In neutral geometry the angle measure sum for a triangle cannot be equal to $180°$.

**(f)** In hyperbolic geometry there may be more than one line through a point $P$ perpendicular to a given line $\Lambda$.

**(g)** All of the triangle congruence criteria from Euclidean geometry are in fact true in neutral geometry.

**4.25**. Suppose in triangle $\triangle ABC$ that $\angle A$ and $\angle B$ are both acute and $|CA| < |CB|$. Prove that there is a point $D$ on $AB$ so that $\triangle ADC$ is isosceles with top vertex $C$.

**4.26**. In general there is no "SSA" congruence criterion. However, if $ABC$ and $DEF$ are right triangles with $\angle C$ and $\angle F$ right angles, and if $AB \cong DE$ and $AC \cong DF$, then $\triangle ABC \cong \triangle DEF$. Give a neutral geometry proof of this.

# Chapter 5

# Euclidean Geometry

We set out in Chapter 4 with a goal of proving from axioms all of the basic toolbox theorems of Euclidean geometry (as set forth in Chapter 1). We found that several of these could be proved in neutral geometry, but a glance back at Section 1B tells us that some important theorems remain to be proved. These remaining toolbox theorems belong strictly to Euclidean geometry – their proofs require the use of a parallel postulate. The situation is summarized in the following table.

**Basic Toolbox Theorems**

| **Neutral Geometry** | **Requires Parallel Postulate** |
|---|---|
| Alternate Interior Angles Theorem<br>(a.i.a. congruent $\Longrightarrow$ parallel lines) | Alternate Interior Angles Theorem<br>(parallel lines $\Longrightarrow$ a.i.a. congruent) |
| Corresponding Angles Theorem<br>(c.a. congruent $\Longrightarrow$ parallel lines) | Corresponding Angles Theorem<br>(parallel lines $\Longrightarrow$ c.a. congruent) |
| Vertical Angles Theorem | $180°$ Sum Theorem |
| Triangle Congruence Criteria | Parallelogram Theorem |
| Isosceles Triangle Theorem | Inscribed Angle Theorem |
| Perpendicular Bisector Theorem | Area Formulae |
| | Pythagorean Theorem |
| | Similar Triangles Theorem |

So now it is time to supplement our dozen axioms of neutral geometry by adding the Euclidean Parallel Axiom. Instead of the complicated wording of Euclid's 5th postulate, we will use the equivalent and simply-worded statement of Playfair's Postulate (see p.118).

**Euclidean Parallel Axiom.** *If $P$ is a point not on a line $\Lambda$ then there is exactly one line through $P$ that is parallel to $\Lambda$.*

With this axiom in hand we are ready to begin proving the theorems in the right-hand column of the above table.

- In Section A we'll prove the remaining parts of the Alternate Interior Angle and Corresponding Angle Theorems, as well as the 180° Sum Theorem.
- In Section B we'll investigate some basic constructions using the Euclidean Parallel Axiom. We'll show that three noncollinear points determine a unique circle, and we'll show how to construct rectangles. The discussion of quadrilaterals makes this a natural place to include the elementary Parallelogram Theorem.
- The remaining theorems in the list require more difficult proofs. The first of them, the Inscribed Angle Theorem, will be given in Section C.
- In Section D we add one more axiom – the Area Axiom – and use it to derive the area formulae for parallelograms and triangles.
- Area proves crucial to our proof in Section E of the most famous of all geometry theorems, the Pythagorean Theorem.
- Finally, in Section F we use the Pythagorean Theorem to establish the Similar Triangles Theorem.

That is our battle plan. So now, off to work!

## A. Basic Consequences of the Parallel Axiom

This section is devoted to the most basic consequences of the Euclidean Parallel Axiom. We begin with the observation that another way to word the axiom would be to say that the words "*is parallel to*" describes a *transitive* relation on lines.

**Fact 5.1.** *If $\Lambda_1 \parallel \Lambda_2$ and $\Lambda_2 \parallel \Lambda_3$ then $\Lambda_1 \parallel \Lambda_3$.*

*Proof:* The proof is by contradiction. Assume that $\Lambda_1$ is not parallel to $\Lambda_3$. Then they meet at some point $P$. The contradiction is immediate – there are two lines ($\Lambda_1$ and $\Lambda_3$) through $P$ both of which are parallel to $\Lambda_2$, contrary to the Euclidean Parallel Axiom. □

You might recall that the neutral geometry half of the Alternate Interior Angles Theorem (Theorem 4.12) required a somewhat tricky proof by contradiction. The proof of the "strictly Euclidean half" is easier, though also indirect. Note that we need to use the already-proved neutral half of the theorem at a crucial spot.

THEOREM **5.2.** *Suppose that line $\Lambda_0$ transverses lines $\Lambda_1$ and $\Lambda_2$, and that $\Lambda_1$ and $\Lambda_2$ are parallel. Then pairs of alternate interior angles of this transversal are congruent.*

*Proof:* Let lines $\Lambda_0$, $\Lambda_1$, and $\Lambda_2$ be as described in the statement of the theorem, and let points $P$, $Q$, $A$, and $B$ be as depicted in Figure 5.1. Assume to reach a contradiction that $\angle BPQ$ is *not* congruent to $\angle AQP$.

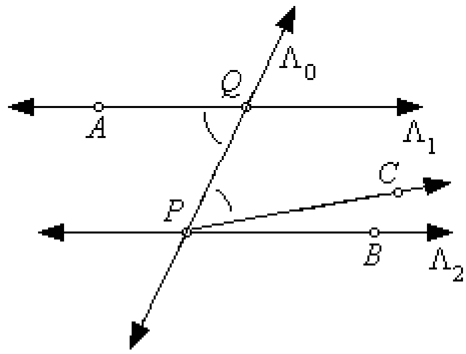



Figure 5.1:

- By Fact 4.7 **(i)** there is a ray $\overrightarrow{PC}$ on the same side of line $\overleftrightarrow{PQ}$ as point $B$ such that $m\angle CPQ = m\angle AQP$.
- Then $\angle CPQ \cong \angle AQP$ by the definition of congruence for angles.
- By Theorem 4.12 the lines $\overleftrightarrow{PC}$ and $\overleftrightarrow{QA} = \Lambda_1$ are parallel.
- Since angles $\angle BPQ$ and $\angle CPQ$ cannot be congruent, the lines $\overleftrightarrow{PC}$ and $\overleftrightarrow{PB} = \Lambda_2$ are distinct.
- This is a contradiction to the Euclidean Parallel Axiom – we cannot have two distinct lines ($\Lambda_2$ and $\overleftrightarrow{PC}$) parallel to $\Lambda_1$ and both passing through point $P$. □

Similar to Corollary 4.13 following the neutral half of the theorem, Theorem 5.2 above has the following corollary. We leave the easy proof as an exercise.

**COROLLARY 5.3.** *If $\Lambda_1 \parallel \Lambda_2$ and $\Lambda_3 \perp \Lambda_1$ then $\Lambda_3$ is also perpendicular to $\Lambda_2$.*

We can also easily extend Theorem 5.2 to a proof of the "strictly Euclidean half" of the Corresponding Angles Theorem. Again, we leave the proof to the reader.

**THEOREM 5.4.** *Suppose that line $\Lambda_0$ transverses lines $\Lambda_1$ and $\Lambda_2$, and that $\Lambda_1$ and $\Lambda_2$ are parallel. Then pairs of corresponding angles of this transversal are congruent.*

We conclude this section by proving one of the most useful facts in all of Euclidean geometry – the 180° Sum Theorem. It's almost amazing how easy its proof becomes once we have Theorem 5.2 at our disposal.

**THEOREM 5.5.** *The measures of the angles of any triangle sum to 180.*

*Proof:* Let $ABC$ be a triangle. We need to prove that $m\angle A + m\angle B + m\angle C = 180$.

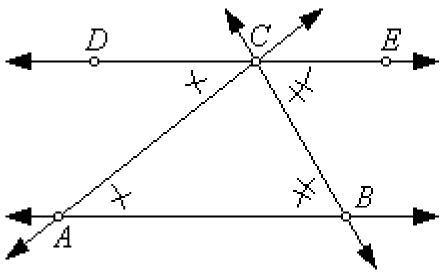



Figure 5.2:

- The point $C$ is not on the line $\overleftrightarrow{AB}$, so by Corollary 4.15 (we could use the Euclidean Parallel Axiom here, but it isn't necessary!) we may construct a line through $C$ that is parallel to $\overleftrightarrow{AB}$.
- Let this line be $\overleftrightarrow{DE}$ with $D * C * E$ as depicted in Figure 5.2.
- Since lines $\overleftrightarrow{AB}$ and $\overleftrightarrow{DE}$ are parallel and are transversed by $\overleftrightarrow{AC}$, Theorem 5.2 tells us that $\angle ACD \cong \angle CAB$.
- Similarly, considering the transversal of $\overleftrightarrow{BC}$ over $\overleftrightarrow{AB}$ and $\overleftrightarrow{DE}$ we conclude that $\angle ECB \cong \angle ABC$.
- Thus, $m\angle ACD = m\angle CAB$ and $m\angle ECB = m\angle ABC$.
- But a consequence of Axiom AC is that $m\angle ECB + m\angle BCA + m\angle ACD = m\angle DCE = 180$.
- Substituting, we get $m\angle ABC + m\angle BCA + m\angle CAB = 180$. □

### Exercises

**5.1**. Prove Corollary 5.3.

**5.2**. Prove Theorem 5.4.

**5.3**. Analyze the proof of Theorem 5.5 to find where the Euclidean Parallel Axiom is needed.

## B. Some Euclidean Constructions

You might recall that Euclid's five postulates in the *Elements* were motivated by straightedge and compass constructions. Let's pause now to consider our axiom system in relation to the following two elementary constructions:

**1.** Given three noncollinear points $A$, $B$, and $C$, construct a circle through all three. That is, find a point $P$ so that $|PA| = |PB| = |PC|$.

**2.** Given a segment $AB$, construct $C$ and $D$ so that $ABCD$ is a rectangle (or a square).

These simple tasks serve well to highlight the difference between neutral geometry and Euclidean geometry. For remarkably, neither of these have solutions that can be proved to work in neutral geometry!

Consider the first task. By the Perpendicular Bisector Theorem (which we proved in neutral geometry as Theorem 4.22) the set of points equidistant from $A$ and $B$ is the perpendicular bisector of segment $AB$, and the set of points equidistant from $B$ and $C$ is the perpendicular bisector of segment $BC$. All we need to do, then, is find the point of intersection of these two lines – this point should be equidistant from all three of $A$, $B$, and $C$. However, how do we know that the two perpendicular bisectors will intersect? This, it turns out, can only be demonstrated with the Euclidean Parallel Axiom. The following lemma provides the key.

LEMMA **5.6.** *Suppose* $\Lambda_1$, $\Lambda_2$, $\Lambda_3$, *and* $\Lambda_4$ *are lines with* $\Lambda_1 \perp \Lambda_3$ *and* $\Lambda_2 \perp \Lambda_4$. *Then* $\Lambda_1 \parallel \Lambda_2$ *if and only if* $\Lambda_3 \parallel \Lambda_4$.

*Proof.* Assume that $\Lambda_1$ and $\Lambda_2$ are parallel. We will prove that $\Lambda_3$ and $\Lambda_4$ are also parallel.

- First, since $\Lambda_3$ is perpendicular to $\Lambda_1$, it is also (by Corollary 5.3) perpendicular to $\Lambda_2$.
- So, $\Lambda_2$ transverses the lines $\Lambda_3$ and $\Lambda_4$ with congruent alternate interior angles. (All interior angles of the transversal are right angles!)
- By Theorem4.12 lines $\Lambda_3$ and $\Lambda_4$ must be parallel.

The proof of the converse is, of course, exactly analogous. □

**FACT 5.7.** *Through any three noncollinear points there is exactly one circle.*

*Proof:* Let $A$, $B$, and $C$ be any three noncollinear points. We will prove that there is at least one circle passing through $A$, $B$, and $C$, leaving the proof that there cannot be more than one such circle as an exercise.

- The lines $\overleftrightarrow{AB}$ and $\overleftrightarrow{BC}$ are distinct (since $A$, $B$, and $C$ are noncollinear) and clearly not parallel (since $B$ lies on both lines).
- Let line $\Lambda_1$ be the perpendicular bisector of $AB$ and line $\Lambda_2$ be the perpendicular bisector of $BC$. By Lemma 5.6 the lines $\Lambda_1$ and $\Lambda_2$ cannot be parallel.
- Let $P$ be the point at which these two lines meet.
- Since $P$ is on the perpendicular bisector of $AB$, Theorem 4.22 tells us that $|PA| = |PB|$.
- Similarly, $|PB| = |PC|$ since $P$ is on the perpendicular bisector of $BC$.
- So, the circle with center $P$ and radius $|PA|$ passes through all three points $A$, $B$, and $C$. □

Let's now turn our attention to the second construction listed at the beginning of this section – that of constructing a rectangle or square on a given segment $AB$. You might think of two possible approaches to this construction, so let's consider each one.

**Construction A:** (See the left side of Figure 5.3.) Given segment $AB$:

- Construct the lines $\Lambda_A$ through $A$ and $\Lambda_B$ through $B$, each perpendicular to $\overleftrightarrow{AB}$ (see Fact 4.20).
- On one side of $\overleftrightarrow{AB}$ find points $C$ on $\Lambda_B$ and $D$ on $\Lambda_A$ so that $|BC| = |AD|$ ($= |AB|$ for a square).

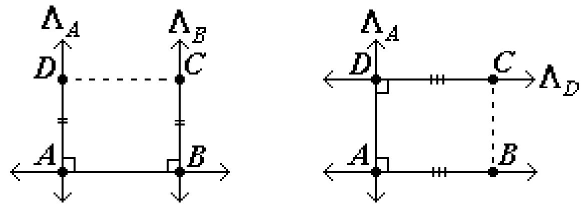



Figure 5.3: Construction A (left) and Construction B (right)

**Construction B:** (See the right side of Figure 5.3.) Given segment $AB$:

- Construct the line $\Lambda_A$ through $A$ and perpendicular to $\overleftrightarrow{AB}$ (see Fact 4.20).
- Let $D$ be a point other than $A$ on this line. (For a square, choose $D$ so that $|AD| = |AB|$.)
- Construct a line $\Lambda_D$ through $D$ that is perpendicular to $\Lambda_A$ and thus parallel to $\overleftrightarrow{AB}$. (The Euclidean Parallel Axiom guarantees that there is a unique such line. But even in neutral geometry we know that at least one such line exists – see Corollary 4.13.)
- Let $C$ be the point on $\Lambda_D$ on the same side of $\Lambda_A$ as $B$ and such that $|DC| = |AB|$.

As the figure suggests, both constructions produce a quadrilateral with two adjacent right angles and a pair of congruent opposite sides, one incident to each of the two known right angles. Such a figure is called a *Saccheri quadrilateral* and will play an important role in Chapter 6. It isn't hard to prove that a Saccheri quadrilateral is a rectangle if we allow ourselves use of the Euclidean Parallel Axiom (see part **(c)** of Exercise 5.7). However, one cannot prove this in neutral geometry. In fact, we'll see in Chapter 6 that while Saccheri quadrilaterals exist in hyperbolic geometry (for the steps in either of the above constructions can certainly be carried out using only neutral geometry maneuvers), there are *no rectangles* in hyperbolic geometry!

Suppose $ABCD$ is a Saccheri quadrilateral as in Figure 5.4, and let $E$ and $F$ be the midpoints of the sides $AB$ and $CD$. It isn't too difficult to show (using only neutral geometry – see Exercise 5.6) that $EF$ is perpendicular to both $AB$ and $CD$. So, we might expect that $|EF|$ would be the same as $|AD|$ and $|BC|$. (After all, $\overleftrightarrow{FC}$ is then parallel to $\overleftrightarrow{AB}$ by Corollary 4.13, and we expect the distance between two parallel lines to be constant at all locations along the lines.) However, this cannot be proved in neutral geometry.[1] But with the Euclidean Parallel Axiom we have the following fact. The proof is left up to you.

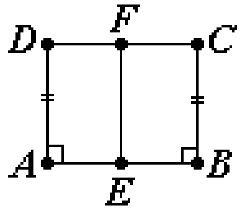



Figure 5.4:

**Fact 5.8.** *Suppose that lines $\Lambda_1$ and $\Lambda_2$ are parallel. Let $A$ and $E$ be points on $\Lambda_1$ and let $D$ and $F$ be points on $\Lambda_2$ such that $AD$ and $EF$ are both perpendicular to $\Lambda_1$ (see Figure 5.5). Then $|AD| = |EF|$.*

We previously defined the distance from a point to a line in neutral geometry (see p.151) as being measured along a perpendicular segment from the point to the line. The above fact justifies the following definition of *distance between two parallel lines* in Euclidean geometry.

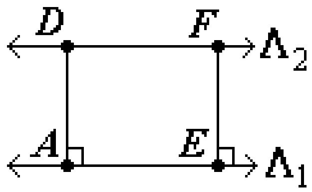



Figure 5.5:

**Definition.** Let $\Lambda_1$ and $\Lambda_2$ be parallel lines. The *distance* between $\Lambda_1$ and $\Lambda_2$ is the distance between $\Lambda_1$ and any point of $\Lambda_2$.

Since this section has led us into a discussion of quadrilaterals, we shall conclude it with the contents of the single toolbox theorem from Section 1B dealing with quadrilaterals. We divide it here into a couple of parts, each one being easily proved. If you have not as yet worked out these proofs for yourself, now is the time to do so. (The definitions of quadrilateral types are as set forth on p.18.)

**Theorem 5.9.** *Every rectangle is a parallelogram and every rhombus is a parallelogram.*

---

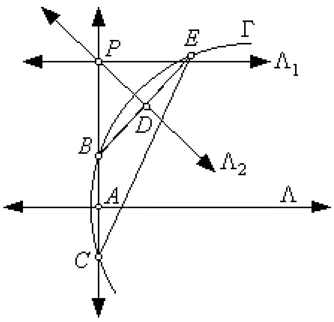
[1] In fact, like the existence of rectangles, the equidistance of parallel lines *fails to be true* in hyperbolic geometry!

**Theorem 5.10.** *If $ABCD$ is a parallelogram then $\angle A \cong \angle C$, $\angle B \cong \angle D$, $|AB| = |CD|$, and $|BC| = |DA|$.*

## Exercises

**5.4**. Complete the proof of Fact 5.7 by showing that any circle through $A$, $B$, and $C$ has the same center and radius as the circle constructed in the first part of the proof, and is thus that same circle.

**5.5**. Find the unjustified step in the following "proof" of the Euclidean Parallel Postulate.




- Let $\Lambda$ be a line and let $P$ be a point not on $\Lambda$.
- Let $A$ be the point on $\Lambda$ such that $PA \perp \Lambda$ (use Fact 4.20).
- Let $\Lambda_1$ be the line through $P$ that is perpendicular to $PA$ (and thus parallel to $\Lambda$ by Corollary 4.13).
- We will prove that any line through $P$ other than $\Lambda_1$ must meet $\Lambda$ (thus showing that there is only one line through $P$ parallel to $\Lambda$).
- So, let $\Lambda_2$ be any line through $P$ other than $\Lambda_1$ and $\overleftrightarrow{PA}$. We will show that $\Lambda_2$ must intersect $\Lambda$.
- Let $B$ be a point of $\overleftrightarrow{PA}$ between $P$ and $A$, and let $C$ be the unique point of $\overleftrightarrow{PA}$ such that $B * A * C$ and $|AB| = |AC|$ (Axiom L2).
- Let $D$ be the point of $\Lambda_2$ such that $BD \perp \Lambda_2$ (Fact 4.20 again).
- Let $E$ be the unique point of $\overleftrightarrow{BD}$ such that $B * D * E$ and $|BD| = |ED|$ (Axiom L2 again).

- Let $\Gamma$ be the circle through $B$, $C$, and $E$.
- Since $\Lambda$ is the perpendicular bisector of $BC$, the center of $\Gamma$ must be on $\Lambda$.
- Since $\Lambda_2$ is the perpendicular bisector of $BE$, the center of $\Gamma$ must be on $\Lambda_2$.
- So, $\Lambda$ and $\Lambda_2$ must meet at the center of $\Gamma$, and so cannot be parallel.

**5.6**. Let $ABCD$ be a Saccheri quadrilateral with $AD \perp AB$, $BC \perp AB$, and $AD \cong BC$. Let $E$ and $F$ be the midpoints of $AB$ and $CD$.

**(a)** Prove using only neutral geometry that $EF$ is perpendicular to both $AB$ and $CD$.

**(b)** Prove using only neutral geometry that $\angle C \cong \angle D$.

**5.7**. Prove (in Euclidean geometry) that a Saccheri quadrilateral is a rectangle. Where does your proof use the Euclidean Parallel Axiom?

**5.8**. Can you find a Saccheri quadrilateral in spherical geometry?

**5.9**. Prove Fact 5.8.

**5.10**. Prove Theorem 5.9.

**5.11**. Prove Theorem 5.10.

**5.12**. Show that in neutral geometry, any Saccheri quadrilateral that is also a rhombus must be a rectangle. What does this say about Saccheri quadrilaterals in hyperbolic geometry?

# C. Inscribed Angles

Our goal for this section is a single toolbox theorem – the Inscribed Angle Theorem. We'll warm-up, however, with a famous special case of that theorem – the Theorem of Thales (see p.21).

**Theorem 5.11.** *Let $\Gamma$ be a circle with center $C$ and let $A$ and $B$ be points on $\Gamma$ so that $C$ is on $AB$ ($AB$ is a* diameter *of $\Gamma$). Let $D$ be any other point of $\Gamma$. Then $\angle ADB$ is a right angle.*

*Proof:* Refer to Figure 5.6 in the following steps.

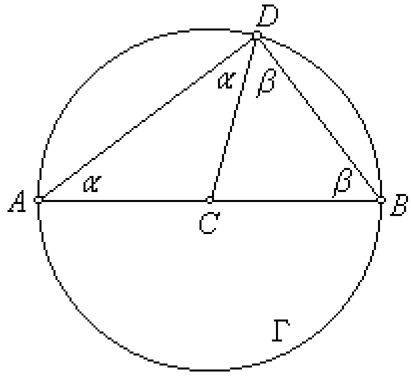



Figure 5.6:

- Since $C$ is the center of $\Gamma$ then $|CA| = |CB| = |CD|$. So, the triangles $ADC$ and $BDC$ are both isosceles with top vertex $C$.
- By the Isosceles Triangle Theorem (Theorem 4.18) $m\angle CAD = m\angle CDA = \alpha$ and $m\angle CBD = m\angle CDB = \beta$ (as indicated in Figure 5.6).
- So by the 180° Sum Theorem (Theorem 5.5) applied to triangle $ABD$ we have

$$\begin{aligned} m\angle A + m\angle B + m\angle D &= 180 \\ \alpha + \beta + (\alpha + \beta) &= 180 \\ 2\alpha + 2\beta &= 180 \\ \alpha + \beta &= 90 \end{aligned}$$

- So $m\angle ADB = \alpha + \beta = 90$, which proves the theorem. □

Note the role Figure 5.6 plays in the above proof: it helps to clarify and illustrate the main idea of the proof, and having it to refer to certainly lessens the number of necessary words! *But*, it does not by itself provide the justification for any step. Any assumptions we made about this diagram can be backed up with the axioms. (See, for example, Exercise 5.16). We should become increasingly comfortable with using diagrams like this in our proofs. We need only be sure that every claim we make about a figure can in fact be verified directly from the axioms (even if we do not directly give the verification, which may well be very lengthy).

**Theorem 5.12.** *Let $\Gamma$ be a circle with center $C$ and let $\overset{\frown}{AQB}$ be an arc on $\Gamma$ and $R$ a point of $\Gamma$ not on $\overset{\frown}{AQB}$. Then $m\overset{\frown}{AQB}$ is equal to twice the measure of the inscribed angle $\angle ARB$ for that arc.*

*Proof:* We may choose our protractor function $p(t) = P_t$ for $\Gamma$ so that $R = P_{-180} = P_{180}$. Then let $\alpha < \beta$ be numbers such that $A = P_\alpha$ and $B = P_\beta$. Note that by definition, $m\overset{\frown}{AQB} = \beta - \alpha$, so we must show that $m\angle ARB = \frac{1}{2}(\beta - \alpha)$. We may assume that $\beta > 0$, and we consider three cases for the value of $\alpha$, all illustrated in Figure 5.7.

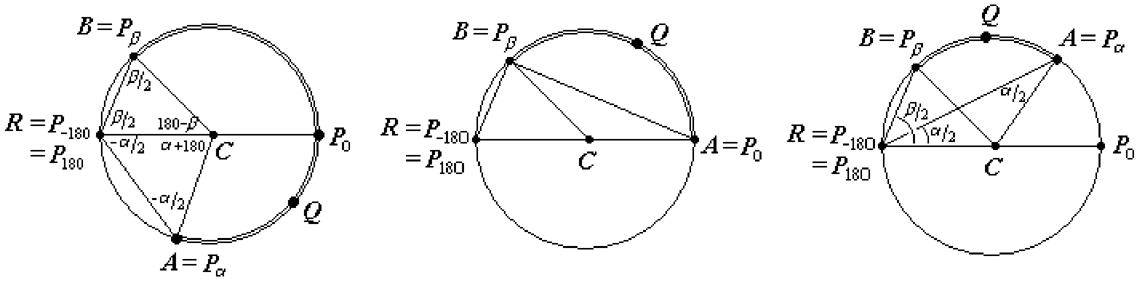



Figure 5.7:

**Case 1:** First suppose $\alpha < 0$, as in the leftmost part of the figure.

- Then by Axiom CA and our definition of angle measure[2], $m\angle RCA = \alpha - (-180) = \alpha + 180$.
- But triangle $RCA$ is isosceles (with top vertex $C$) because $RC$ and $AC$ are both radii of $\Gamma$.
- So by the Isosceles Triangle Theorem (Theorem 4.18) and the 180 Sum Theorem (Theorem 5.5) we see that $m\angle CAR = m\angle CRA = \frac{1}{2}(180 - m\angle RCA) = \frac{1}{2}(180 - (\alpha + 180)) = -\alpha/2$. (Note that this is actually a positive number, as an angle measure should be according to our definition!)
- Similar reasoning shows that $m\angle CBR = m\angle CRB = \frac{1}{2}(180 - m\angle RCB) = \frac{1}{2}(180 - (180 - \beta)) = \beta/2$.
- So[3] $m\angle ARB = m\angle ARC + m\angle CRB = -\alpha/2 + \beta/2 = \frac{1}{2}(\beta - \alpha)$, as desired.

[2] There's more to this than might meet the eye at first! To calculate $m\angle RCA$ from our definition we would adjust our protractor by rotating it by 180 (that is, $P'_t = P_{t+180}$ – the last statement in Axiom CA guarantees this relationship between the two protractor functions) so that $R = P'_0$ and $A = P'_{\alpha+180}$. Then, since $\alpha + 180 < 180$ we would have $m\angle RCA = \alpha + 180$.

[3] There's another fine point here that can be argued from the axioms if one is a stickler for such details. We need to know that the ray $\overrightarrow{RC} = \overrightarrow{RP_0}$ is interior to the angle $\angle ARB$. This can be verified to follow from our assumptions that $\alpha < 0$ and $\beta > 0$, but we have omitted the fine points here to keep the main argument as clear as possible.

**Case 2:** Suppose now that $\alpha = 0$ so that $A = P_0$ (so that $R * C * A$, as in the middle portion of Figure 5.7). We leave it as Exercise 5.13 to verify the conclusion in this case.

**Case 3:** Finally, suppose $\alpha > 0$ as depicted in the right portion of Figure 5.7.

- In this case $m\angle RCA = 180 - \alpha$ by definition.
- As in Case 1, triangle $RCA$ is isosceles, and it is easy to derive that $m\angle CAR = m\angle CRA = \alpha/2$.
- Also similar to Case 1, triangle $RCB$ is isosceles and $m\angle CBR = m\angle CRB = \beta/2$.
- Finally, then, $m\angle ARB = m\angle CRB - m\angle CRA = \beta/2 - \alpha/2 = \frac{1}{2}(\beta - \alpha)$. □

## Exercises

**5.13**. Complete Case 2 in the proof of Theorem 5.12.

**5.14**. Let $\Gamma$ be a circle with center $C$ and let $A$, $B$, and $P$ be three distinct points on $\Gamma$. Let $\Lambda$ be a line through $A$ with $CA \perp \Lambda$ ($\Lambda$ is tangent to $\Gamma$ at $A$) and let $D$ and $E$ be points on $\Lambda$ with $D * A * E$. Prove that $m\angle APB$ is equal to either $m\angle BAD$ or $m\angle BAE$.

**5.15**. If you didn't already work Exercise 2.38, do so now.

**5.16**. Why is it justified in the proof of Thales' Theorem (Theorem 5.11) to say $m\angle ADB = \alpha + \beta$? Can you show that $C$ is in the interior of angle $\angle ADB$?

# D. The Area Axiom

To the ancient Greek mathematicians, the Pythagorean Theorem was principally a statement about areas. It appears, with its converse, as the last two theorems in Book 1 of Euclid's *Elements*, where it is stated as:

> *In right-angled triangles the square on the side subtending the right angle is equal [in area] to the [sum of the areas of the] squares on the sides containing the right angle.*

In fact, the proof given by Euclid (essentially the same proof we will give in Section E) relies on the concept of area to carry the argument. In this section we will equip our axiom system for this proof (and for the subsequent discussion of ratios and similarity in Section F) by supplying the elements needed for discussion of area. We will need to augment our axiom system with a new axiom and some new facts that follow as consequences of that axiom. Our immediate goal will be to develop material sufficient to allow us to compute the areas of convex polygons. (Recall that we defined the interior of a convex polygon on p.132.) In particular, we will prove the area formulae for triangles and parallelograms from our basic toolbox of theorems in Section 1B.

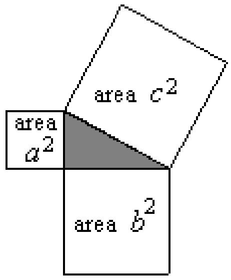



Figure 5.8: The Pythagorean Theorem as a statement about areas

From the Parallelogram Theorem we know that every rectangle and every rhombus is a parallelogram. Our first simple fact is to conclude that these are all convex polygons. Recall that a polygon is *convex* if for each line determined by two consecutive vertices, all of the other vertices lie on one side of that line.

**FACT 5.13.** *All triangles and all parallelograms are convex polygons.*

*Proof:* To prove that a triangle is convex there is really nothing to show: it is impossible for two vertices to be on opposite sides of a line determined by a side of the triangle since there is only one vertex not on that line.

Now let $ABCD$ be a parallelogram.

- The segment $CD$ cannot intersect the line $AB$ since by definition of a parallelogram $CD \parallel AB$.
- Consequently, $C$ and $D$ are on the same side of $\overleftrightarrow{AB}$.
- A similar argument works for the other sides of the 4-gon.

This shows that $ABCD$ is a convex 4-gon. $\square$

To complete the proof of the Pythagorean Theorem in Section E we need to be able to compute areas for the interiors of parallelograms and triangles. We will actually do considerably more by showing how the area of the interior of any convex polygon can be computed. Of course, so far we have not even stated what *area* is! To give us a starting point we append the following additional axiom to our axiom system.

**THE AREA AXIOM.** *To the interior of each convex polygon $P_1P_2\cdots P_n$ we may associate a number denoted by* area$(P_1P_2\cdots P_n)$ *and called the* area of $P_1P_2\cdots P_n$. *This association has the following properties.*

**(1)** *If $ABCD$ is a rectangle then* $\text{area}(ABCD) = |AB||BC|$.

**(2)** *If $\triangle ABC \cong \triangle DEF$ then* $\text{area}(ABC) = \text{area}(DEF)$.

**(3)** *For $3 \leq k \leq n-1$ we have* $\text{area}(P_1P_2\cdots P_n) = \text{area}(P_1P_2\cdots P_k) + \text{area}(P_kP_{k+1}\ldots P_nP_1)$. *Also, if $A$ is a point on side $P_kP_{k+1}$ ($2 \leq k \leq n-1$) then* $\text{area}(P_1P_2\cdots P_n) = \text{area}(P_1P_2\ldots P_kA) + \text{area}(AP_{k+1}P_{k+2}\ldots P_nP_1)$.

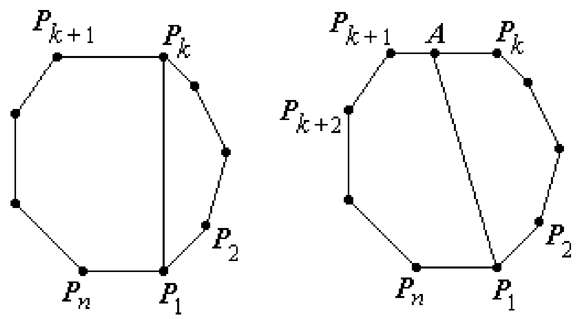



Figure 5.9: Part 3 of the Area Axiom

Though technically area is computed on the interior of a polygon, and not on the polygon itself, we will often say for brevity merely "the area of the polygon".

Consider for a moment what each part of the Area Axiom says:

**(1)** gives a "baseline" for evaluating areas of convex polygons. Any measure of area should coincide with "length-times-width" for rectangles.

**(2)** guarantees the intuitive notion that since congruent triangles "have the same size and shape" they should also have the same area.

**(3)** states that our concept of area behaves well with respect to "cutting" a convex polygon into two pieces (see Figure 5.9). This assumption (together with our assumption on the area of rectangles) will be used to derive the area of triangles, and from there to derive the area of any convex polygon.

As we proceed from here, keep in mind the role played by the Area Axiom. We have set the above three properties as defining characteristics for area, so that now the word "area" to us now designates a method of measurement obeying these rules. Using only these three rules (and the other axioms we have adopted) we will prove the basic facts one would expect concerning areas of convex polygons.

This process has the following consequence: *any method of computing areas that satisfies these three rules must also satisfy the theorems we will prove.* There are many ways to compute areas in the coordinate plane (for example, using integrals or using formulae involving coordinates of vertices), and it is entirely conceivable that different methods might give different valuations of area on some sets. However, as a consequence of what we will prove in this section, we can be assured that if two methods for computing area both satisfy the three rules in the Area Axiom, they must give the same values on every convex polygon.

**FACT 5.14.** *Let $ABC$ be a right triangle with right angle $\angle ABC$. Then* $\text{area}(ABC) = \frac{1}{2}|AB||BC|$.

*Proof:* Refer to Figure 5.10 in the following steps.

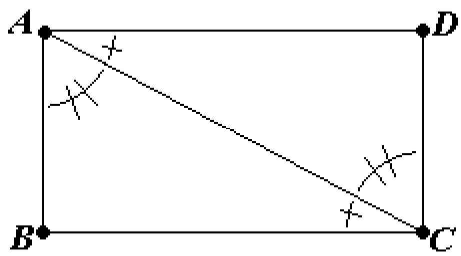



Figure 5.10:

- First, let $D$ be the point such that $ABCD$ is a rectangle (recall the constructions in Section B).
- Considering the transversal of $\overleftrightarrow{AC}$ over the two parallel lines $\overleftrightarrow{AB}$ and $\overleftrightarrow{CD}$ we see that $\angle BAC \cong \angle DCA$ by Theorem 5.2.
- For the same reason (considering the transversal of $\overleftrightarrow{AC}$ over $\overleftrightarrow{BC}$ and $\overleftrightarrow{AD}$) we have $\angle BCA \cong \angle DAC$.
- Trivially, we have $AC \cong CA$.

- By the last three steps and the ASA criterion (Theorem 4.16) we see that $\triangle ABC \cong \triangle CDA$.
- By the Area Axiom, then, we have $\text{area}(ABC) = \text{area}(CDA)$.
- So, $\text{area}(ABCD) = \text{area}(ABC) + \text{area}(CDA)$ by the Area Axiom.
- Combining the last two steps we have $\text{area}(ABCD) = 2\text{area}(ABC)$, so (using the Area Axiom to evaluate $\text{area}(ABCD)$) we conclude $\text{area}(ABC) = \frac{1}{2}|AB||BC|$. □

Our next task is to extend this to a formula for the area of any triangle.

**Definition.** We say that the *height* of triangle $ABC$ corresponding to the side $AB$ is the distance from $C$ to $\overleftrightarrow{AB}$.

**Theorem 5.15.** *Let $ABC$ be a triangle with height $h$ corresponding to the side $AB$. Then* $\text{area}(ABC) = \frac{1}{2}|AB|h$.

*Proof:* Let $D$ be the point on $\overleftrightarrow{AB}$ so that $CD \perp AB$. (Note that in this case $|CD| = h$.) There are two cases. We give the proof here for the case that $D$ is between $A$ and $B$, and leave the other case for Exercise 5.18.

- Since $D$ is between $A$ and $B$ we have $|AD| + |BD| = |AB|$.
- The two triangles $ADC$ and $BDC$ are both right triangles (with the right angle at vertex $D$), so by Fact 5.14 we have $\text{area}(ADC) = \frac{1}{2}|AD||CD|$ and $\text{area}(BDC) = \frac{1}{2}|BD||CD|$.
- By the Area Axiom, then, we have $\text{area}(ABC) = \text{area}(ADC) + \text{area}(BDC) = \frac{1}{2}|AD||CD| + \frac{1}{2}|BD||CD| = \frac{1}{2}(|AD| + |BD|)|CD| = \frac{1}{2}|AB|h$

□

With the area formula for a triangle in hand, it isn't difficult to derive the area formula for a parallelogram. We leave the proof of the following to the reader.

**Theorem 5.16.** *Let $ABCD$ be a parallelogram and let $h$ be the distance between the parallel lines $\overleftrightarrow{AB}$ and $\overleftrightarrow{CD}$. Then* $\text{area}(ABCD) = |AB|h$.

Having now derived the area formulae for some familiar convex polygons (though see Exercise 5.20 for one more) we will close this section with a formula expressing the area of any convex $n$-gon as a sum of areas of triangles. We leave to the reader the (relatively easy) task of using mathematical induction to prove the following.

THEOREM **5.17.** *Let $P_1P_2\cdots P_n$ be a convex $n$-gon. Then the area of its interior is equal to* $\text{area}(P_1P_2P_3) + \text{area}(P_1P_3P_4) + \text{area}(P_1P_4P_5) + \cdots + \text{area}(P_1P_{n-1}P_n)$.

### Exercises

**5.17**. Show that the points inside a circle form a convex set by expressing this set as the intersection of some collection of halfplanes.

**5.18**. Complete the proof of Theorem 5.15 in the case where $D$ is not between $A$ and $B$. (You may assume without loss of generality that $A * B * D$.)

**5.19**. Prove Theorem 5.16.

**5.20**. Find and prove a formula for the area of a trapezoid.

**5.21**. Use mathematical induction to prove Theorem 5.17.

**5.22**. Prove that if $ABCD$ is a convex 4-gon with perpendicular diagonals $AC$ and $BD$, then $\text{area}(ABCD) = \frac{1}{2}|AC||BD|$.

**5.23**. Let $ABCD$ be a rectangle with $P$ any point on $BC$ and $Q$ any point on $CD$. Prove that $|AB||BC| = 2\ \text{area}(AXY) + |BX||DY|$.

## E. The Pythagorean Theorem

This section's goal is straightforward: to prove the Pythagorean Theorem and its converse. It would be difficult to overstate the importance of this theorem to Euclidean geometry. As we saw in Chapter 2, the Pythagorean Theorem is a gatekeeper to many of geometry's most interesting results, the proofs of which require its use.

The proof we give here is essentially the one used by Euclid in the *Elements*. It is a deservedly famous proof, for it is extremely clever and elegant.

**THEOREM 5.18. (The Pythagorean Theorem)** *Let $ABC$ be a right triangle with $\angle C$ a right angle, then $a^2 + b^2 = c^2$.*

*Proof:* Let $ABC$ be as described. Begin by using the construction in Section B to construct points $D, E, F, G, H$, and $I$ such that

- $BCDE$ is a square with $D$ and $E$ on the side of $\overleftrightarrow{BC}$ not containing $A$.
- $ACFG$ is a square with $F$ and $G$ on the side of $\overleftrightarrow{AC}$ not containing $B$.
- $ABHI$ is a square with $H$ and $I$ on the side of $\overleftrightarrow{AB}$ not containing $C$.

Now let $K$ be the point on $AB$ so that $CK \perp AB$ (see Fact 4.25). The line $\overleftrightarrow{CK}$ (being perpendicular to $AB$) is parallel to both $AI$ and $BH$ (see Corollary 4.13). But $A$ and $B$ are on different sides of this line, and thus $I$ and $H$ are also on different sides. Thus $\overleftrightarrow{CK}$ must meet $HI$ at a point $J$. See Figure 5.11 for reference.

To avoid getting lost in the details without seeing the "big picture" of the proof, we will first outline the major steps. After this outline we will fill in the details for each of the major steps.

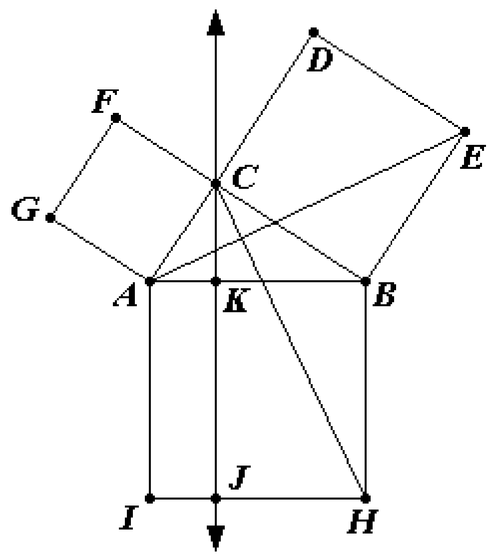



Figure 5.11: Configuration for the proof of the Pythagorean Theorem

**(1)** The area of the triangle $EBA$ is one-half the area of the square $BCDE$, or in other words, $\frac{1}{2}a^2$.

**(2)** The area of the triangle $CBH$ is one-half the area of the rectangle $BHJK$.

**(3)** The triangles $EBA$ and $CBH$ are congruent, and so have equal areas.

**(4)** So, the area of rectangle $BHJK$ is $a^2$.

**(5)** Similar steps show that the rectangle $JKAI$ has area $b^2$.

**(6)** But the sum of the areas of these two rectangles is the area of the square $ABHI$, or $c^2$.

**Step (1):**

- First, angles $\angle ACB$ and $\angle BCD$ are both right angles (because $ABC$ is a right triangle and $BCDE$ is a square), so $m\angle ACD = 180$ which means that the points $A$, $C$, and $D$ are collinear.

- So the height of $EBA$ corresponding to side $EB$ is $a$.

- By the area formula for triangles (Theorem 5.15) we have area$(EBA) = \frac{1}{2}|EB|a = \frac{1}{2}a^2$.

**Step (2):**

- As noted above, $\overleftrightarrow{CK} \parallel BH$, so the height of $CBH$ corresponding to side $BH$ is the distance between $\overleftrightarrow{BH}$ and $\overleftrightarrow{CK}$, which is $|BK|$.

- So by our formula for triangle area and the Area Axiom we have area$(CBH) = \frac{1}{2}|BH||BK| = \frac{1}{2}$area$(BHJK)$.

**Step (3):**

- We have

$$\begin{aligned} m\angle EBA &= m\angle EBC + m\angle CBA \\ &= 90 + m\angle CBA \\ &= m\angle ABH + m\angle CBA \\ &= m\angle CBH \end{aligned}$$

- We constructed $D$ and $E$ so that they were on the side of $\overleftrightarrow{BC}$ not containing $A$. From this we see that $\overrightarrow{BC}$ is between $\overrightarrow{BE}$ and $\overrightarrow{BA}$.

- Similarly we conclude $\overrightarrow{BA}$ is between $\overrightarrow{BC}$ and $\overrightarrow{BH}$.

  so $\angle EBA \cong \angle CBH$.

- Because $BCDE$ and $ABHI$ are squares, we have $BA \cong BH$ and $EB \cong CB$.

- Then by Axiom SAS we have $\triangle EBA \cong \triangle CBH$.

**Step (4):**

This follows immediately from **(1)**, **(2)**, and **(3)**:

$$\begin{aligned} \text{area}(BHJK) &= 2\text{area}(CBH) \\ &= 2\text{area}(EBA) \\ &= a^2 \end{aligned}$$

**Step (5):**

The steps are similar to **(1)** through **(4)** above – see Exercise 5.24.

**Step (6):**

$$\begin{aligned} c^2 &= |AB||BH| \\ &= \text{area}(ABHI) \\ &= \text{area}(BHJK) + \text{area}(JKAI) \\ &= a^2 + b^2 \ . \end{aligned}$$

This completes the proof. □

The converse of the Pythagorean Theorem, though not nearly as well-known, is also an extremely useful geometric theorem. Its proof is fairly easy from here.

**THEOREM 5.19.** *If in triangle $ABC$ we have $a^2 + b^2 = c^2$, then $\angle C$ is a right angle.*

*Proof:* Construct a triangle $PQR$ such that $p = a$, $q = b$, and $\angle R$ is a right angle.

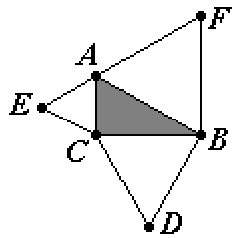
- Applying the Pythagorean Theorem to this right triangle (together with our hypothesis) we see that $r^2 = p^2 + q^2 = a^2 + b^2 = c^2$, so $r = c$.
- By the SSS criterion (Theorem 4.23) we have $\triangle ABC \cong \triangle PQR$.
- Finally, then, we have $\angle R \cong \angle C$, so $\angle C$ must also be a right angle. □

### Exercises

**5.24**. Carry out the steps to show that area$(JKAI) = b^2$ in the proof of Theorem 5.18.

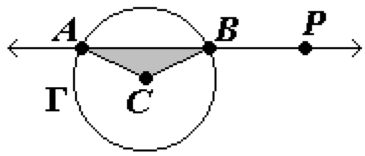
**5.25**. Let $ABC$ be a right triangle with $\angle C$ a right angle. Suppose we find points $D$, $E$, and $F$ so that triangles $BCD$, $ACE$, and $ABF$ are all equilateral. Prove:

$$\text{area}(BCD) + \text{area}(ACE) = \text{area}(ABF)\,.$$




**5.26**. Let $\Gamma$ be a circle with center $C$ and radius $r$. Show that if we are allowed to choose two points $A$ and $B$ anywhere on $\Gamma$, the maximum of area$(ABC)$ will occur when $\angle C$ is a right angle.

**5.27**. Let $\Gamma$ be a circle with center $C$ and let $P$ be a point outside of $\Gamma$. Give a straightedge and compass construction to find the line $\Lambda$ through $P$ intersecting $\Gamma$ in points $A$ and $B$ such that area$(ABC)$ is maximum. (*Hint:* look at Exercise 5.26. The construction in Exercise 2.68 might also be useful.)




**5.28**. Let $ABC$ be a triangle and let $D$ be a point on $AB$ such that $CD \perp AB$. Prove that $\angle ACB$ is a right angle if and only if $|CD|^2 = |AD||DB|$.

**5.29**. Let $\angle ACB$ be a right angle and let $D$ be any point on $BC$. Prove that $|BC|^2 - |DC|^2 = |AB|^2 - |AD|^2$.

# F. Similar Triangles

This section is devoted to introducing another important tool in Euclidean geometry: the notion of triangle similarity (defined as it was in Chapter 1 – see p.17). Here we will see that equality of the ratios of corresponding side lengths in similar triangles can be obtained without too much work from the Pythagorean Theorem. The results in this section are definitely in *Euclidean* geometry, and not in *neutral* geometry. In fact, we will see in Chapter 6 that triangle similarity is actually a non-issue in hyperbolic geometry, since hyperbolic geometry has an "AAA" congruence criterion!

In Chapter 1 we grouped the results of this section under the single heading of the "Similar Triangles Theorem" – see p.17. There are really several statements to prove in that theorem, and our first goal will be to establish that if $\triangle ABC \sim \triangle PQR$ then $p/a = q/b = r/c$. We will need some lemmas to get there.

**LEMMA 5.20.** *Let $ABC$ be a right triangle with right angle $\angle C$. Let a number $0 < t < 1$ be given and let $D$ and $E$ be the points on $AB$ and $AC$ respectively such that $|AD| = tc$ and $|AE| = tb$. Then $DE \parallel BC$, $|DE| = ta$, and $\triangle ABC \sim \triangle ADE$.*

*Proof.* Let $\Lambda_1$ be the line through $E$ parallel to $BC$ (and thus perpendicular to $AC$ by Theorem 5.3) as in Figure 5.12. We will show that $D$ is on $\Lambda_1$ so that $\Lambda_1$ is in fact $\overleftrightarrow{DE}$. The desired conclusions will follow from this.

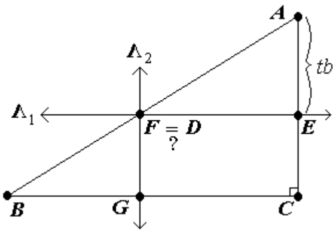



Figure 5.12:

- First, $A$ and $C$ are on opposite sides of $\Lambda_1$ while $B$ and $C$ are on the same side of $\Lambda_1$. So $A$ and $B$ must be on opposite sides of $\Lambda_1$, implying that $\Lambda_1$ intersects $AB$ at a point, say $F$. (We will show that $F = D$.)

- Let $\Lambda_2$ be the line through $F$ parallel to $\overleftrightarrow{AC}$ (and thus perpendicular to $\overleftrightarrow{BC}$).

- From an argument similar to that used for $\Lambda_1$ above, the line $\Lambda_2$ must intersect $BC$ at a point $G$.

- By the Area Axiom we have

$$\begin{aligned} \text{area}(ABC) &= \text{area}(BGF) + \text{area}(GCAF) \\ &= \text{area}(BGF) + \text{area}(GCEF) + \text{area}(AEF) \end{aligned}$$

  where $GCEF$ is a rectangle and both $BGF$ and $AEF$ are right triangles.

- Using our area formulae, this equality becomes

$$\begin{aligned}\frac{1}{2}|AC||BC| &= \frac{1}{2}|FG||BG| + |CE||CG| + \frac{1}{2}|FE||AE| \\ \frac{1}{2}ba &= \frac{1}{2}|CE|(a - |CG|) + |CE||FE| + \frac{tb}{2}|FE| \\ &= \frac{1}{2}(b - tb)(a - |FE|) + (b - tb)|FE| + \frac{tb}{2}|FE|\end{aligned}$$

  which reduces to $|FE| = ta$.

- Since $AEF$ and $ABC$ are both right triangles, we may apply the Pythagorean Theorem to obtain

$$\begin{aligned}|AF| &= \sqrt{|FE|^2 + |AE|^2} \\ &= \sqrt{(ta)^2 + (tb)^2} \\ &= t\sqrt{a^2 + b^2} \\ &= t\sqrt{|BC|^2 + |AC|^2} \\ &= t|AB| \\ &= tc\end{aligned}$$

- Since $F$ and $D$ are points on $\overrightarrow{AB}$ for which $|AF| = tc = |AD|$, they must coincide by Axiom L2.

- This shows that $\Lambda_1 = \overleftrightarrow{DE}$, so
  - $DE \parallel BC$ (since $\Lambda_1$ was constructed parallel to $BC$),
  - $|DE| = |FE| = ta$, and
  - the congruence of corresponding angles (Theorem 5.4) easily gives us $\triangle ABC \sim \triangle ADE$ □

**Lemma 5.21.** *If $ABC$ and $PQR$ are right triangles and $\triangle ABC \sim \triangle PQR$ then $p/a = q/b = r/c$.*

*Proof.* We will assume that the right angles are $\angle C$ and $\angle R$. If $r = c$ then $\triangle ABC \cong \triangle PQR$ by by the SAA criterion. So, assume without loss of generality that $r/c = t < 1$. We will show $p/a = q/b = t$.

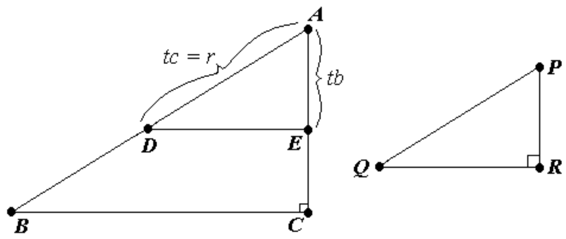



Figure 5.13:

- Choose points $D$ on $AB$ and $E$ on $AC$ so that $|AD| = tc$ $(= r)$ and $|AE| = tb$ (see Figure 5.13).
- By Lemma 5.20 we have $|DE| = ta$ and $\triangle ABC \sim \triangle ADE$.
- By transitivity, $\triangle PQR \sim \triangle ADE$.
- But $|AD| = r = |PQ|$, so $\triangle PQR \cong \triangle ADE$ by the SAA criterion.
- From this congruence we have $p = |DE|$ and $q = |AE|$.
- So, $p/a = |DE|/a = ta/a = t$, and $q/b = |AE|/b = tb/b = t$.

□

The first part of the Similar Triangles Theorem is now within our grasp.

**Theorem 5.22.** *If $\triangle ABC \sim \triangle PQR$ then $p/a = q/b = r/c$.*

*Proof.* Assume $\triangle ABC \sim \triangle PQR$. We will show directly (by cutting these into right triangles and applying Lemma 5.21) that $p/a = q/b = r/c$. Refer to Figure 5.14 in the steps below.

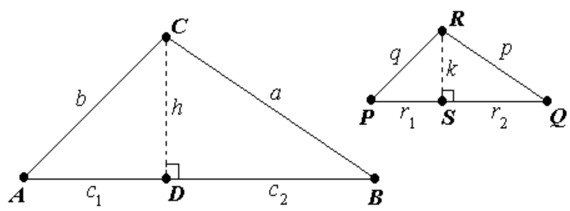



Figure 5.14:

- By Fact 4.25 we may assume that there are points $D$ on $AB$ and $S$ on $PQ$ so that $ADC$, $BDC$, $PSR$, and $QSR$ are all right triangles.

- Let the lengths of the edges of these triangles be labeled as in Figure 5.14. Note that $c = c_1 + c_2$ and $r = r_1 + r_2$.
- By the similarity of $\triangle ABC$ and $\triangle PQR$ we have $\angle CAD \cong \angle RPS$. Also, $\angle ADC \cong \angle PSR$ since these are both right angles. Thus, by Theorem 5.5 we have also $\angle ACD \cong \angle PRS$. So $\triangle ADC \sim \triangle PSR$.
- Similarly, $\triangle BDC \sim \triangle QSR$.
- By Lemma 5.21, then, we have $p/a = r_2/c_2 = k/h = r_1/c_1 = q/b$.
- Let $t$ be the common value of all these ratios. All remains only to show $r/c = t$.
- But since $r_1 = tc_1$ and $r_2 = tc_2$ we have

$$r/c = \frac{r_1 + r_2}{c_1 + c_2} = \frac{tc_1 + tc_2}{c_1 + c_2} = t\,.$$

□

We can now prove the second part of our Similar Triangles Theorem as stated in Section 1B – a sort of "Side-Angle-Side" criterion for similarity.

**THEOREM 5.23.** *Suppose $ABC$ and $PQR$ are triangles such that $\angle A \cong \angle P$ and $q/b = r/c$. Then $\triangle ABC \sim \triangle PQR$.*

*Proof.* We may assume that $q = tb$ and $r = tc$ for some $t < 1$. Refer to Figure 5.15 in the steps below.

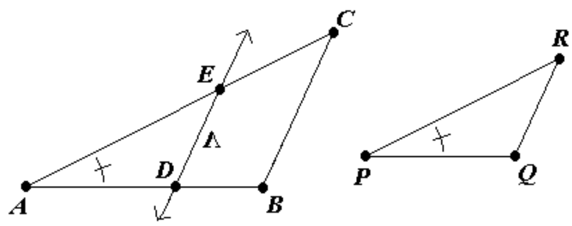



Figure 5.15:

- Let $D$ be the point on $AB$ such that $|AD| = tc$ $(= r)$.
- Let $\Lambda$ be the line through $D$ parallel to $BC$.
- Then $A$ and $B$ are on opposite sides of $\Lambda$ (since $AB$ meets $\Lambda$) but $C$ and $B$ must be on the same side of $\Lambda$ (since $BC$ does *not* meet $\Lambda$).

- So, $A$ and $C$ are on opposite sides of $\Lambda$, which means that $AC$ must meet $\Lambda$ at a point $E$.
- Then $\angle ADE \cong \angle ABC$ and $\angle AED \cong \angle ACB$ since they are pairs of corresponding angles in transversals of parallel lines (Theorem 5.4).
- This shows $\triangle ADE \sim \triangle ABC$, so we must have $|AE|/b = |AD|/c$ by Theorem 5.22.
- Then $|AE| = tb = q$. So by Axiom SAS we have $\triangle ADE \cong \triangle PQR$.
- So, $\triangle ABC \sim \triangle ADE \cong \triangle PQR$ which completes the proof.

$\square$

The only part of the Similar Triangles Theorem yet to be proved is the converse to Theorem 5.22 above. This is really a sort-of "SSS similarity criterion". We leave its proof as an exercise.

THEOREM **5.24.** *Suppose that triangles $ABC$ and $PQR$ satisfy $p/a = q/b = r/c$. Then $\triangle ABC \sim \triangle PQR$.*

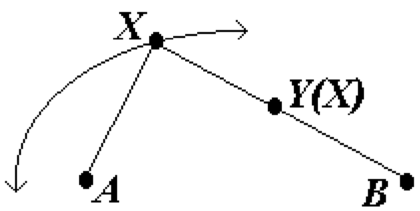
## Exercises

**5.30**. Prove Theorem 5.24.

**5.31**. Assume that $\overleftrightarrow{AB} \parallel \overleftrightarrow{DE}$, $\overleftrightarrow{BC} \parallel \overleftrightarrow{EF}$, and $\overleftrightarrow{AC} \parallel \overleftrightarrow{DF}$. Prove that $\triangle ABC \sim \triangle DEF$.

**5.32**. Suppose you are given three segments $AB$, $CD$, and $EF$. Give a straightedge and compass construction to find the point $G$ on $EF$ so that $|EG|/|FG| = |AB|/|CD|$.

**5.33**. Let $A$ and $B$ be points and let $d$ be a number greater than $|AB|$. The *ellipse* with *foci* $A$ and $B$ and *major axis* $d$ is the set of all points $X$ such that $|AX|+|BX| = d$. For each such point $X$, let $Y(X)$ be the midpoint of the segment $BX$. What is the locus of points $Y(X)$ as $X$ varies over the ellipse?



**5.34**. Let $ABC$ be any triangle and let $D$ be a point on $BC$ so that $\angle BAD \cong \angle CAD$. Prove that $|DC|/|DB| = |AC|/|AB|$.

**5.35**. Let $ABC$ be any triangle and let $D$ and $E$ be points on $\overleftrightarrow{AB}$ and $\overleftrightarrow{AC}$ respectively so that $CD \perp \overleftrightarrow{AB}$ and $BE \perp \overleftrightarrow{AC}$. Prove that $|AB|/|AC| = |AE|/|AD|$.

**5.36**. Prove that if $\triangle ABC \sim \triangle PQR$ and $r = ct$ then $\text{area}(PQR) = t^2\text{area}(ABC)$.

# Chapter 6

# Hyperbolic Geometry

There are two aspects to the material in this chapter. There is, of course, a train of mathematical results we will follow. They lead to the theorems of *hyperbolic geometry*, or the geometry that results when the Euclidean Parallel Axiom is replaced by its negation. But alongside those results and interwoven with them is a story, the story of one of the greatest mathematical discoveries in human history. And to understand that story is to understand the significance of the theorems themselves.

In Chapter 3 we briefly discussed what has come to be known as the "parallel postulate controversy". Recall that, perhaps because the wording of Euclid's fifth postulate simply *sounds* more like a theorem than an axiom, many mathematicians questioned its place among the other axioms. It "ought even to be struck out of the Postulates altogether; for it is a *theorem* involving many difficulties" and its nature is "alien to the special character of postulates". So wrote Proclus (410 - 485 AD), head of the Platonic Academy in Athens at the decline of the Roman empire. It is from Proclus' wonderful commentary on Euclid's *Elements* that we learn much of what we know about early Greek geometry, for Proclus included with his commentary a summary of the work of his mathematical forebears. According to Proclus, doubt as to the place of the parallel postulate began very early after Euclid's career. Most who had such doubts agreed with Proclus' opinion stated above, that the parallel postulate should be a theorem instead of an axiom. In other words, they suspected that it was not independent of Euclid's other assumptions. Had their attempts to prove this postulate using only Euclid's other assumptions succeeded, it would have established Euclidean geometry and neutral geometry as one and the same, effectively crowning Euclidean geometry as the only alternative satisfying the basic properties of the earlier axioms.

Of course, this was not to be. Euclidean geometry is but one of the possible options for a geometric reality encompassing neutral geometry, but the realization of that fact was very slow in coming. We can attribute the more than 21 century delay between Euclid's career and the discovery of another possible geometric world, now called hyperbolic geometry, to two major factors.

First, the mathematical analysis required for the discovery was difficult and required tremendous subtlety. Many of the greatest mathematicians in those intervening centuries attempted proofs of the parallel postulate, and many thought they had succeeded. Only careful subsequent analysis showed the fatal flaws of circularity in their reasoning – they were all guilty of, in one way or another, making an assumption equivalent to what they were attempting to prove. The very fact that they fell prey to such flaws should be viewed as evidence that human understanding of the axiomatic method had simply not grown to a level of sophistication sufficient to motivate the discovery.

Second, the prevailing philosophical assumption was that the inevitable conclusion of the struggle over the parallel postulate would be, in one way or another, the establishment of Euclidean geometry as the one reigning geometric reality. The geometry Euclid described by his axioms was assumed to be fundamental to the nature of the universe. It was utterly inconceivable that there might be other equally valid geometries to discover with equal claim to being the "reality" by which the universe is framed. The influential 18th century philosopher Immanuel Kant went so far as to propose that the structure of Euclidean space is inherent in the human mind, calling it an "inevitable necessity of thought." With this attitude firmly entrenched, it is not surprising that nobody was really *looking* to discover a geometry different from that of Euclid. It makes the discovery of hyperbolic geometry even more remarkable to consider that it occurred, for the most part, accidentally.

Our approach to the material of this chapter will reflect somewhat the way it developed historically, in that we will work *within neutral geometry* as long as possible.

- We begin in Section A with neutral geometry attempts to approximate the 180° Sum Theorem.

- Sections B and C then take us through some of the failed attempts to establish the parallel postulate as a neutral geometry theorem. Though none of these "proofs" accomplished what they set out to do, they did accomplish something very meaningful, for the insights they developed along the way led to the very doorstep of the new geometric world.

- In Section D we finally cross this threshold by introducing the Hyperbolic Parallel Axiom as a substitute for its Euclidean counterpart. The neutral geometry investigations of the previous sections then immediately yield the surprising fundamental results of the new geometry.

- Section E introduces a model by which we know that this new geometry is no less consistent than the more familiar Euclidean geometry.

- Finally, Sections F, G, and H work out some technical consequences of the Hyperbolic Parallel Axiom.

## A. A Little More Neutral Geometry

As noted in the above introduction, the discovery of hyperbolic geometry grew out of attempts to prove the Euclidean Parallel Axiom as a theorem of neutral geometry. So, as we begin our journey toward this discovery ourselves, we must begin with neutral geometry.

In this section we will trace a short sequence of theorems that are best understood as an attempted neutral geometry approach to the 180° Sum Theorem. With Theorem 5.5 we saw how easily that famous result is proved using the Euclidean Parallel Axiom. But attempts to prove it within neutral geometry always came up short. The main result of this section, the Saccheri-Legendre Theorem (which states that the sum of the measures of the three angles in a triangle is *no more than* – but possibly *less than* – 180°), is as close as one can get within neutral geometry.

In general, the results below do not lead (as the neutral geometry theorems of Chapter 4 did) to interesting facts in Euclidean geometry. Instead, they push in a different direction, and will be crucial to our development of hyperbolic geometry as a different extension of the neutral geometry material. Our starting point is the result usually called the Exterior Angle Theorem.[1]

**Theorem 6.1.** *Suppose $ABC$ is a triangle and $D$ is a point on $\overleftrightarrow{BC}$ so that $B * C * D$. Then $m\angle ACD > m\angle A$ and $m\angle ACD > m\angle B$.*

[1] The angle $\angle ACD$ in the statement of Theorem 6.1 is often called an *exterior angle* to the triangle $ABC$.

*Note:* This result is often stated "*the measure of an exterior angle to a triangle is greater than the measure of either of its remote interior angles.*" It would follow easily from Theorem 5.5, for by that theorem we have $m\angle A + m\angle B = 180 - m\angle ACB = m\angle ACD$. Here, of course, we must prove our theorem solely within the confines of neutral geometry.

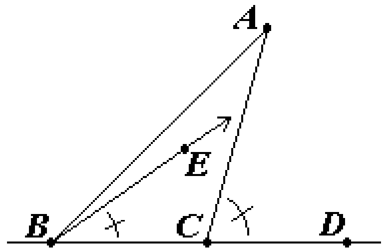



Figure 6.1:

*Proof:* Let the points be named as described. We will prove by contradiction that $m\angle ACD > m\angle B$, leaving the (similar) proof that $m\angle ACD > m\angle A$ as an exercise. Assume, then, (to reach a contradiction) that $m\angle ACD \leq m\angle B$.

- Then we may find a ray $\overrightarrow{BE}$ with $E$ either on $\overrightarrow{BA}$ or interior to $\angle ABC$ so that $m\angle EBC = m\angle ACD$ (see Figure 6.1). Note that in either case $\overrightarrow{BE}$ meets the segment $AC$. (We invite you to work out how that detail can be justified using Axioms AC and C4.)

- Then the transversal of $\overleftrightarrow{BE}$ and $\overleftrightarrow{CA}$ by $\overleftrightarrow{BC}$ has a pair of congruent corresponding angles, so by Theorem 4.14 we have $\overleftrightarrow{BE} \parallel \overleftrightarrow{CA}$.

- This is, of course, our contradiction. These lines cannot be parallel becuase, as noted, $\overrightarrow{BE}$ meets $AC$. □

The following corollary to Theorem 6.1 is well-known and quite useful. We leave its proof as an exercise.

COROLLARY **6.2.** *In any triangle ABC the smaller angle is always opposite the shorter side. That is, if $a < b$ then $m\angle A < m\angle B$.*

We also leave as an exercise the easy proof of the following consequence of Theorem 6.1.

LEMMA **6.3.** *The sum of the measures of any two angles in a triangle is less than 180.*

We are now ready to accomplish the principal goal of this section: a proof of the Saccheri-Legendre Theorem. (We will learn something of the mathematicians for whom this result is named in the next section.) It should be noted that this theorem is definitely the strongest statement about angle measure sums for triangles that can be made within neutral geometry. For we know already that in Euclidean geometry the sum is exactly 180, and we will shortly see that in hyperbolic geometry the sum is always a number less than 180.

**Saccheri-Legendre Theorem.** *If $ABC$ is any triangle then $m\angle A + m\angle B + m\angle C \leq 180$.*

*Proof:* Assume, to reach a contradiction, that $ABC$ is a triangle with angle measure sum equal to $180 + \delta$ for some $\delta > 0$. We will show below that we can construct, beginning with $ABC$, a triangle with the same angle measure sum, but with one angle measuring less than $\delta$. This, of course, will give us our contradiction, for the remaining two angles of this triangle would have measures summing to more than 180, contrary to Lemma 6.3.

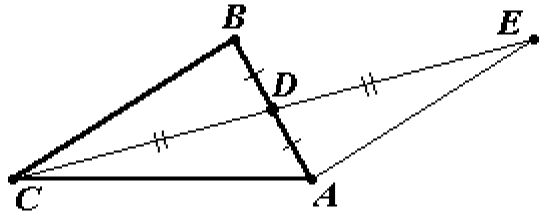



Figure 6.2:

It remains, then, only to show how this triangle is produced. Let's suppose that $m\angle A = \alpha$, $m\angle B = \beta$, and $m\angle C = \gamma$, with $\alpha \geq \beta \geq \gamma$. It will be enough to show that we can produce a triangle with angle sum equal to that of $ABC$ and with an angle of measure no more than $\gamma/2$, for then repeated application of the process would lead to a triangle with angle sum $180 + \delta$ and an angle of measure as small as we like. The following construction accomplishes this.

- Let $D$ be the midpoint of $AB$ and let $E$ be the point on $\overrightarrow{CD}$ so that $|CE| = 2|CD|$ (so that $D$ is the midpoint of $CE$ – see Figure 6.2).
- The vertical angles $\angle ADE$ and $\angle BDC$ are congruent, so by Axiom SAS we have $\triangle ADE \cong \triangle BDC$.
- It follows from this congruence that $\angle BCD \cong \angle E$ and $\angle DAE \cong \angle B$ (so that $m\angle DAE = \beta$).
- We have $m\angle CAE = m\angle CAB + m\angle BAE = \alpha + \beta$ and $m\angle ACE = m\angle ACB - m\angle BCD = \gamma - m\angle E$.
- The angle measure sum of $ACE$ is then $(\alpha + \beta) + (\gamma - m\angle E) + m\angle E = \alpha + \beta + \gamma$, the same as for $ABC$.

- But since $m\angle CAE = \alpha + \beta$, the other two angles of this triangle have measures whose sum is $\gamma$, and so one of them must have sum no larger than $\gamma/2$.

- This proves our claim and thus establishes the entire theorem. $\square$

Angle measure sums for triangles and convex quadrilaterals will play an important role in the material we develop for Section B. To facilitate the discussion it is traditional to introduce the following definition.

**DEFINITION.** Let $ABC$ be a triangle. The *defect* of this triangle is the number $d(ABC) = 180 - m\angle A - m\angle B - m\angle C$. Similarly, if $EFGH$ is a convex quadrilateral then its defect is defined to be $d(EFGH) = 360 - m\angle E - m\angle F - m\angle G - m\angle H$.

Defects behave nicely with regard to decomposing triangles and quadrilaterals, as noted in the following fact.

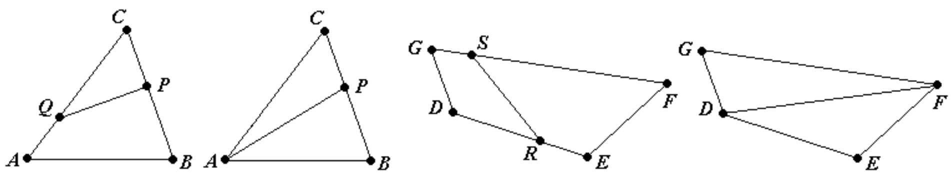



Figure 6.3:

**THEOREM 6.4.** *Let $ABC$ be a triangle with $P$ on side $BC$ and $Q$ on side $AC$, and let $DEFG$ be a convex quadrilateral with $R$ on side $DE$ and $S$ on side $FG$ (as in Figure 6.3). Then:*

- $d(ABC) = d(ABPQ) + d(CPQ) = d(ABP) + d(ACP)$.
- $d(DEFG) = d(DRSG) + d(EFSR) = d(DEF) + d(FGD)$.

*Proof:* We will give here the proof to one part; the others are proved very similarly and the details are left as Exercise 6.5. To show $d(ABC) = d(ABP) + d(ACP)$

we simply calculate:

$$\begin{aligned}
d(ABC) &= 180 - m\angle BAC - m\angle B - m\angle C \\
&= 180 - (m\angle BAP + m\angle PAC) - m\angle B - m\angle C \\
&= 180 - m\angle BAP - m\angle PAC - m\angle B - m\angle C \\
&\qquad + (180 - m\angle BPA - m\angle CPA) \\
&= (180 - m\angle BAP - m\angle B - m\angle BPA) \\
&\qquad + (180 - m\angle PAC - m\angle C - m\angle CPA) \\
&= d(ABP) + d(ACP)
\end{aligned}$$

□

We close this section by restating (in terms of defect) the fact we will need for the material in Section B.

THEOREM **6.5.** *The defect of any triangle or any convex quadrilateral is always a nonnegative number.*

*Proof.* For triangles, this is just a restatement of the Saccheri-Legendre Theorem. For a convex quadrilateral, merely use Theorem 6.4 in conjunction with the Saccheri-Legendre Theorem. □

## Exercises

**6.1**. Complete the proof of Theorem 6.1 by showing that $m\angle ACD > m\angle A$.

**6.2**. Prove Corollary 6.2. (*Hint:* Suppose in triangle $ABC$ that $a < b$. Let $D$ be a point on $AC$ with $|DC| = |BC|$ and apply Theorem 6.1 to triangle $ABD$.)

**6.3**. Prove Lemma 6.3.

**6.4**. State and prove a theorem on angle measure sums for convex $n$-gons.

**6.5**. Prove the remaining parts of Theorem 6.4:

**(a)** $d(ABC) = d(ABPQ) + d(CPQ)$

**(b)** $d(DEFG) = d(DRSG) + d(EFSR)$

**(c)** $d(DEFG) = d(DEF) + d(FGD)$.

**6.6**. Refer back to Figure 3.2. The triangle shown there clearly has angle measure sum greater than 180. Why doesn't this violate the Saccheri-Legendre Theorem?

# B. Struggles Over the Parallel Postulate

As we have noted, Euclid's *Elements* is undoubtedly one of the most successful texts of all time. It remained as *the* standard approach to learning geometry for more than 2000 years. Yet, as we have also noted, critics of Euclid's use of the parallel postulate were quick to emerge. Since the ideas brought forward by these critics proved critical in the development of hyperbolic geometry, we give here an account (though brief and incomplete) of their work.

There is a common thread running through the attempts we will describe. Each sought to remove the parallel postulate from the list of required assumptions for geometry. If we let $\{\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3, \mathbf{E}_4, \mathbf{E}_5\}$ represent Euclid's set of five axioms (with $\mathbf{E}_5$ being the parallel postulate) we can indicate the goal of these would-be parallel postulate provers as

$$\text{Desired proof :} \quad \{\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3, \mathbf{E}_4\} \implies \mathbf{E}_5$$

But each of their proofs, under careful scrutiny, was found to include use of some additional (unacknowledged) assumption. If this additional assumption is denoted by $\mathbf{A}$, then what the proof actually shows is

$$\{\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3, \mathbf{E}_4, \mathbf{A}\} \implies \mathbf{E}_5$$

Now the additional assumption $\mathbf{A}$ (which varied from case to case) was always a fact from Euclidean geometry, so we also have the implication

$$\{\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3, \mathbf{E}_4, \mathbf{E}_5\} \implies \mathbf{A}$$

Thus, $\mathbf{E}_5$ and $\mathbf{A}$ are then *equivalent under the other assumptions of neutral geometry* – if one of them is true then they both must be true. Each attempted proof of the parallel postulate did not actually remove $\mathbf{E}_5$ from the list of required axioms, but instead *replaced* it with a different but equivalent assumption!

The end result of these attempts, then, was the compilation of a list of statements equivalent to the parallel postulate (relative to the other assumptions of neutral geometry). Specifically, consider the following list of statements.

**A.** If line $\Lambda_0$ transverses lines $\Lambda_1$ and $\Lambda_2$ and the interior angles of this transversal on one side of $\Lambda_0$ have measures adding to less than 180 then $\Lambda_1$ and $\Lambda_2$ intersect at a point on that side of $\Lambda_0$.

**B.** If $P$ is a point not on a line $\Lambda$ then there is exactly one line through $P$ parallel to $\Lambda$.

**C.** If $\overleftrightarrow{PQ}$ is parallel to a line $\Lambda$ then the distance from $P$ to $\Lambda$ equals the distance from $Q$ to $\Lambda$.

**D.** There exists a triangle with zero defect.

**E.** There exists a rectangle.

**F.** Every triangle has zero defect.

**G.** Every convex quadrilateral has zero defect.

Statements A and B are, respectively, Euclid's and Playfair's parallel postulates, and their equivalence has already been noted (see Exercise 3.22). We already know that accepting either A or B as an axiom (along with the axioms of neutral geometry) allows us to prove any of the statements C through G (see Fact 5.8, Theorem 5.5, and Exercise 6.8). In this section and the next, motivated by attempted proofs from the parallel postulate's critics, we will see that *all* of these statements have equal power in neutral geometry: accepting any one of A through G as an axiom implies the truth of them all!

**Early critics: Posidonius and Ptolemy**

In the first century BC, Posidonius (a follower of the Stoic school of philosophy) contended that Euclid's last axiom could be proved if only a better definition of parallel lines was accepted. Posidonius proposed replacing Euclid's definition of parallel lines (lines that do not intersect each other) with the following:

*Lines $\Lambda$ and $\Lambda'$ are parallel if and only if they are everywhere equidistant from each other.*

In other words, Posidonius would $\Lambda'$ be considered parallel to $\Lambda$ only if all points of $\Lambda'$ are equal distance from $\Lambda$. His writings have not survived to the present, so we can only infer his reasoning, but Posidonius might have arrived at this definition as follows (see Figure 6.4 for reference).

- Suppose $\Lambda$ is a line and $P$ is a point not on $\Lambda$, and suppose that the distance from $P$ to $\Lambda$ is $d$.

- Consider a ray perpendicular to $\Lambda$ and lying on the same side of $\Lambda$ as does $P$. This ray can contain only one point $Q$ whose distance to $\Lambda$ is $d$.

- Now suppose $\Lambda'$ is a line through $P$ parallel to $\Lambda$ (by Posidonius' definition). Then if $\Lambda'$ meets the ray we have constructed, it must pass through $Q$ – that is, $\Lambda'$ must *be* the line $\overleftrightarrow{PQ}$.

- So $\overleftrightarrow{PQ}$ is the only candidate for a parallel to $\Lambda$ passing through $P$.

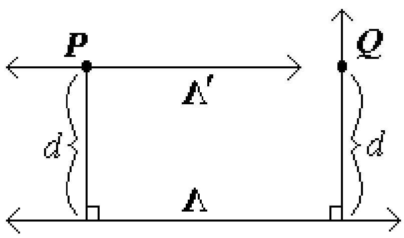



Figure 6.4: Posidonius' reasoning

This is perfectly sound reasoning as far as it goes. The trouble, however, is that we are indeed asking too much. Note the wording "*if* $\Lambda'$ meets the ray" (as opposed to asserting that it *does* meet the ray) and "$\overleftrightarrow{PQ}$ is the only *candidate* for a parallel" (as opposed to asserting that is *is* a parallel). There is a good reason for such hedging in our words! If we had Theorem 5.3 available we would indeed be able to say with certitude that $\Lambda'$ meets the ray in question. (Check this!) But of course, we don't have Theorem 5.3 unless we are willing to assume the parallel postulate, for it is not a theorem of neutral geometry. Furthermore, there is no guarantee that the line $\overleftrightarrow{PQ}$ meets Posidonius' definition of a parallel to $\Lambda$ *unless* we also accept the truth of statement C from our list. Without that assumption it may be that $\overleftrightarrow{PQ}$ is distance $d$ from $\Lambda$ at $P$ and $Q$ only, with other points on $\overleftrightarrow{PQ}$ being either further from or closer to $\Lambda$. So to guarantee that $\overleftrightarrow{PQ}$ meets Posidonius' definition of a parallel to $\Lambda$ we must accept statement C and thus also the entire list of equivalent statements. Far from freeing us of it, Posidonius' definition *necessitates* the parallel postulate! Using his definition alone we could not even prove the *existence* of parallel lines in neutral geometry.

Ptolemy of Alexandria (2nd century AD) (who we have already met in Chapter 2) may have been one of the first to attempt a proof of Euclid's fifth postulate based on the other postulates. This is quite different in nature from what Posidonius did, for Ptolemy made no changes in Euclid's definitions. While Ptolemy's effort is notable in being the first such attempt, it was far from successful. In effect, Ptolemy proved Euclid's version of the parallel postulate (statement A in the above list) by using the other axioms plus an unacknowledged assumption essentially like statement B.

Many mathematicians would follow Ptolemy's footsteps in the centuries that followed, each searching for the elusive proof of the parallel postulate. But all such attempts would have one of two endings: either (like Ptolemy's) they would produce a proof convincing to its author but found under later examination to contain a hidden assumption of one of the statements in our list, or else (for those fortunate enough not to fool themselves into a false proof) would reach no conclusion at all.

**Proclus**

The next notable attempt to prove the parallel postulate was conducted by Proclus of Alexandria (410 - 485 AD), mentioned in the introduction to this chapter. Proclus followed his criticism of Ptolemy's attempted proof with one of his own. His reasoning can be summarized as follows (see Figure 6.5).

- Let $\Lambda$ be a line and let $P$ be a point not on $\Lambda$. Let $Q$ be the point on $\Lambda$ so that $PQ \perp \Lambda$.
- Let $\Lambda'$ be the line through $P$ perpendicular to $PQ$. Note that $\Lambda' \parallel \Lambda$ by the neutral half of the Alternate Interior Angles Theorem (Theorem 4.12).
- Suppose (to reach a contradiction, or so Proclus hoped!) that there is another parallel to $\Lambda$ passing through $P$, say $\Lambda''$, and consider the ray determined by this line on the same side of $\Lambda'$ as $Q$.
- For each point $R$ on this ray let $R'$ be the point on $\Lambda'$ so that $RR' \perp \Lambda'$.
- Proclus reasoned that as $R$ is moved further out from $P$ along this ray the distance $|RR'|$ would tend to infinity and thus would at some point exceed $|PQ|$.
- This led Proclus to declare a contradiction, for he concluded that for this to occur $R$ must then be on the side of $\Lambda$ opposite $P$. In other words, $\Lambda''$ must contain points on both sides of $\Lambda$ and thus must intersect $\Lambda$.

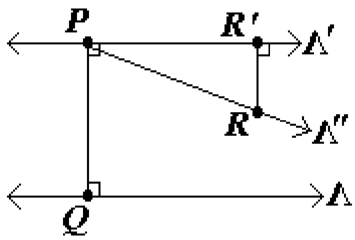



Figure 6.5: Proclus' reasoning

Can you tell what assertion is unsupported? If you guessed the claim that $|RR'|$ tends to infinity, you are partly correct. We have certainly not proved this and did not give a reason for it in the above summary. However, this assertion turns out to be true (though quite difficult to prove). The fatal flaw in Proclus' reasoning is in an assumption hidden in his declaration of a contradiction. When he concludes that $\Lambda''$ must cross $\Lambda$ in order for $|RR'|$ to surpass $|PQ|$, Proclus is assuming that $\Lambda$ maintains a constant distance (namely $|PQ|$) from $\Lambda'$. In other words, he is assuming statement C from our list. Without this assumption Proclus' contradiction disappears: $|RR'|$ may grow to infinity at the same time the distance from $R$ to $\Lambda$ also diverges! (It is understandable if this seems unbelievable to you at this point, especially as you examine Figure 6.5. But bear with us patiently! We will see in Section 6H that if the Euclidean Parallel Axiom is false then this is exactly what happens.)

Proclus' attempt is important for at least two reasons. First, his technique is useful, though not in the way he envisioned. We will use the above construction later (in the proof of Theorem 6.7) to establish a major step in our development of hyperbolic geometry. Second, Proclus' work illustrates well just how careful we must be to avoid hidden assumptions. We must be certain that *every assertion* in our proof is backed up with a reason, either an axiom or a previously proved theorem. Unless this precaution is followed we may make a seemingly obvious claim that in fact hides an assumption equivalent to what we are trying to prove. Such was the case with Proclus.

**Arab contributions in the middle ages**

During the centuries of the middle ages when the intellectual life of Europe withered, much of the learning of ancient Greece was preserved and refined by Arabic scholars. Arabic mathematicians translated the Greek works and enhanced them with many of their own original contributions. These Arabic manuscripts would later provide an important bridge between between the ancient world and a new mathematical awakening across Europe in the 15th and 16th centuries.

The problem of the parallel postulate seemed to hold extraordinary interest for several of the Arabic mathematicians in the middle ages. Three names of note are ibn-al-Haitham (more commonly known as Alhazen, 965 - 1039 AD), Omar Khayyam (famous as a poet, but also a respected scientist, 1050 - 1123 AD), and Nasir Eddin al-Tusi (1201 - 1274 AD). These three men realized the importance of statement E in our list. That is, they knew that if a rectangle could be constructed in neutral geometry then the parallel postulate of Euclid would follow as a theorem. (We'll prove this ourselves shortly.)

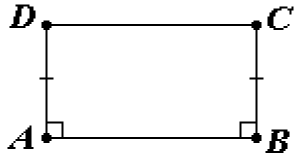



Figure 6.6: A Saccheri quadrilateral

Recall that in Section 4B we briefly discussed the construction of rectangles and its relation to neutral and Euclidean geometry (see p.163). We found that it is easy to construct the quadrilateral $ABCD$ in Figure 6.6, and we called such a figure a *Saccheri quadrilateral*. Two of its angles ($\angle A$ and $\angle B$) are right angles by construction, so to show that $ABCD$ is a rectangle we need only verify that $\angle C$ and $\angle D$ are right angles. This was the battle plan of Omar Khayyam and al-Tusi[2] – they aimed to show $ABCD$ was a rectangle without using the parallel postulate. Predictably, their proofs do not work, poisoned by unjustified assumptions that turn out to be equivalent to the parallel postulate itself. Khayyam, for instance, assumed that two lines not at constant distance from one another must of necessity intersect – something you should recognize as the contrapositive of statement C from our list.

Saccheri quadrilaterals will play an important role in what lies ahead, so we'll take this opportunity now to give them a more formal definition and to outline some of their basic properties.

**Definition.** A *Saccheri quadrilateral* is a quadrilateral $ABCD$ such that $\angle A$ and $\angle B$ are right angles and $AD \cong BC$ (as in Figure 6.6). The segment $AB$ is called its *base* while $CD$ is called its *summit*. The angles $\angle C$ and $\angle D$ are called its *summit angles*.

**Theorem 6.6.** *Under the assumptions of neutral geometry, Saccheri quadrilaterals have the following properties.*

**(a)** *The segment joining the midpoints of the summit and base of a Saccheri quadrilateral is perpendicular to both summit and base.*

**(b)** *The summit angles of a Saccheri quadrilateral are congruent.*

The challenge to prove these statements was included previously as Exercise 5.6. If you did not do so previously, take the time now to work through the proofs.

Note that from part (a) of Theorem 6.6, we can use one side of a Saccheri quadrilateral together with the midpoints of its summit and base to create a new kind of "near-rectangle" – one with right angles at three of its vertices. This type of quadrilateral also bears the name of a mathematician we will meet shortly.

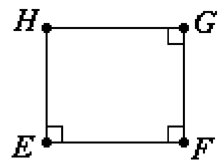


Figure 6.7: A Lambert quadrilateral

**Definition.** A *Lambert quadrilateral* is a quadrilateral $EFGH$ such that $\angle E$, $\angle F$, and $\angle G$ are right angles (as in Figure 6.7).

[2]It is perhaps only by cultural bias that the Saccheri quadrilateral was named for a post-Renaissance Italian mathematician (of whom we will learn shortly) instead of for one of the Arabic mathematicians who considered it centuries earlier.

Alhazen began his investigations with a Lambert quadrilateral[3], attempting to prove that the fourth angle was also a right angle. He managed to do so, but his proof carries the usual flaw of implicitly assuming statement $C$ from our list.

**Saccheri**

Several centuries after al-Tusi, the Jesuit teacher Girolamo Saccheri (1667 - 1733 AD) revisited the near-rectangle now named for him. His strategy was to use proof by contradiction applied to al-Tusi's construction. He reasoned that if the summit angles of the quadrilateral were other than right angles then (since the summit angles are congruent) they must either be both greater than 90 in measure or both less than 90 in measure. He sought then to eliminate each of these possibilities, called respectively the *obtuse angle hypothesis* and the *acute angle hypothesis*. The Saccheri-Legendre Theorem (which Saccheri derived) soon eliminates the obtuse angle hypothesis, for the defect of the Saccheri quadrilateral cannot be negative (see Theorem 6.5). This left only the acute angle hypothesis to eliminate by contradiction.

Saccheri worked diligently toward this goal, deriving a list of unlikely-sounding consequences of the acute angle hypothesis, desperately trying to find a contradiction among them. Unable to *show* a contradiction in a satisfactorily straightforward manner, Saccheri was at last reduced to *declaring* one by insisting that the consequences he had proved for the acute angle hypothesis are "repugnant to the nature of the straight line!" The notion he found so "repugnant" was the possibility that statement C from our list might not hold. Whether or not Saccheri recognized statement C as equivalent to the parallel postulate, his discomfort with the manner in which he at last dismissed the acute angle hypothesis is evident in his own statement.

> It is well to consider here a notable difference between the foregoing refutations of the two hypotheses. For in regard to the hypothesis of the obtuse angle the thing is clearer than midday light...
>
> But on the contrary I do not attain to proving the falsity of the other hypothesis, that of the acute angle, without previously proving that the line, all of whose points are equidistant from an assumed straight line lying in the same plane with it, is equal to [a] straight line.

[3]Again, if historical justice had held more influence and cultural bias held less, this object might well go by the name *Alhazen quadrilateral*.

Of course, Saccheri had not really "proved" statement C, but merely insisted that it *must* be true to uphold the "nature of the straight line."

Saccheri displayed remarkable skill in proving the consequences of the acute angle hypothesis. But he lacked whatever was required (courage or foresight or both) to recognize that this material was a new geometry, separate from Euclidean geometry. He published his work in a book titled *Euclides ab Omni Naevo Vindicatus*, or *Euclid Cleared of Every Flaw*, an ironic title considering what its contents nearly accomplished. Had Saccheri made but one leap – had he accepted his own results for what they really were instead of twisting them into a contradiction they did not actually contain – it would be to him that we credit the discovery of hyperbolic geometry. As it was, though, that leap would not be taken by anyone for yet another century.

**Lambert**

During that century another attempt similar to Saccheri's would be waged by Johann Heinrich Lambert (1728 - 1777 AD), a Swiss-German philosopher and mathematician. Lambert also tried to demonstrate the existence of a rectangle by showing that the obtuse angle and acute angle hypotheses lead to contradictions. (This is the Lambert for whom the Lambert quadrilateral, being "half of a Saccheri quadrilateral", is named.) Like Saccheri, he successfully ruled out the obtuse angle hypothesis, and also like Saccheri he manufactured a contradiction from the acute angle hypothesis only by using an unjustified assumption.

Along the way he also derived the striking consequences of the acute angle hypothesis that had so infuriated Saccheri. Lambert's reaction to these consequences was quite different, however, for he demonstrated at least the ability to conceive of them as valid theorems. He wrote that some of the consequences were tantalizing enough to make one "wish that the Hypothesis of the Acute Angle were true." He continued, however, by explaining that he ultimately "would not want it to be so, for this would result in countless inconveniences."

Lambert never published his work (it was published only after his death), perhaps because he was never completely comfortable with his "proof" that the acute angle hypothesis cannot hold. Lambert approached his work thoughtfully, and though he did not live to see the matter finally resolved, he summarized well all of the efforts up to his day in the following statement.

> Proofs of the Euclidean postulate can be developed to such an extent that apparently a mere trifle remains. But a careful analysis shows that in this seeming trifle lies the crux of the matter; usually

> it contains either the proposition that is being proved or a postulate equivalent to it.

As we've seen in this brief account, statement C from our list was the most frequently assumed equivalent postulate or overlooked "trifle." In the next section we'll examine that entire list of statements and prove their equivalence. This will clarify the role played by the parallel postulate in Euclidean geometry and delineate the consequences for denying the parallel postulate in hyperbolic geometry.

### Exercises

**6.7**. Prove that if $ABCD$ is a Saccheri quadrilateral with base $AB$ then $\overleftrightarrow{CD} \parallel \overleftrightarrow{AB}$. (Your proof, of course, should use nothing outside of neutral geometry.)

**6.8**. Prove that in Euclidean geometry, every Saccheri quadrilateral is a rectangle.

**6.9**. Show formally that Proclus' reasoning is correct if we assume statement C to be true.

**6.10**. Prove that under the hypothesis of the acute angle there exist triangles of positive defect.

## C. Axioms Equivalent to the Parallel Postulate

Not everyone who encountered the problem of the parallel postulate attempted a proof of the axiom. Instead, some took up the task of *replacing* the parallel postulate with some other less controversial assumption. Among these attempts was that of the French mathematician Alexis Claude Clairaut (1713 - 1765 AD) who proposed as a replacement for the parallel postulate what came to be known as *Clairaut's Axiom: there exists a rectangle.*

Certainly putting this forth as a new axiom is a safer and easier task than trying to *prove* it within neutral geometry, as did Saccheri and Lambert! But in this section's one theorem we will show that Clairaut was correct. This simple proposition (or any of the others on our list from the last section) *cannot be proved in neutral geometry, but may be adopted as an axiom to replace the* *parallel postulate.* This theorem becomes truly tantalizing when we consider its contrapositive: *if we assume the parallel postulate is false, then all of the statements A through G must fail*! That realization, as we will see in the next section, is where the threshold into hyperbolic geometry is crossed.

THEOREM **6.7.** *Under the assumptions of neutral geometry the following statements are equivalent to each other.*

**A.** *If line $\Lambda_0$ transverses lines $\Lambda_1$ and $\Lambda_2$ and the interior angles of this transversal on one side of $\Lambda_0$ have measures adding to less than 180 then $\Lambda_1$ and $\Lambda_2$ intersect at a point on that side of $\Lambda_0$.*

**B.** *If $P$ is a point not on a line $\Lambda$ then there is exactly one line through $P$ parallel to $\Lambda$.*

**C.** *If $\overleftrightarrow{PQ}$ is parallel to a line $\Lambda$ then the distance from $P$ to $\Lambda$ equals the distance from $Q$ to $\Lambda$.*

**D.** *There exists a triangle with zero defect.*

**E.** *There exists a rectangle.*

**F.** *Every triangle has zero defect.*

**G.** *Every convex quadrilateral has zero defect.*

The proof of this theorem consists of several pieces as each implication indicated in Figure 6.8 is established. You may check (by following the arrows) that this shows each statement in the list implies each of the others. Fortunately, several of these implications are already known to us. Exercise 3.22 outlines the proof of **A** $\iff$ **B**, while our development of Euclidean geometry established the fact that **B** implies each of **C, D, E, F**, and **G**. It remains then to prove the implications between these last five statements plus the implication **G** $\implies$ **B**. Some of these proofs are outlined in the exercises and it is left to you to fill in the details. The details are not difficult, and you will gain a greater understanding of this theorem by working through them.

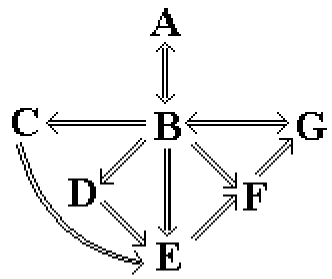



Figure 6.8: The plan for establishing Theorem 6.7

*Proof of* **C** $\implies$ **E***:* Assuming that statement C holds true we will use a procedure reminiscent of Posidonius' reasoning (see p.193) to construct a rectangle. All of the steps in the following construction are justified in neutral geometry. See Figure 6.9 for reference.

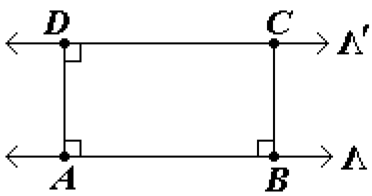



Figure 6.9:

- From a point $A$ on a line $\Lambda$, construct segment $AD$ perpendicular to $\Lambda$.
- Form the line $\Lambda'$ through $D$ perpendicular to $\overleftrightarrow{DA}$.
- From the neutral half of the Alternate Interior Angles Theorem (Theorem 4.12) we know that $\Lambda \parallel \Lambda'$.
- Let $C$ be a point on $\Lambda'$ other than $D$ and let $B$ be the point on $\Lambda$ so that $BC \perp \Lambda$ (see Fact 4.20).
- From statement C we conclude $|AD| = |BC|$, so $ABCD$ is a Saccheri quadrilateral with base $AB$.
- Theorem 6.6 tells us that the summit angles of this quadrilateral are congruent.
- Since the summit angle at $D$ is a right angle, both summit angles must be right angles so that $ABCD$ is a rectangle.

*Proof of* **D** $\implies$ **E**: See Exercise 6.11.

*Proof of* **E** $\implies$ **F**: Assume that a rectangle $ABCD$ exists. We will prove that every triangle then has zero defect. There are three major steps outlined here, with details of the first step left as an exercise.

- First, by "doubling" $ABCD$ we may construct a rectangle with side lengths $2|AB|$ and $2|BC|$ (see Exercise 6.12). If follows then that given any number $r > 0$ we may find a rectangle with both sides of length greater than $r$.
- Next, let $PQR$ be any right triangle with right angle $\angle R$. We will show that $d(PQR) = 0$. (See Figure 6.10 for reference.)

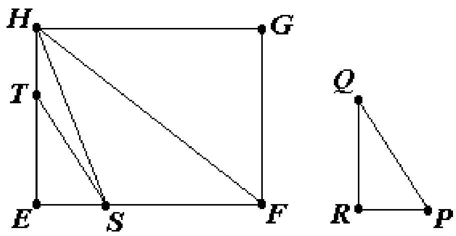



Figure 6.10:

  - By the previous step we may find a rectangle $EFGH$ with minimum side length greater than either $|PR|$ or $|QR|$.

- Let $S$ be a point on $EF$ and $T$ a point on $EH$ so that $ES \cong RP$ and $ET \cong RQ$.
- By Axiom SAS we have $\triangle STE \cong \triangle PQR$.
- But now using the fact that $d(EFGH) = 0$ together with Theorem 6.4 and Theorem 6.5 we see that each of the triangles $FGH$, $FHS$, $STH$, and $STE$ must have zero defect.
- So (since congruent triangles have identical angle measures and thus have the same defect), $d(PQR) = d(STE) = 0$.

• Finally, since every right triangle has zero defect, if follows from Fact 4.25 and Theorem 6.4 that all triangles have zero defect.

*Proof of* **F** $\implies$ **G***:* See Exercise 6.13.

*Proof of* **G** $\implies$ **B***:* This is probably the subtlest portion of the proof. We will follow Proclus' reasoning (as outlined in the previous section), but work under the hypothesis that every convex quadrilateral has defect zero. So, as in Proclus' construction, let $\Lambda$ be a line, $P$ a point not on $\Lambda$, and $Q$ the point on $\Lambda$ so that $PQ \perp \Lambda$. Also let $\Lambda'$ be the line through $P$ perpendicular to $\Lambda$ (so that $\Lambda' \parallel \Lambda$ by Theorem 4.12) and assume (to reach a contradiction) that there is another line $\Lambda''$ parallel to $\Lambda$ that includes $P$. Define points shown in Figure 6.11 as follows.

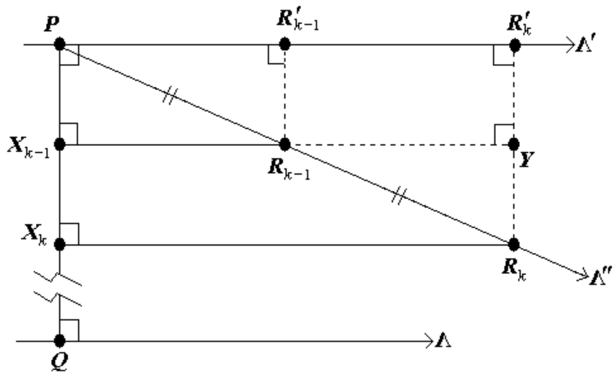



Figure 6.11:

• Let $R_0$ be a point on $\Lambda''$ on the same side of $\Lambda'$ as $Q$.

• Having chosen $R_0$, find the points $R_1, R_2, R_3, \ldots$ so that $P * R_{k-1} * R_k$ and $|PR_k| = 2|PR_{k-1}|$ for all $k$. (This means that $R_k$ is on $\overrightarrow{PR_0}$ twice as far out from $P$ as is $R_{k-1}$.)

- For each $k \geq 0$ let $X_k$ be the point on $\overleftrightarrow{PQ}$ so that $R_kX_k \perp \overleftrightarrow{PQ}$.

We will prove below that $|PX_k| = 2|PX_{k-1}|$ so that $|PX_k| = 2^k|PX_0|$ for each $k \geq 0$. Once this is established we will have a contradiction just as Proclus envisioned:

- Clearly there is an integer $k$ so that $|PX_k| = 2^k|PX_0| > |PQ|$.
- It's then easy to see that $P$ and $X_k$ are on opposite sides of $\Lambda$.
- Then since $R_kX_k \parallel \Lambda$ (Theorem 4.12 again), $P$ and $R_k$ are on opposite sides of $\Lambda$.
- This means that $\Lambda''$ contains points on both sides of $\Lambda$ and so must meet $\Lambda$, a contradiction since $\Lambda$ and $\Lambda'$ were assumed parallel!

All that remains, then, is to demonstrate our claim that $|PX_k| = 2|PX_{k-1}|$. Remember that we are allowed to use our assumption that all convex quadrilaterals have zero defect.

- Let $R'_{k-1}$ and $R'_k$ be the points of $\Lambda'$ so that $R_{k-1}R'_{k-1}$ and $R_kR'_k$ are each perpendicular to $\Lambda'$.
- Also let $Y$ be the point on $\overleftrightarrow{R_kR'_k}$ so that $R_{k-1}Y \perp \overleftrightarrow{R_kR'_k}$ (as in Figure 6.11).
- The two quadrilaterals $PR'_{k-1}R_{k-1}X_{k-1}$ and $R'_{k-1}R'_kYR_{k-1}$ each have three right angles, so by our assumption (statement **G**) angles $\angle R'_{k-1}R_{k-1}X_{k-1}$ and $\angle R'_{k-1}R_{k-1}Y$ must also be right angles.
- This means that $X_{k-1}$, $R_{k-1}$, and $Y$ are collinear, so $\angle X_{k-1}R_{k-1}P$ and $\angle YR_{k-1}R_k$ are vertical angles and are thus congruent.
- By the SAA criterion we then have $\triangle X_{k-1}R_{k-1}P \cong \triangle YR_{k-1}R_k$ so $|PX_{k-1}| = |R_kY|$.
- Now $X_{k-1}X_kR_kY$ is a quadrilateral with three right angles, so again (by our assumption that statement **G** is true) it must be a rectangle.
- Thus $X_{k-1}X_k$ and $YR_k$, being opposite sides of a rectangle, are congruent (see Exercise 6.14). So $|X_{k-1}X_k| = |YR_k| = |PX_{k-1}|$.
- Finally, then, we have $|PX_k| = |PX_{k-1}| + |X_{k-1}X_k| = 2|PX_{k-1}|$.

At last the proof is complete! $\square$

### Exercises

**6.11**. Prove that **D** $\implies$ **E** by doing the following.

- Beginning with a triangle of zero defect, show that there exists a right triangle with zero defect.
- From a right triangle with zero defect, construct a rectangle.

**6.12**. Fill in the missing detail to the proof that **E** $\implies$ **F** by showing that if $ABCD$ is a rectangle then there is a rectangle with side lengths $2|AB|$ and $2|BC|$.

**6.13**. Prove that **F** $\implies$ **G** by using Theorem 6.4.

**6.14**. By the Parallelogram Theorem every rectangle is a parallelogram and every parallelogram has opposite sides of equal length. But the Parallelogram Theorem is not a neutral geometry theorem, so we could not use this in the proof of **G** $\implies$ **B** as our justification that opposite sides of a rectangle are congruent. To fill this gap, prove in neutral geometry that if a rectangle $ABCD$ exists then $AB \cong CD$.

**6.15**. Prove directly that **G** $\implies$ **E** by considering a Saccheri quadrilateral.

## D. Geometry's New Universe

In this section we will finally introduce some basic facts of hyperbolic geometry. Actually, they are for the most part already proved and waiting for us in the form of Theorem 6.7, just as they were for Saccheri and Lambert. Though these two men failed to see the possibility of the new geometry, three men ultimately did about a century later. We begin the section with more of our historical survey, covering the stories of these three men as well as the last of those who clung to the hope of proving the parallel postulate.

#### Legendre

To a mathematician who believed staunchly that the parallel postulate should be a theorem and not an axiom, Theorem 6.7 just provided a bigger target. Instead of trying to prove Euclid's fifth axiom (statement A from the list) directly, there are now several other statements for which a successful proof would accomplish the same goal.

This was exactly the attitude of Adrien Marie Legendre (1752 - 1833 AD), one of the leading mathematicians in the French school. Legendre wrote a text called *Elements de Géométrie* that dominated the instruction of geometry for a century. In a dozen editions of the text published during his lifetime, Legendre included several different "proofs" of statements B and F from Theorem 6.7. Like Saccheri before him, Legendre successfully proved the defect of a triangle cannot be negative (it is his name now attached to the Saccheri-Legendre Theorem), and he believed that he had given several different proofs that the defect could also not be positive. Each of these demonstrations was criticized by his contemporaries as being circular in their logic (which they of course were), but Legendre stubbornly insisted on their correctness. He attributed the criticism and lack of acceptance with which his arguments were met to their difficulty.

> These considerations ... leave little hope of obtaining a proof of the theoory of parallels or the theorem on the sum of the three angles of the triangle, by means as simple as those which one uses for proving the other propositions of the *Elements.*
>
> It is no less certain that the theorem on the sum of the three angles of the triangle must be regarded as one of those fundamental truths which is impossible to dispute, and which are an enduring example of mathematical certitude, which one continually pursues and which one obtains only with great difficulty in the other branches of human knowledge. Without doubt one must attribute to the imperfection of common language and to the difficulty of giving a good definition of the straight line the little success which geometers have had until now when they have wanted to deduce this theorem only from the notions of equality of triangles contained in the first book of the *Elements.*

These words are taken from a compilation of Legendre's various "proofs" of the parallel postulate that he published in 1833, the year of his death. What he did not know when he composed these sentences is that elsewhere in the world the tides had begun to turn. Three men in separate places had encountered the parallel postulate, struggled to prove it, and come away believing in the possibility of a geometry in which it did not hold true.

**Gauss**

Carl Friedrich Gauss (1777 - 1855 AD) was certainly one of the greatest mathematicians the world has ever seen. His contributions span the entire breadth of the subject; in fact, he can be credited with helping to found several branches of modern mathematics. Such a mind as his could not have been kept from the problem of the parallel postulate! Gauss began working on the problem as early as 1792 and found his efforts utterly frustrating. He wrote to a friend, Wolfgang Bolyai, who was also working on the problem, that his work compelled him to "doubt the truth of geometry itself."

After years of fermenting, Gauss' ideas gradually converged on the notion that the parallel postulate was *not* a necessary belief, and that disbelieving it would never lead to a contradiction, but rather to a new geometry. He wrote (in another private letter) in 1824:

> The assumption that the sum of the three angles is less than 180° leads to a curious geometry, quite different from ours, but thoroughly consistent, which I have developed to my entire satisfaction.

Gauss called this new curious geometry *anti-Euclidean*, or *non-Euclidean*. But though (by his own words) he was convinced of the correctness of this new geometry, he never announced his findings publicly. In fact, he warned those with whom he *did* discuss it to not divulge it to others.

The reason for this is a matter of some mystery. It seems hard to believe that someone of Gauss' caliber would fail to see the import of this discovery. Yet it is also hard to understand how, knowing its importance, someone who devoted their life to mathematical research (as Gauss had done) would not champion the idea enthusiastically. Most likely, the answer is that Gauss wanted to avoid controversy. He knew that the primacy of Euclidean geometry was firmly entrenched in the thinking of his day, and he suspected that most of his contemporaries would scoff at the suggestion that other geometries existed. Perhaps he concluded that even his reputation could not prevent uproar over something as unthinkable as a non-Euclidean geometry.

Indeed, it may be argued that Gauss' fear was well founded. He lived to see others make the same discoveries and announce them to the world, only to be ignored or dismissed. Perhaps a vote of confidence from Gauss would have quickened acceptance of the idea; we'll never know that for certain. We do know that it was only after Gauss' death, when his work on the subject was made known to the mathematical world at large, that the notion of "non-Euclidean" geometry was taken seriously.

### Bolyai

Janos Bolyai (1802 - 1860 AD) was the son of Wolfgang Bolyai, the aforementioned friend of Gauss. When young Janos showed interest in the problem of the parallel postulate, his father strongly discouraged the effort. He warned Janos to turn back from the problem and not become entangled in what he considered from his own experience to be a vain and fruitless search.

Undaunted by this warning, Janos persisted in his investigations. Eventually he wrote back to his father, reporting proudly that his work was progressing better than he had hoped, and that "out of nothing" he had created a "strange new universe." Finally, in about 1829 his work brought him to the same conclusion Gauss had reached. He wrote down his discoveries and sent them off to his father.

Wolfgang, proud of his son's accomplishment and anxious that it receive quick publication, included the work as an appendix to one of his own treatises. He sent a copy to his friend Gauss for appraisal (unaware, of course, that Gauss had already worked out the same material), thoroughly expecting praise for the bold discoveries his son had made.

But Gauss' reaction was lukewarm, for, while he praised the manner in which Janos had carried out the task, he noted that he could not praise too highly something he himself had previously done. Gauss never gave the work the public praise that both of the Bolyais hoped he would. The experience was so discouraging to Janos that he never again published anything.

### Lobachevsky

Gauss was German and Bolyai Hungarian. Neither was aware that in another corner of the world a young Russian mathematician by the name of Nikolai Ivanovich Lobachevsky (1793 - 1856 AD) was arriving at the same conclusion regarding the parallel postulate. Lobachevsky developed what he called "imaginary geometry" by replacing the parallel postulate with the assumption that through a point $P$ not on a line $\Lambda$ there passed more than one line parallel to $\Lambda$. He worked out from this new axiom system the same results Gauss had discovered previously and Bolyai was then working on.

Lobachevsky was the first to publish his work (in 1829), though it was published only in Russian and did not receive any significant attention. When a German translation finally came to Gauss' attention in 1840, Gauss gave it private praise, but (as he had done with Bolyai's work) did not champion it publicly.

Lobachevsky and Bolyai are both somewhat sad figures in our story. Though they were the first to publicly proclaim the existence of an alternative to Euclidean geometry, neither's work was much appreciated in their lifetimes.

### The new geometry

The body of theorems Gauss called "non-Euclidean geometry" and Lobachevsky called "imaginary geometry" is today known as *hyperbolic geometry.* It is the geometry that results from replacing the Euclidean Parallel Axiom with its negation. Specifically, for the remainder of this chapter we use as our axioms all of the axioms of neutral geometry together with the following.

> **The Hyperbolic Parallel Axiom.** There exists a line $\Lambda$ and a point $P$ not on $\Lambda$ so that through $P$ there are at least two distinct lines parallel to $\Lambda$.

Note that this is strictly the negation of the Euclidean Parallel Axiom, in that it stipulates only a single counterexample to there being a unique parallel. The reason for this wording is to make clear that Euclidean geometry and hyperbolic geometry are the *only* possible extensions of neutral geometry.[4] In Section F we will prove that the multiple parallels assumed at one point by the above axiom in fact exist *generically* (at every point). Specifically, we will prove that given *any* line $\Lambda$ and *any* point $P$ not on $\Lambda$ there are *infinitely many* lines through $P$ parallel to $\Lambda$.

For the remainder of this section, though, we will easily establish some of the most remarkable characteristics of hyperbolic geometry by using Theorem 6.7. The fact that statements C through G follow from the Euclidean Parallel Axiom (statement B) was proved in Chapter 5. So the new content to Theorem 6.7 was the converse of this: *if any of statements C through G are correct then the Euclidean Parallel Axiom is also satisfied.* Legendre saw this as an invitation to attempt a proof of one of C through G and thus (he hoped) to prove the parallel postulate. Gauss, Bolyai, and Lobachevsky, on the other hand, eventually realized that its contrapositive was an invitation to a new geometry: *if in some geometry the Euclidean Parallel Axiom is false (but the axioms of neutral geometry hold) then each of the statements C through G must be false also*!

[4] Playfair's postulate is an undecideable statement in neutral geometry – we are free to assume that it is either true or false since it cannot be proved to be either. If we assume it is true, we get Euclidean geometry – if we assume it is false, we get hyperbolic geometry.

Since the Hyperbolic Parallel Axiom is exactly the negation of the Euclidean Parallel Axiom, the negations of statements C through G are now theorems in our new geometry. For instance, negating statement D (and recalling Theomem 6.5) we get the following striking theorem.

**THEOREM 6.8.** *The defect of every triangle is positive. (That is, the sum of the measures of the three angles in a triangle is less than 180.)* □

The negation of statement E would be simply that there exist no rectangles in hyperbolic geometry. However, more than this is actually true. We leave it as an easy exercise to fill in the proof for the following fact.

**THEOREM 6.9.** *The defect of every convex quadrilateral is positive. (That is, the sum of the measures of the four angles in a convex quadrilateral is less than 360.)*

The negation of statement C would be that there exists *at least one example* of parallel lines $\Lambda$ and $\overleftrightarrow{PQ}$ in which the distance from $P$ to $\Lambda$ is not equal to the distance from $Q$ to $\Lambda$. It would be false to claim that given *arbitrary* parallel lines $\Lambda$ and $\overleftrightarrow{PQ}$ these distances can never be equal, for it could be that $P$ and $Q$ are the summit vertices of a Saccheri quadrilateral with base on $\Lambda$. There is, however, a related universal statement that holds as a theorem in hyperbolic geometry. Its proof is left as Exercise 6.19.

**THEOREM 6.10.** *If $\Lambda$ and $\Lambda'$ are any two parallel lines and $d > 0$ is any distance then there are at most two points of $\Lambda'$ at distance $d$ from $\Lambda$.*

Next comes perhaps the single most striking fact from hyperbolic geometry: the "angle-angle-angle" triangle congruence criterion. Surprisingly, in hyperbolic geometry *there do not exist similar but noncongruent triangles*! So, unlike Euclidean geometry where length is relative and two figures may "look the same" even though they are on different scales, in hyperbolic geometry we cannot change the scale of a figure without altering its angles as well as its lengths. It was this very property that caused Lambert to ultimately reject the idea of a geometry without the parallel postulate (recall his comments on p. 199). Though he at least toyed with the notion, he ultimately could not believe in a universe in which distance was as absolute as angle.

**THEOREM 6.11.** *Let $ABC$ and $DEF$ be triangles such that $\angle A \cong \angle D$, $\angle B \cong \angle E$, and $\angle C \cong \angle F$. Then $\triangle ABC \cong \triangle DEF$.*

*Proof:* Let the two triangles be as stated and assume, to reach a contradiction, that they are not congruent. It must be, then, that the corresponding side lengths are not equal. Assume without loss of generality that $|AB| > |DE|$. Refer to Figure 6.12 in the following steps.

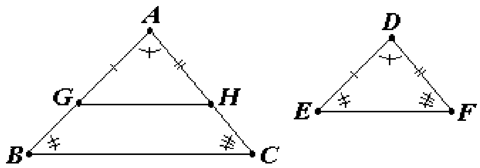



Figure 6.12:

- Let $G$ be the point on $AB$ so that $|AG| = |DE|$.
- Let $H$ be the point on $\overrightarrow{AC}$ so that $|AH| = |DF|$.
- By Axiom SAS, $\triangle AGH \cong \triangle DEF$.
- So, $\angle AGH \cong \angle DEF$ and $\angle AHG \cong \angle DFE$.
- By Theorem 4.12, $GH \parallel \overleftrightarrow{BC}$.
- Thus, H must be on the same side of $\overleftrightarrow{BC}$ as $G$ (which is the same side as $A$), so $H$ is between $A$ and $C$.
- So, $m\angle HGB = 180 - m\angle HGA = 180 - m\angle B$ and $m\angle GHC = 180 - m\angle GHA = 180 - m\angle C$.
- But then the quadrilateral $BCHG$ has defect zero, a contradiction to Theorem 6.9. This completes the proof. $\square$

**Riemann and Elliptic Geometry**

The counterintuitive nature of hyperbolic geometry's theorems makes it somewhat understandable why mathematicians in the 19th century were reluctant to accept it. But viewed another way, perhaps to believe in non-Euclidean geometries should not have been so difficult, for a clear example of a non-Euclidean geometry was right under their noses: the geometry on the surface of a sphere.

Around the same time that hyperbolic geometry was finally catching on, Georg Friedrich Bernhard Riemann (1826-1866) developed a formal generalization of spherical geometry known now as *elliptic geometry*. To best see the place of this third alternative relative to Euclidean and hyperbolic geometries, recall that Saccheri had reduced his investigation of the Saccheri quadrilateral to three cases: the summit angles were either acute, obtuse, or right angles (see p.198).

- We now know that the *right angle hypothesis* (the assumption that the summit angles are right angles) leads to Euclidean geometry, for the existence of a single rectangle is equivalent to the parallel postulate. Under this hypothesis, all triangles have angle sums equal to 180°.
- We also now know that the *acute angle hypothesis* leads to hyperbolic geometry, for it gives us the existence of a quadrilateral with positive defect. Under this hypothesis, all triangles have angle sums of *less* than 180°.
- The *obtuse angle hypothesis* would mean the existence of a quadrilateral with negative defect and thus triangles with angle sums *greater* than 180°. Saccheri had been able to show that this possibility cannot occur under the axioms of neutral geometry (recall the Saccheri-Legendre Theorem).

But we have seen (see Figure 3.2) that triangles in spherical geometry *can* have angle sums greater than 180°. Riemann's elliptic geometry was the formal resolution of this apparent difficulty. He showed that by altering parts of the neutral axioms (such as the assumption in our Axiom L2 that lines have unbounded length) that there was no longer a contradiction in the obtuse angle hypothesis. Thus all three possibilities for the summit angles of a Saccheri quadrilateral (or the angle sums of a triangle) have geometric realizations. The angle sum of a triangle is *exactly* 180° in Euclidean geometry, *less than* 180° in hyperbolic geometry, and *greater than* 180° in elliptic geometry.

## Exercises

**6.16**. True or False? (Questions for discussion)

**(a)** In hyperbolic geometry every triangle has positive defect.

**(b)** In hyperbolic geometry the summit angles of a Saccheri quadrilateral cannot be congruent.

**(c)** Suppose that $ABCD$ is a Lambert quadrilateral in hyperbolic geometry, with right angles at $A$, $B$, and $C$. Then $|AD|$ cannot equal $|BC|$.

**(d)** In neutral geometry we cannot prove whether or not Lambert quadrilaterals exist.

**(e)** In Euclidean geometry, every Lambert quadrilateral is a rectangle.

**(f)** Not all neutral geometry theorems are true in elliptic geometry.

**(g)** In elliptic geometry, quadrilaterals may have negative defects.

**6.17**. Prove Theorem 6.9.

**6.18**. Prove that in hyperbolic geometry there does not exist a pair of distinct parallel lines with more than one common perpendicular.

**6.19**. Prove Theorem 6.10 by using facts about Saccheri quadrilaterals.

**6.20**. Describe how to find a Saccheri quadrilateral in the model for spherical geometry. What can you say about its summit angles?

## E. A Model For Hyperbolic Geometry

Saccheri, Lambert, Legendre, Gauss, Bolyai, and Lobachevsky all (in their own way) investigated the consequences of the Hyperbolic Parallel Axiom. The first three were determined to find a contradiction within those consequences, while the latter three ultimately formed the opinion that there was no such contradiction. But how can we know for certain? How can we decide if the system of axioms for hyperbolic geometry is consistent? Theorem 3.1 tells us that a model is just what is called for, and several such models were proposed in the latter part of the 19th century. These models proved that hyperbolic geometry is *at least as consistent as Euclidean geometry* – as much as we can expect to know in light of Gödel's Theorem.

By Gödel's Theorem we cannot know for certain that the axioms of the real numbers are consistent. Since our geometric axioms rely on these notions, we ultimately cannot determine the consistency of Euclidean geometry. We have a model for Euclidean geometry, however, which proves that Euclidean geometry is consistent *provided that the axioms of set theory and the real numbers are consistent*. The reason is that our usual model for the Euclidean plane "lives" within the universe of these other systems – the Cartesian plane model is after all an object constructed from the real numbers. As long as the foundations of set theory and the real numbers are free of contradictions, our model is on firm footing and its existence proves that Euclidean geometry is just as consistent. Because of this we say that Euclidean geometry is *consistent relative to* the axiom systems for set theory and the real numbers.

In 1868 the Italian mathematician Eugenio Beltrami (1835 - 1900 AD) proved that hyperbolic geometry is consistent relative to Euclidean geometry by exhibiting a model for hyperbolic geometry that "lives" within Euclidean geometry.

That is, he found a way of associating subsets of the Euclidean plane with "hyperbolic lines" in a way which obeys all of the axioms for hyperbolic geometry. Thus, a valid model for hyperbolic geometry exists (showing that hyperbolic geometry is consistent) provided that Euclidean geometry is consistent!

So, if you have found it hard to swallow some of the theorems of hyperbolic geometry – if you, like Saccheri, suspect that somewhere in those strange-sounding consequences there *must* be a contradiction – consider the following chain of reasoning. Suppose that tomorrow you were to find an indisputable contradiction within hyperbolic geometry. Then:

- You would know for certain that Euclidean geometry also contains contradictions! For because of Beltrami's model it cannot be that Euclidean geometry is consistent while hyperbolic geometry is not.

- But the situation is even worse than this: the usual model of Euclidean geometry shows that set theory and the real numbers cannot be consistent if Euclidean geometry is inconsistent. So your little discovery would have proved the existence of flaws in the very foundations of mathematics!

The pragmatic view of the situation is this: in order to *do* mathematics, we would like to at least *believe* in the consistency of its basis, though we know we can never *prove* it. But the consistency of this basis implies the consistency of Euclidean geometry which in turn implies the consistency of hyperbolic geometry. In a very real sense, mathematicians have no choice but to accept the new universe of Gauss, Bolyai, and Lobachevsky.

Beltrami's model had another consequence. It was the first model known to satisfy all the axioms for neutral geometry but not the Euclidean Parallel Axiom. So, by the principal of Theorem 3.2, the question at the heart of our two-millennium-long story at last had an answer: the parallel postulate is *not* a theorem of neutral geometry, and the search for a proof within neutral geometry had been futile (though fruitful in unexpected ways). Beltrami had proved once and for all that the parallel postulate is independent of the neutral geometry axioms.

Other models for hyperbolic geometry would soon follow after Beltrami's. The French mathematician Henri Poincaré (1854 - 1912 AD) posed two such models, and we will now briefly introduce one of them.

### The Poincaré halfplane model

Before introducing the model, let's remind ourselves of a few basics. A model for a geometry is some system of associating objects with the lines of the geometry in such a way that the axioms are obeyed. The objects to which lines are associated need not "look" in any way like a straight line, for the term "line" in a geometry is just an abstraction and should carry no connotation. In the model we are about to introduce the "lines" are certain subsets of the Euclidean plane, most of which do not look at all like the lines in the usual model for Euclidean geometry. To avoid confusion as much as possible we will refer to them as "H-lines", and will likewise place an "H" in front of any term that should be distinguished from its counterpart in the Euclidean model.

The universe of points in our model is the set $\Pi$ of all points $(x, y)$ in the Cartesian plane for which $y > 0$. In other words, it is the "upper halfplane" of points strictly above the $x$-axis. The *H-lines* of the model are the subsets of $\Pi$ meeting one of the following descriptions:

- intersections of $\Pi$ with vertical lines $x = c$ ($c$ any constant).
- intersections of $\Pi$ with circles $(x - c)^2 + y^2 = r^2$ ($c$ and $r$ any constants) centered on the $x - axis$.

Figure 6.13 shows three H-lines through a point $P$ in this model. H-segments are then either segments of vertical lines or arcs of circles centered on the $x$-axis.

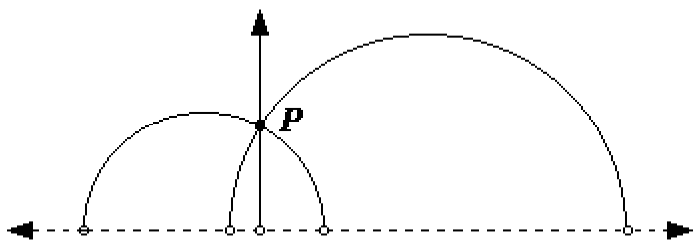



Figure 6.13: Some H-lines through the point $P$

It is an easy exercise (see Exercise 6.21) to prove that this model satisfies Axiom L1: given any two points in $\Pi$ there is a unique H-line through them both. We illustrate the procedure with an example.

EXAMPLE **6.12.** *Consider the points $P = (1, 4)$ and $Q = (7, 2)$. Find the equation of the H-line $\overleftrightarrow{PQ}$.*

*Solution:* We are obviously looking for a circle passing through $P$ and $Q$ with center on the $x$-axis. The center of this circle must then be the point at which the $x$-axis intersects the perpendicular bisector of the Euclidean segment $PQ$ (see Figure 6.14). The midpoint of this Euclidean segment is (4,3) and its slope is $-1/3$. The perpendicular slope is then 3, so it is easy to see that the center of the circle is the point $(3, 0)$. It's radius is then the distance from $(3, 0)$ to $(7, 2)$ which is $\sqrt{20}$. So, the equation of the H-line $\overleftrightarrow{PQ}$ is $(x-3)^2 + y^2 = 20$, or $y = \sqrt{20 - (x-3)^2}$. (Note that only the positive square root is needed since the H-line exists only in the upper halfplane.) $\square$

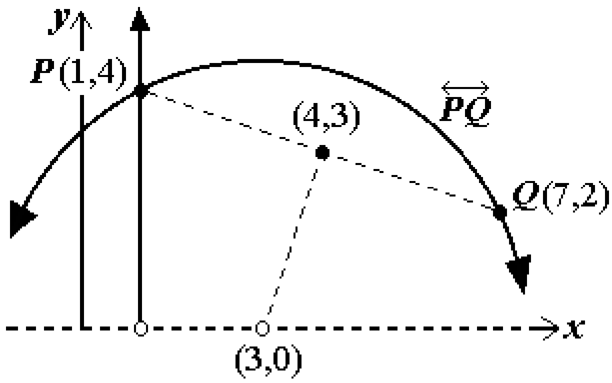


Figure 6.14: Deriving the equation of an H-line

It is a fact that this model satisfies all of our neutral geometry axioms. Some (for instance, Axiom L4 – see Exercise 6.22) can be checked without too much difficulty. But to check the axioms that involve lengths we would need to know how distances are measured in this model. Clearly, the H-length of an H-segment cannot just be its Euclidean length, for then some of our H-lines would have finite total length, contrary to Axiom L2. Instead, H-length is measured using a formula that "compresses" distances near the $x$-axis: points with small $y$-coordinate that appear to be close together may in fact be far apart in the H-metric. Though we will not derive it here, we can easily state a valid distance formula for this model:

> The H-distance between points $(x_1, y_1)$ and $(x_2, y_2)$ in the upper halfplane model is
>
> $$d_H((x_1, y_1), (x_2, y_2)) = \text{Arccosh}\left(1 + \frac{(x_1-x_2)^2+(y_1-y_2)^2}{2y_1y_2}\right).$$

where "Arccosh" is the inverse hyperbolic cosine function given by $\text{Arccosh}(x) = \ln(x + \sqrt{x^2-1})$.

To formally prove that this distance function satisfies the requirements of the neutral axioms requires long and complex calculations and a good deal of background material. We won't develop those proofs here. We will, however, make occasional use of the above formula in working examples.

Along with a distance function, the neutral axioms require us to provide our model with a notion of angle measure. And while distances are quite distorted by pictures using the halfplane model, measures of H-angles are what you would probably expect. The measure of the H-angle formed by two intersecting H-lines is determined using the (Euclidean) tangent lines to the two H-lines. The slopes of these tangent lines are easily computed by differentiation, and it is not difficult to see that this results in an angle measure consistent with our axioms.

Example **6.13.** *In addition to points $P$ and $Q$ from the previous example, let $R$ be the point $(1,5)$. Find $m\angle QPR$.*

*Solution:* Differentiating the equation for $\overleftrightarrow{PQ}$ from Example 6.13 we have $\frac{dy}{dx} = \frac{3-x}{\sqrt{20-(x-3)^2}}$, so the tangent slope of $\overleftrightarrow{PQ}$ at $P$ is $\frac{dy}{dx}\Big|_{x=1} = 1/2$. The H-line $\overleftrightarrow{PR}$ is the vertical ray $x = 1$. So $m\angle QPR$ will be the measure of the Euclidean angle shown in Figure 6.15. This is easily computed as Arctan(2), or about 63.435 degrees. □

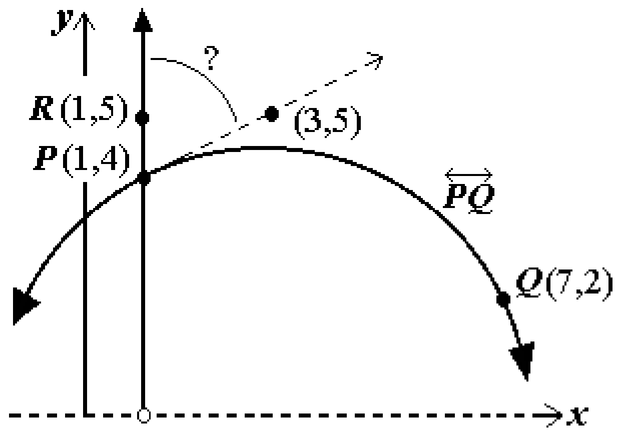


Figure 6.15: Calculating H-angle measure

To complete our discussion of the axioms, we note that it is abundantly evident that the Poincaré halfplane model does *not* satisfy the Euclidean Parallel Axiom! Figure 6.16 depicts the infinite family of parallels through a point $P$ not on a line $\Lambda$. Note that this family fills out a region of the plane, and that two of them play a special role in that they bound this region and seem to converge to $\Lambda$ as they approach the $x$-axis. (It is important to remember that the points on the $x$-axis are *not* part of the model, so these lines really *are* parallel to $\Lambda$ even though they seem to touch $\Lambda$ at the $x$-axis.) This is not just an artifact of our model. The final sections of this chapter will be devoted to proving that parallel lines in hyperbolic geometry behave exactly as suggested by this figure.

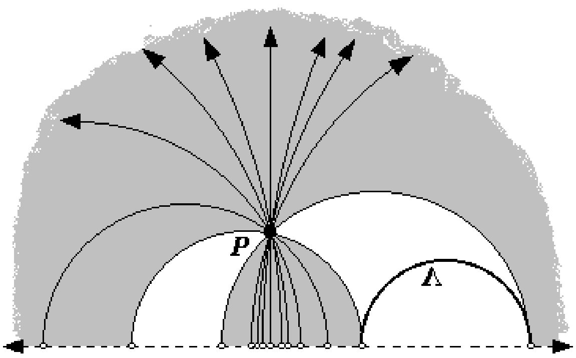



Figure 6.16: The H-parallels to $\Lambda$ through $P$

**Proving what theorems are *not* in neutral geometry**

We'll now construct counterexamples to a few Euclidean geometry theorems. We stated when proving some of these theorems that the Euclidean Parallel Axiom was needed for their proof. You probably believed us at the time, but you might have wondered how one would *know* that a theorem cannot be proved without a certain axiom. The answer to this question lies in the power of models. We now have a model satisfying the axioms of neutral geometry but not the Euclidean Parallel Axiom. So, if a counterexample to a theorem exists in this model we are assured that it cannot be proved by the neutral geometry axioms alone.

EXAMPLE **6.14.** *Show that without the Euclidean Parallel Axiom we cannot prove that alternate interior angles in a transversal of parallel lines are congruent (Theorem 5.2).*

*Solution:* You should recall that we proved the converse of this theorem (namely that if the alternate interior angles of a transversal are congruent then the lines being transversed are parallel) in neutral geometry, so this converse should (and does!) hold true in the Poincaré halfplane model. But Figure 6.17 depicts an H-line $\Lambda_0$ transversing parallel H-lines $\Lambda_1$ and $\Lambda_2$ so that the alternate interior angles of the transversal are obviously not congruent. (One is a right angle and the other is clearly not.) □

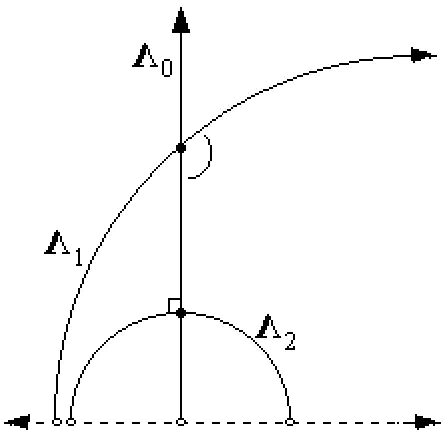



Figure 6.17: Parallel lines not giving congruent alternate interior angles

EXAMPLE **6.15.** *Show that without the Euclidean Parallel Axiom we cannot prove that every set of three distinct noncollinear points will determine a unique circle.*

*Solution:* Let $P$, $Q$, and $R$ be the points $(0,1)$, $(1,1)$, and $(2,1)$ respectively (see Exercise 6.29). Then the perpendicular bisectors of the H-segments $PQ$ and $QR$ are $x = 1/2$ and $x = 3/2$. (Since we have not defined length, you'll need to take our word that these H-lines bisect the H-segments in question.) So all points H-equidistant from $P$ and $Q$ are on $x = 1/2$ and all points H-equidistant from $Q$ and $R$ are on $x = 3/2$, meaning that there is *no* point H-equidistant from $P$, $Q$, *and* $R$! Thus, no H-circle passes through all three, for it could have no center. (With this example in mind, it might be instructive to again review the proof to Fact 5.7 to see where the Euclidean Parallel Axiom is used.) □

We will leave it to you (see Exercises 6.27 and 6.28) to supply counterexamples to some other familiar Euclidean geometry theorems.

**The role of models and diagrams (again)**

We close this section with a thought on the appropriate use of models and diagrams in geometry, specifically in hyperbolic geometry. As we have seen, the very existence of models for hyperbolic geometry has profound consequences for the consistency of the subject matter. And we have seen that the model itself is useful for proving what Euclidean geometry theorems are not part of neutral geometry. We even used the model to predict behavior of lines in hyperbolic geometry, and will now devote the rest of this chapter to proving those predictions correct.

But as we return to proofs from the axioms, we will also return to our old system of diagrams in which we tried to indicate lines as straight figures, imitating the model for Euclidean geometry. For even though this system has some disadvantages, it is probably better suited than is the Poincaré halfplane model for the task of illustrating our proofs. (We won't shelve the halfplane model completely. We'll continue to use it for motivation and in examples and exercises.) Indeed, we should never be slaves to any one depiction or model of geometry, but rather should feel free to adopt whatever serves the present purpose best. As we proceed with the arguments in the following section we invite you to draw your own figures using the model we just learned. It will help cement your understanding of the model, and may bring some better understanding of the proofs.

## Exercises

**6.21**. Prove that the Poincaré halfplane model satisfies Axiom L1.

**6.22**. Prove that the Poincaré halfplane model satisfies Axiom L4. (For H-lines that are Euclidean semicircles in the model it may appear that one of their sides is not convex. But remember that H-segments are not straight unless they are vertical, so a set may be H-convex even if it is not convex in a Euclidean sense.)

**6.23**. Find the equation of the H-line $\overleftrightarrow{PQ}$ if

**(a)** $P$ is $(3, 1)$ and $Q$ is $(7, 1)$.

**(b)** $P$ is $(-2, 3)$ and $Q$ is $(-2, 6)$.

**(c)** $P$ is $(-2, 5)$ and $Q$ is $(0, 1)$.

**(d)** $P$ is $(1, 4)$ and $Q$ is $(2, 1)$.

**(e)** $P$ is $(-5, 1)$ and $Q$ is $(0, 7)$.

**(f)** $P$ is $(2, 1)$ and $Q$ is $(3, 10)$.

**6.24**. Find $m\angle BAC$ if

**(a)** $A$ is $(1, 1)$, $B$ is $(2, 1)$, and $C$ is $(3, 1)$

**(b)** $A$ is $(3, 2)$, $B$ is $(3, 3)$, and $C$ is $(1, 1)$

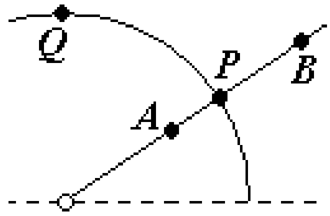
**(c)** $A$ is $(0, 3)$, $B$ is $(4, 3)$, and $C$ is $(6, 3)$

**(d)** $A$ is $(-1, 4)$, $B$ is $(5, 4)$, and $C$ is $(0, 1)$

**6.25**. In each instance, find a point $Q$ on the H-line $\Lambda$ so that $PQ \perp \Lambda$, then calculate $|PQ| = d_H(P, Q)$ using the formula for H-distance on p.216.

**(a)** $P$ is $(2, 5)$ and $\Lambda$ is given by $y = \sqrt{4x - x^2}$.

**(b)** $P$ is $(-4, 3)$ and $\Lambda$ is given by $x = 0$ $(y > 0)$.




**6.26**. In the halfplane model, let $A$ and $B$ be the points $(\frac{\sqrt{3}}{3}, \frac{1}{3})$ and $(\frac{3\sqrt{3}}{4}, \frac{3}{4})$, and let $P$ and $Q$ be the points $(\frac{\sqrt{3}}{2}, \frac{1}{2})$ and $(0, 1)$, respectively (as in the figure at right).

**(a)** Show that the H-line $\overleftrightarrow{PQ}$ is the perpendicular bisector of the H-segment $AB$.

**(b)** Show that $d_H(A, P) + d_H(P, B) \neq d_H(A, B)$. How can this be?

**6.27**. Let $\varepsilon > 0$ be given. Show how we may construct (in the Poincaré halfplane model) an H-triangle and an H-convex quadrilateral, each with all angles having measure less than $\varepsilon$.

**6.28**. Give a counterexample in the Poincaré halfplane model to Corollary 5.3.

**6.29**. Explain why the three points in Example 6.15 are not H-collinear.

**6.30**. In Example 6.15 we demonstrated that three H-noncollinear points may fail to determine an H-circle. But does the uniqueness portion of Theorem 5.7 still hold true? That is, if there *is* an H-circle through points $P$, $Q$, and $R$, is there only one such H-circle? Explain and justify your answer.

# F. Angle of Parallelism

The rest of this chapter covers difficult material. We will explore the workings of hyperbolic geometry in greater detail, and following some of the proofs in these explorations will require considerable concentration. Expect to go slowly working through these pages.

Let $\Lambda$ be a line, and $P$ a point not on $\Lambda$. Let $\mathcal{P}(\Lambda, P)$ denote the set of all lines through $P$ that are parallel to $\Lambda$. In this section we will study $\mathcal{P}(\Lambda, P)$ and prove that it has the properties suggested by our model in Figure 6.16. That figure suggests the following:

**(i)** $\mathcal{P}(\Lambda, P)$ is always an infinite set, and

**(ii)** $\mathcal{P}(\Lambda, P)$ is exactly the set of lines lying within an angle formed by two of its members.

Note that we are not guaranteed either of these behaviors simply by the axioms. The Hyperbolic Parallel Axiom is certainly related to **(i)**, but it states only that $\mathcal{P}(\Lambda, P)$ contains at least *two* elements for *some* choice of $\Lambda$ and $P$, not *infinitely many* elements for *all* choices of $\Lambda$ and $P$. But below we will see that these predicted properties are in fact consequences of our new axiom. Specifically, we will prove the following facts about $\mathcal{P}(\Lambda, P)$.

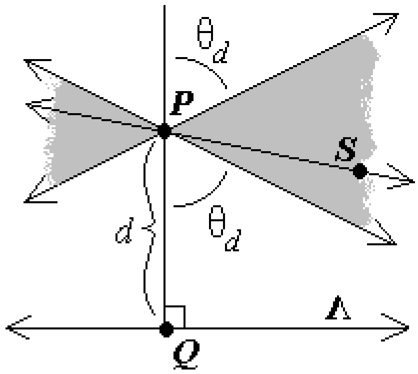



Figure 6.18: The statement of Theorem 6.16

THEOREM **6.16.** *For each distance $d > 0$ there is an angle measure $\theta_d$ satisfying the following properties.*

**(a)** *If $P$ is a point not on a line $\Lambda$, $Q$ is the point on $\Lambda$ so that $PQ \perp \Lambda$, and if $|PQ| = d$, then a line $\overleftrightarrow{PS}$ through $P$ is in $\mathcal{P}(\Lambda, P)$ if and only if $\theta_d \leq m\angle QPS \leq 180 - \theta_d$ (see Figure 6.18).*

**(b)** *For each $d$ we have $0 < \theta_d < 90$.*

**(c)** *If $d < d'$ then $\theta_d > \theta_{d'}$.*

**(d)** *The value of $\theta_d$ varies continuously with $d$.*

**(e)** $\lim_{d\longrightarrow 0} \theta_d = 90$ *and* $\lim_{d\longrightarrow\infty} \theta_d = 0$

Part **(a)** of this theorem describes exactly the situation suggested by Figure 6.16 and motivates the following definitions.

DEFINITIONS. The number $\theta_d$ is traditionally called the *angle of parallelism for the distance d* (though note that in our terminology it should be called an angle measure and not an angle). If $P$, $\Lambda$, and $Q$ are as in part **(a)** of Theorem 6.16 then the lines through $P$ which make angles of measure $\theta_d$ and $180 - \theta_d$ with the ray $\overrightarrow{PQ}$ are called the *bounding parallels* for the family $\mathcal{P}(\Lambda, P)$.

Before giving a proof of Theorem 6.16 we will use some neutral geometry and calculus to give a formula for $\theta_d$.

Let $P$ be a point at distance $d > 0$ from a line $\Lambda$ and let $Q$ be the point on $\Lambda$ so that $PQ \perp \Lambda$. Choose one of the rays with endpoint $Q$ determined by $\Lambda$ and for each number $t > 0$ let $R_t$ be the point on this ray at distance $t$ from $Q$ (see Figure 6.19**(a)**). Define $\theta_d(t)$ to be the measure of the angle $\angle QPR_t$. Note that the value of $\theta_d(t)$ depends only on $d$ and $t$ (and not on $P$ or $\Lambda$). For by the side-angle-side congruence criterion, any two right triangles with sides adjacent to the right angle having lengths $d$ and $t$ will have an angle of this same measure opposite the side of length $t$. The functions $\theta_d(t)$ have the following two useful properties.

**(i)** If $t < t'$ then $\theta_d(t) < \theta_d(t')$.

**(ii)** If $d < d'$ then for each value of $t$ we have $\theta_d(t) > \theta_{d'}(t)$.

The proofs of these properties are illustrated in parts **(b)** and **(c)** of Figure 6.19.

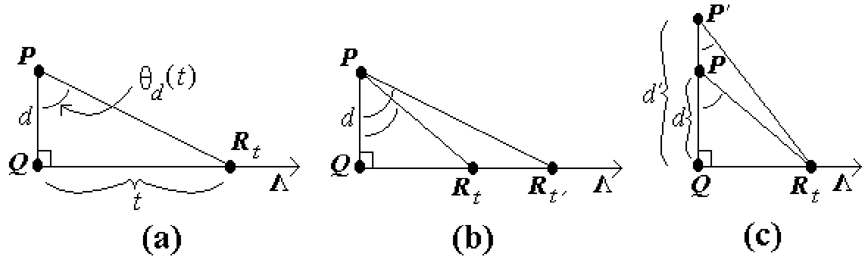



Figure 6.19: **(a)** The definition of $\theta_d(t)$, **(b)** $\theta_d(t) < \theta_d(t')$, **(c)** $\theta_d(t) > \theta_{d'}(t)$

For part **(c)** of the figure, we see that $\theta_d(t) = m\angle QPR_t > m\angle QP'R_t = \theta_{d'}(t)$ by Theorem 6.1 applied to triangle $PP'R_t$.

By property **(i)**, $\theta_d(t)$ is a strictly increasing function of $t$. But $\theta_d(t)$ is also bounded above by 90 (since the angle measure sum for $QPR_t$ cannot exceed 180 by the Saccheri-Legendre Theorem). From calculus we know that the limit of a bounded but increasing function of $t$ exists as $t \longrightarrow \infty$, which allows us to make the following definition:

$$\theta_d = \lim_{t \longrightarrow \infty} \theta_d(t) \,.$$

With this formula in hand we can now proceed with the proof of Theorem 6.16. We will show that the value of $\theta_d$ given by this formula satisfies statements **(a)**, **(b)**, **(c)**, **(d)**, and **(e)** of the theorem.

*Proof of part* **(a)***:* Let $P$ and $\Lambda$ be given and let $d$, $Q$, and the points $R_t$ be as in the discussion above. Let a line $\overleftrightarrow{PS}$ be given. We need to show that $\overleftrightarrow{PS} \parallel \Lambda$ if and only if $\theta_d \leq m\angle QPS \leq 180 - \theta_d$. If $S'$ is a point on $\overleftrightarrow{PS}$ so that $S * P * S'$ then either

- both $m\angle QPS$ and $m\angle QPS'$ are in the range $[\theta_d, 180 - \theta_d]$, or
- neither $m\angle QPS$ nor $m\angle QPS'$ is in the range $[\theta_d, 180 - \theta_d]$.

So there is no loss of generality in assuming that $m\angle QPS$ is the smaller of the two – that is, that $m\angle QPS \leq 90$. We need only prove then that $\overleftrightarrow{PS} \parallel \Lambda$ if and only if $m\angle QPS \geq \theta_d$. We do this in a contrapositive manner. By Axiom AC and Axiom C4, the inequality $m\angle QPS < \theta_d$ holds true

- if and only if $m\angle QPS < \theta_d(t) = m\angle QPR_t$ for some $t$ (recall the definition of $\theta_d$ as a limit!)
- if and only if $S$ is in the interior of $\angle QPR_t$
- if and only if $\overrightarrow{PS}$ intersects $QR_t$

(see Figure 6.20). So $m\angle QPS < \theta_d$ if and only if $\overleftrightarrow{PS}$ is not parallel to $\Lambda$, which is equivalent to statement **(a)**. □

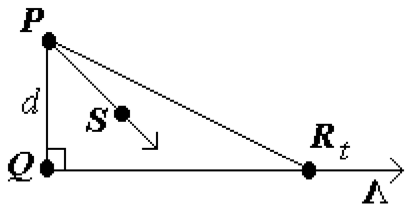


Figure 6.20:

Now everything so far has been done in neutral geometry, and so would be equally valid (but not very interesting) in Euclidean geometry – see Exercise 6.39. The truly striking fact about this theorem (and the part for which we will need to use the Hyperbolic Parallel Axiom) is part **(b)** – that $\theta_d$ is *less than* 90 for all values of $d$.

*Proof of part* **(b)***:* First note that the conclusion $\theta_d > 0$ is trivial. (Can you explain why?) So we need only show that $\theta_d < 90$ for all $d$. Assume, then, (to reach a contradiction) that $\theta_c = 90$ for some distance $c > 0$. We will first show that $\theta_{2c}$ is also 90.

- Let $P$, $Q$, and $\Lambda$ be as in Figure 6.21 (so that the distance from $P$ to $\Lambda$ is $|PQ| = c$) and let $P'$ be the point on $\overrightarrow{QP}$ so that $|P'Q| = 2c$.

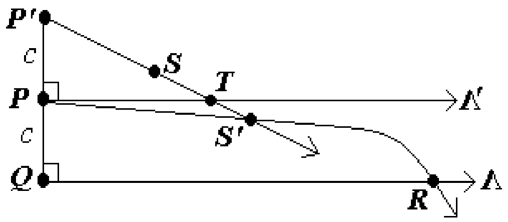



Figure 6.21:

- Let $\overleftrightarrow{P'S}$ be any line through $P'$ so that $m\angle QP'S < 90$. We need to show that $\overleftrightarrow{P'S}$ intersects $\Lambda$. By part **(a)** (which we have already proved) this will imply that $\theta_{2c} = 90$.
- Let $\Lambda'$ be the line through $P$ perpendicular to $\overleftrightarrow{PQ}$. (Again, see Figure 6.21.)
- Then the distance from $P'$ to $\Lambda'$ is $c$, so $\overleftrightarrow{P'S}$ is not in $\mathcal{P}(\Lambda', P')$ and so must meet $\Lambda'$ at some point $T$.

It might be tempting at this point to claim that $\overrightarrow{P'S}$ must meet $\Lambda$ because $T$ is distance $c$ from $\Lambda$. However, this would be making the (by now familiar) mistake of assuming that lines $\Lambda$ and $\Lambda'$ remain at constant distance – something we know to be false in hyperbolic geometry from Theorem 6.10. But we can still conclude that $\overrightarrow{P'S}$ meets $\Lambda$, as the following steps show.

- Let $S'$ be a point on $\overleftrightarrow{P'S}$ on the opposite side of $\Lambda'$ from $P'$ (so that $P' * T * S'$) and consider the ray $\overrightarrow{PS'}$.
- Since $m\angle QPS' < 90 = \theta_c$, we can again apply part **(a)** to conclude that $\overrightarrow{PS'}$ meets $\Lambda$ at some point $R$.
- Then $\overleftrightarrow{P'S}$ meets the side $PR$ of triangle $PQR$ (at $S'$) and so must meet one of the other sides of this triangle by Fact 4.8.
- But $\overleftrightarrow{P'S}$ already intersects $\overleftrightarrow{PQ}$ at $P'$, so it cannot intersect the side $PQ$.
- This means it must intersect $QR$, so $\overleftrightarrow{P'S}$ intersects $\Lambda$ as desired.

This shows that $\theta_{2c} = 90$ and repeating the argument we can conclude that $\theta_{2^k c} = 90$ for all $k$. But property **(ii)** of the functions $\theta_d(t)$ (see p.223) implies that if $d < d'$ then

$$\theta_d = \lim_{t \longrightarrow \infty} \theta_d(t) \geq \lim_{t \longrightarrow \infty} \theta_{d'}(t) = \theta_{d'} \,.$$

For any $d > 0$ there is some $k$ for which $d < 2^k c$ and thus

$$\theta_d \geq \theta_{2^k c} = 90\,.$$

This shows that $\theta_d = 90$ for all $d > 0$. But this is our contradiction, for the Hyperbolic Parallel Axiom obviously requires that $\theta_d$ be less than 90 for at least one value of $d$. □

In the above proof for part **(b)** we noted that if $d < d'$ then $\theta_d \geq \theta_{d'}$. So it may not seem like much should remain to proving part **(c)**. However, there is a good deal of subtlety involved in turning this inequality into a strict inequality. Follow the proof below carefully!

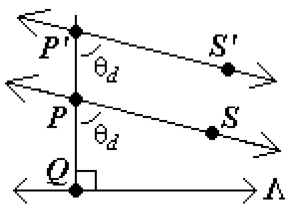



Figure 6.22:

*Proof of part* **(c)***:* Let $d$ and $d'$ be given with $d < d'$. Let $\Lambda$, $P$, and $Q$ be as before (so that $PQ \perp \Lambda$ with $|PQ| = d$) and let $P'$ be a point on $\overrightarrow{QP}$ so that $|QP'| = d'$. Let $\overleftrightarrow{PS}$ be a bounding parallel for $\mathcal{P}(\Lambda, P)$ (so that $m\angle QPS = \theta_d$ – remember that we have already proved part **(a)**!) and let $S'$ be a point on the same side of $\overleftrightarrow{QP'}$ as $S$ so that ,$\angle QP'S' = \theta_d$ (see Figure 6.22). To prove that $\theta_d > \theta_{d'}$ we need to show that $\overleftrightarrow{P'S'}$ is *not* a bounding parallel for $\mathcal{P}(\Lambda, P')$. That is, we need to show that there is a line $\overleftrightarrow{P'T}$ in $\mathcal{P}(\Lambda, P')$ so that $m\angle QP'T < \theta_d$.

First, find points $R$ on $\overleftrightarrow{PS}$ and $R'$ on $\overleftrightarrow{P'S'}$ so that $RR'$ is perpendicular to both lines (see Exercise 6.40). See Figure 6.23 for reference in the steps that follow.

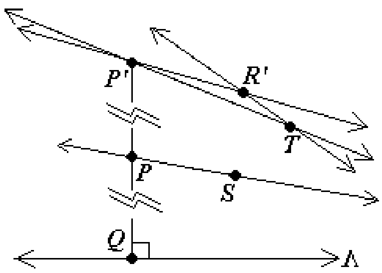



Figure 6.23:

- Letting $|RR'| = r$, we already know already (from part **(b)**) that $\theta_r < 90$.
- Let $T$ be a point so that $\theta_r \leq m\angle RR'T < 90$. Then $\overleftrightarrow{R'T}$ is in the family $\mathcal{P}(\overleftrightarrow{PS}, R')$ by part **(a)**.

- But then the entire line $\overleftrightarrow{R'T}$ must lie on the side of $\overleftrightarrow{PS}$ containing $R'$, and thus cannot intersect $\Lambda$.
- So it is not difficult to verify that $\overleftrightarrow{P'T}$ misses $\Lambda$ (see Exercise 6.41).
- Since $m\angle QP'T < m\angle QP'R' = \theta_d$ (the point $T$ is on the same side of $\overleftrightarrow{P'R'}$ as is $Q$), our proof is complete.

□

Part **(d)** of the theorem claims that the value of $\theta_d$ varies continuously with $d$. Not surprisingly, our Axiom C1 on continuity is the key to proving this.

*Proof of part* **(d)***:* Let $\Lambda$ be a line, $P$ a point not on $\Lambda$, $Q$ a point on $\Lambda$ so that $PQ \perp \Lambda$, and suppose $|PQ| = d$. Let $\overleftrightarrow{PR}$ be a bounding parallel for $\mathcal{P}(P, \Lambda)$ (so that $m\angle QPR = \theta_d$). Using Axiom L2, find a ruler function $s(t) = S_t$ for the line $\overleftrightarrow{PQ}$ so that $S_0 = Q$ and $S_d = P$.

Now let $\varepsilon > 0$ be given. We will show that there is a $\delta > 0$ so that $|\theta_d' - \theta_d| < \varepsilon$ whenever $|d' - d| < \delta$. (Refer to Figure 6.24.)

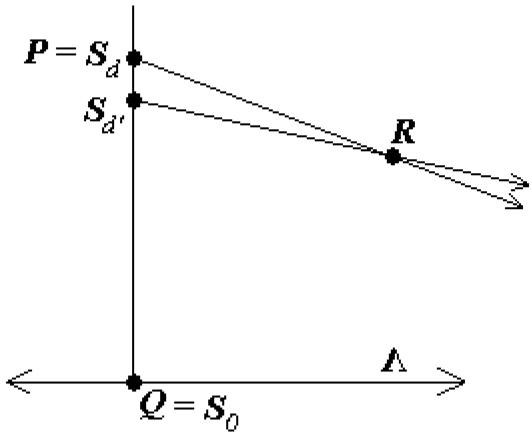



Figure 6.24:

- By Axiom C1 $m\angle QS_tR$ changes continuously with $t$.
- Thus, we may find $\delta$ so that $|m\angle QS_{d'}R - \theta_d| < \varepsilon$ whenever $|d' - d| < \delta$.
- But for $d' < d$, the ray $\overrightarrow{S_{d'}R}$ misses $\Lambda$. For other than the segment $S_{d'}R$ it lies in the halfplane determined by $\overleftrightarrow{PR}$ that does not contain $\Lambda$.
- Thus if $d - \delta < d' < d$ we have $m\angle QS_{d'}R < \theta_d + \varepsilon$ and $\overleftrightarrow{S_{d'}R}$ misses $\Lambda$.
- Since $|QS_{d'}| = d'$, by part **(a)** we have $\theta_{d'} < m\angle QS_{d'}R < \theta_d + \varepsilon$.
- A similar argument shows that $\theta_{d'} > \theta_d - \varepsilon$ when $d < d' < d + \varepsilon$ (see Exercise 6.42). □

*Proof of part* **(e)***:* The proof that $\lim_{d \longrightarrow 0} \theta_d = 90$ is fairly straightforward and we leave it as Exercise 6.43. We will prove $\lim_{d \longrightarrow \infty} \theta_d = 0$ by contradiction. Assume (to reach a contradiction) that $\lim_{d \longrightarrow \infty} \theta_d = \alpha > 0$. Then from parts **(b)** and **(c)**, we know the angle of parallelism for *every* distance is greater than $\alpha$. Consider the construction illustrated in Figure 6.25 and described in the steps below.

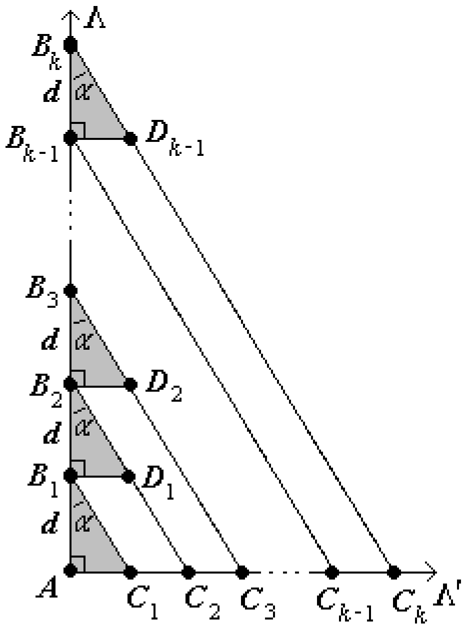



Figure 6.25:

- Let $\Lambda$ and $\Lambda'$ be perpendicular lines meeting at $A$.
- Let $r(t) = R_t$ be a ruler function for $\Lambda$ with $R_0 = A$, and for some $d > 0$ let $B_1 = R_d, B_2 = R_{2d}, B_3 = R_{3d}$, and so on.
- On one side of $\Lambda$ form rays at each of the points $B_1, B_2, B_3, \ldots$ so that the ray at $B_k$ makes an angle of measure $\alpha$ with $\overrightarrow{B_k A}$.
- By assumption, $\alpha$ is less than the angle of parallelism for every possible distance, so the ray from $B_k$ must intersect $\Lambda'$ at a point $C_k$ and (for $k \geq 2$) must also intersect the perpendicular line to $\Lambda$ through $B_{k-1}$ at a point $D_{k-1}$.
- The triangles $\triangle AB_1C_1$, $\triangle B_1B_2D_1$, $\triangle B_2B_3D_2, \ldots, \triangle B_{k-1}B_kD_{k-1}, \ldots$ are all congruent by the ASA criterion. (They each have angles of measure 90 and $\alpha$ bounding a side of length $d$.)
- By Theorem 6.8 the defect of each of these triangles is some positive number $\delta > 0$.
- By Theorem 6.4, then, for $k \geq 2$ we have $d(AB_kC_k) \geq k\delta$.
- But this will clearly cause a contradiction when $k$ is large enough that $k\delta > 180$. $\square$

The five parts of Theorem 6.16 establish the existence and major properties of the angle of parallelism in hyperbolic geometry. We'll close this section with an example of computing $\theta_d$ with the halfplane model from Section E.

EXAMPLE **6.17.** *Let $\Lambda$ be the H-line with equation $y = \sqrt{9-x^2}$ and let $P$ be the point $(0,5)$. Calculate $\theta_d$ where $d$ is the H-distance from $P$ to $\Lambda$.*

*Solution:* Clearly the $H$-segment from $P$ perpendicular to $\Lambda$ is the part of the $y$-axis joining $P$ to the point $Q = (0,3)$. Applying the formula for H-distance (from p.216) we can compute $d$ as follows.

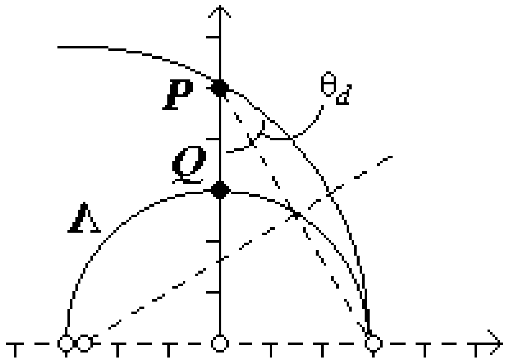



Figure 6.26:

$$\begin{aligned} d &= d_H((0,5),(0,3)) \\ &= \text{Arccosh}\left(1 + \frac{0+4}{30}\right) \\ &= \ln(5/3)\,. \end{aligned}$$

We can calculate $\theta_d$ by finding one of the bounding H-parallels for $\mathcal{P}(\Lambda, P)$.

- This, as illustrated in Figure 6.26, will be an H-line given by a (Euclidean) circle with diameter on the $x$-axis and passing through $(0,5)$ and $(3,0)$.
- The center of this (Euclidean) circle must be where the (Euclidean) perpendicular bisector of the (Euclidean) segment from $(0,5)$ to $(3,0)$ meets the $x$-axis. (Again, see Figure 6.26.)
- The slope of this perpendicular bisector is $3/5$ and it passes through the (Euclidean) midpoint $(\frac{3}{2}, \frac{5}{2})$. So, its equation is $y - \frac{5}{2} = \frac{3}{5}(x - \frac{3}{2})$, or $5y = 3x + 8$.
- It's easy to compute that the $x$-intercept of this line is $-8/3$, and $(-8/3, 0)$ is distance $17/3$ from $(3,0)$, so the equation of the circle we're after is $(x + \frac{8}{3})^2 + y^2 = (\frac{17}{3})^2$.
- The equation for our bounding H-parallel is thus

$$y = \sqrt{\left(\frac{17}{3}\right)^2 - \left(x + \frac{8}{3}\right)^2} = \sqrt{\frac{225}{9} - \frac{16}{3}x - x^2}\,.$$

- Differentiating this equation, we have

$$\frac{dy}{dx} = \frac{-\frac{16}{3} - 2x}{2\sqrt{\frac{225}{9} - \frac{16}{3}x - x^2}}\,.$$

- We are interested in the slope at $(0, 5)$, so we set $x = 0$ to get

$$\left.\frac{dy}{dx}\right|_{x=0} = \frac{-\frac{16}{3} - 0}{2\sqrt{\frac{225}{9} - 0 - 0}} = -8/15\,.$$

  (Alternately, we could get this tangent slope without differentiation by simply realizing that the tangent at $(0, 5)$ will be perpendicular to the radius from $(-8/3, 0)$ to $(0, 5)$. The radius slope is easily computed as $15/8$, so the tangent slope must be $-8/15$.)

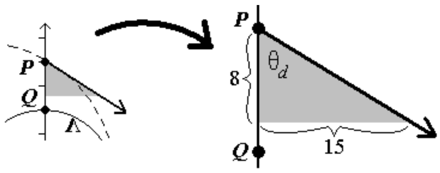



Figure 6.27:

- As shown in Figure 6.27, the angle between this bounding H-parallel and the H-segment from $P$ to $Q$ has measure Arctan(15/8), so we have computed

$$\theta_{\ln(5/3)} = \text{Arctan}(15/8) \approx 61.92751306 \text{ (measured in degrees).}$$

□

We should note that the exact value of $\theta_d$ depends on the distance function being used. The unit of distance is an arbitrary choice, and changing it has the effect of scaling the distance function by a constant multiple. However, regardless of that choice, the following remarkable relation holds:

$$\boxed{\tan\left(\frac{\theta_d}{2}\right) = e^{-kd}}$$

where $k$ is a constant determined by the choice of scale in the distance function. Deriving this formula is well beyond the scope of this text, but you should check (see Exercise 6.38) that it holds for the above calculation with $k = 1$.

## Exercises

**6.31**. In the halfplane model for hyperbolic geometry from Section E, let $\Lambda$ be the H-line $\{(0, y) : y > 0\}$ and let $P$ be the point $(10, 10)$. Make a sketch similar to Figure 6.16 showing the family $\mathcal{P}(\Lambda, P)$ of H-parallels to $\Lambda$ through $P$.

**6.32**. Make a halfplane model sketch illustrating the fact $\lim_{d\longrightarrow\infty} \theta_d = 0$ from part **(e)** of Theorem 6.16

**6.33**. In the halfplane model for hyperbolic geometry from Section E, let $P$ be the point $(0, 4)$ and let $\Lambda$ be the H-line with equation $y = \sqrt{4 - x^2}$.

**(a)** Sketch the family $\mathcal{P}(\Lambda, P)$.

**(b)** Find the equations of the two bounding H-parallels for $\mathcal{P}(\Lambda, P)$.

**(c)** Find $\theta_d$ where $d$ is the H-distance from $P$ to $\Lambda$.

**6.34**. Repeat Exercise 6.33 with the same H-line $\Lambda$ but changing $P$ to $(0, 8)$. Additionally, answer the following:

**(d)** Compare your calculations of $\theta_d$ in this exercise and Exercise 6.33. Is this consistent with part **(c)** of Theorem 6.16? Explain.

**6.35**. Repeat Exercise 6.33 with the same H-line $\Lambda$ but changing $P$ to $(0, 1)$. Additionally, answer the following:

**(d)** Compare your calculation of $\theta_d$ with that from Exercise 6.33. Can you explain this? (Use the formula for H-distance on p.216.)

**6.36**. Repeat Exercise 6.33 with $\Lambda$ as the H-line $x = 0$ $(y > 0)$ and $P$ the point $(4, 1)$.

**6.37**. In the halfplane model for hyperbolic geometry, let $\Lambda$ be the H-line $y = \sqrt{4 - x^2}$ and let $P$ be the point $(4, 2)$. Compute $\theta_d$ where $d$ is the H-distance from $P$ to $\Lambda$. (*Hint:* you don't have to calculate $d$. Instead, find the two bounding H-parallels and use them to calculate $2\theta_d$.)

**6.38**. Check the formula $\tan(\theta_d/2) = e^{-kd}$ (where $k = 1$) for our calculation of $\theta_{\ln(5/3)}$.

**6.39**. Prove that in Euclidean geometry $\theta_d = 90$ for all distances $d$.

**6.40**. Prove that in Figure 6.22 there are points $R$ on $\overleftrightarrow{PS}$ and $R'$ on $\overleftrightarrow{P'S'}$ so that $RR'$ is perpendicular to both of these lines. (*Hint:* let $M$ be the midpoint of $PP'$ and form perpendicular segments from $M$ to both lines.)

**6.41**. Verify in the last part of the proof of part **(c)** to Theorem 6.16 that $\overleftrightarrow{P'T}$ misses $\Lambda$.

**6.42**. Complete the proof of part **(d)** to Theorem 6.16 by explaining why $\theta_{d'} > \theta_d - \varepsilon$ when $d < d' < d + \delta$.

**6.43**. Complete the proof of part **(e)** to Theorem 6.16 by showing that $\lim_{d \longrightarrow 0} \theta_d = 90$. (*Hint:* let $\Lambda$ be perpendicular to $\overleftrightarrow{AC}$ at $A$. Take a ruler function $p(t) = P_t$ for $\Lambda$ with $P_0 = A$ and let $B = P_1$. Now use Axiom C1 on $m\angle BP_tC$ where $t$ approaches zero from the negative side.)

## G. Asymptotic Parallels and Ideal Points

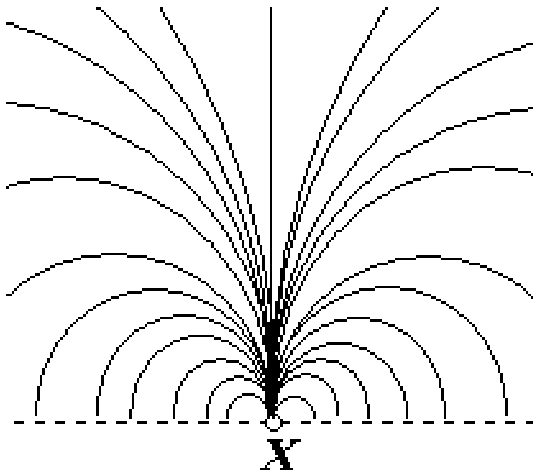



Figure 6.28:

The motivation for our work in this section is found in the halfplane model diagram of Figure 6.28. This depicts a family of parallel H-lines, each converging to a point $X$ on the $x$-axis. Now the $x$-axis is not part of the upper half plane, so the point $X$ is not really present in the hyperbolic plane (thus the H-lines in the figure really *are* parallel). So the way in which we have described this interesting family of H-lines uses a feature of the model (the point $X$) that is not part of the geometry. Is there a counterpart to the behavior in this figure that doesn't rely on the model? Can we transfer this diagram to a description of an interesting family of lines in hyperbolic geometry? That is the goal of this section.

The key, it turns out, is found in the concept of angle of parallelism. Suppose that $\Lambda$ is an H-line in the model converging to the point $X$ on the $x$-axis, and let $P$ be a point not on $\Lambda$. You no doubt have noticed (see Figure 6.29) that the H-line $\Lambda'$ through $P$ converging to $X$ is a bounding H-parallel for the family $\mathcal{P}(\Lambda, P)$. This suggests that the family of H-lines in Figure 6.28 might be identified with bounding parallels of all families $\mathcal{P}(\Lambda, P)$ (where $P$ is allowed to be any point not on $\Lambda$).

That's *almost* the correct notion – the only problem being that each family $\mathcal{P}(\Lambda, P)$ contains *two* bounding parallels, only one of which converges to $X$ (see again Figure 6.29). To distinguish which one we want, we'll need to introduce a new notion.

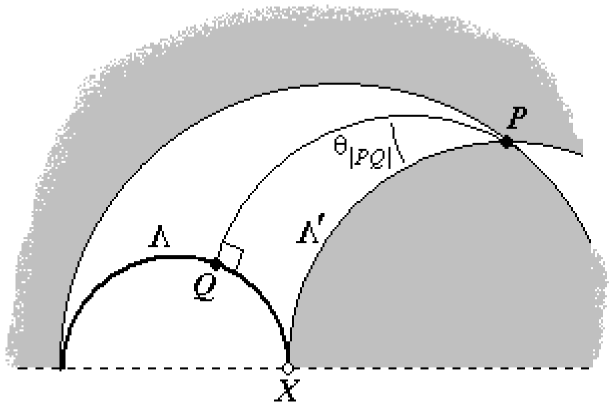



Figure 6.29:

From now on we'll deal not only with lines, but with *oriented lines*. Intuitively, we want our definition of line orientation to designate a "positive end" for a line. Suppose $A$ and $B$ are points on $\Lambda$, and let $r(t) = R_t$ be a ruler function for $\Lambda$. Let $t_A = r^{-1}(A)$ and $t_B = r^{-1}(B)$ (so that $A = R_{t_A}$ and $B = R_{t_B}$). Then clearly there are only two possibilities: either $t_A < t_B$ or $t_B < t_A$. We will say that the ruler function $r$ is of type $(A, B)$ if $t_A < t_B$ and of type $(B, A)$ if $t_B < t_A$. We leave it as Exercise 6.44 to check the unsurprising fact that if $C$ and $D$ are two other points on $\Lambda$, then the division of ruler functions for $\Lambda$ into types $(C, D)$ and $(D, C)$ is exactly the same as for types $(A, B)$ and $(B, A)$. (That is, either every ruler function of type $(A, B)$ is also of type $(C, D)$, or every ruler function of type $(A, B)$ is also of type $(D, C)$.) This justifies the following definition.

**Definition.** An *orientation* of the line $\Lambda = \overleftrightarrow{AB}$ is either one of the two pairs $(A, B)$ or $(B, A)$. The symbol ${}_{(A,B)}\Lambda$ will denote the *oriented line* consisting of $\Lambda$ under the orientation $(A, B)$. As a point set, this will be the same as the unoriented line $\Lambda$. However, any ruler function $r$ for ${}_{(A,B)}\Lambda$ is required to satisfy $r^{-1}(A) < r^{-1}(B)$.

As you consider the next definition, keep in mind that our goal is to describe, in terms independent of the model, the line behavior exhibited in Figure 6.28.

**Definition.** Let $\Lambda$ be a line and $P$ a point not on $\Lambda$. Let $Q$ be the point on $\Lambda$ with $PQ \perp \Lambda$ and let $\Lambda'$ be a bounding parallel for $\mathcal{P}(\Lambda, P)$. Let $R$ be a point on $\Lambda'$ with $m\angle QPR = \theta_{|PQ|}$. Finally, let $S$ be a point on $\Lambda$ on the same side of $\overleftrightarrow{PQ}$ as $R$. Then we say that the oriented line ${}_{(P,R)}\Lambda'$ is *asymptotically parallel* to the oriented line ${}_{(Q,S)}\Lambda$ (see Figure 6.30).

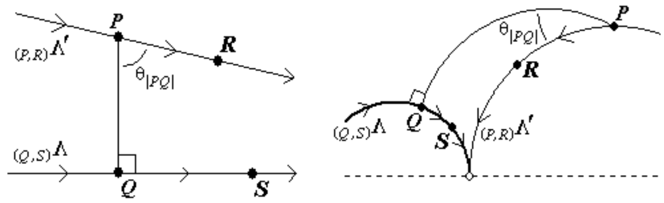



Figure 6.30: ${}_{(P,R)}\Lambda'$ is asymptotically parallel to ${}_{(Q,S)}\Lambda$ (halfplane model diagram on the right)

Note that this definition is *not* immediately symmetric – there is no guarantee in the definition that just because ${}_{(P,R)}\Lambda'$ is asymptotically parallel to ${}_{(Q,S)}\Lambda$ that we will also have ${}_{(Q,S)}\Lambda$ asymptotically parallel to ${}_{(P,R)}\Lambda'$. We'll see shortly that the relation of asymptotic parallelism *is* symmetric, but we'll need to prove that fact (and will do so in Lemma 6.19).

There is another more fundamental issue to address regarding this definition. The alert reader should have noticed a concern about the soundness of the definition itself. It seems to rely not just on the oriented lines, but on the chosen point $P$ as well. Suppose that $P'$ is another point on $\Lambda'$. In what we've proved up to now we have no guarantee that $\Lambda'$ is a bounding parallel for $\mathcal{P}(\Lambda, P')$, thus no guarantee that the above definition is satisfied at $P'$, even if it is satisfied at $P$. Our first lemma will clear up that difficulty.

**Lemma 6.18.** *Suppose that ${}_{(P,R)}\Lambda'$ and ${}_{(Q,S)}\Lambda$ are oriented lines as in the above definition (so that $PQ \perp \Lambda$ and $m\angle QPR = \theta_{|PQ|}$). Let $P'$ be any other point on $\Lambda'$ and let $Q'$ be the point on $\Lambda$ with $P'Q' \perp \Lambda$. Then if $R'$ is a point on $\Lambda'$ with $(P', R')$ and $(P, R)$ giving the same orientation of $\Lambda'$, then $m\angle Q'P'R' = \theta_{|P'Q'|}$.*

*Proof:* There are two cases to consider: either $P'$ is on ray $\overrightarrow{PR}$ or it is not. We will leave the proof in the case $P'$ is on $\overrightarrow{PR}$ as an exercise and handle the latter case here. So, assume $P'$ is not on $\overrightarrow{PR}$. In this case we can choose $R' = R$ (since the orientations $(P, R)$ and $(P', R)$ of $\Lambda'$ are identical). We must show that $m\angle Q'P'R = \theta_{|P'Q'|}$. Assume (to reach a contradiction) that this is not the case, and refer to Figure 6.31 in the following steps.

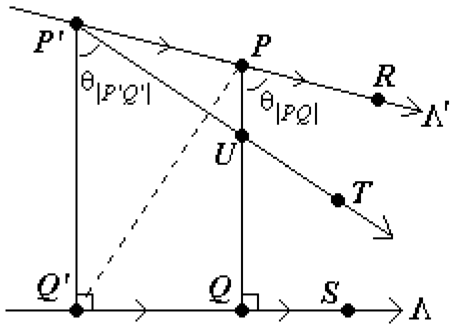



Figure 6.31:

- Since $\Lambda'$ is parallel to $\Lambda$, part **(a)** of Theorem 6.16 implies that $m\angle Q'P'R > \theta_{|P'Q'|}$.
- So there is then a point $T$ interior to $\angle Q'P'R$ with $m\angle Q'P'T = \theta_{|P'Q'|}$.
- Then $\overrightarrow{P'T}$ must meet $PQ'$ by Axiom C4 applied to triangle $Q'P'P$.
- So then by Pasch's Theorem (Fact 4.8) applied to triangle $Q'PQ$, $\overrightarrow{P'T}$ must meet either $QQ'$ or $PQ$. Being parallel to $\Lambda$, it cannot meet $QQ'$ so we conclude that $\overrightarrow{P'T}$ meets $PQ$ at a point $U$. We will assume that (as in the figure) $U$ is between $P'$ and $T$. (If not, we may simply choose a new point $T$.)
- By the Saccheri-Legendre Theorem applied to $P'PU$, $m\angle P'PU + m\angle PUP' < 180$.
- So applying the Vertical Angles Theorem, $m\angle QUT = m\angle PUP' < 180 - m\angle P'PU = \theta_{|PQ|}$.
- But since $\overleftrightarrow{P'T} \parallel \Lambda$, this means $\theta_{|UQ|} \leq m\angle QUT < \theta_{|PQ|}$, which contradicts part **(c)** of Theorem 6.16. This contradiction shows that we must have $m\angle Q'P'R = \theta_{|P'Q'|}$, as claimed. $\square$

What we just proved shores up the definition of asymptotic parallelism – we now know that if ${}_{(P,R)}\Lambda'$ is asymptotically parallel to ${}_{(Q,S)}\Lambda$ at one point of $\Lambda'$, then it will be asymptotically parallel to ${}_{(Q,S)}\Lambda$ at all points of $\Lambda'$. The next two lemmas show that this relation of one oriented line to another is both *symmetric* and *transitive.*

**LEMMA 6.19.** *If ${}_{(C,D)}\Lambda'$ is asymptotically parallel to ${}_{(A,B)}\Lambda$ then ${}_{(A,B)}\Lambda$ is also asymptotically parallel to ${}_{(C,D)}\Lambda'$.*

*Proof:* Refer to Figure 6.32 in the following steps. Since ${}_{(C,D)}\Lambda'$ is asymptotically parallel to ${}_{(A,B)}\Lambda$ we may assume that $CA \perp \Lambda$ with $|CA| = d$ and $m\angle ACD = \theta_d$.

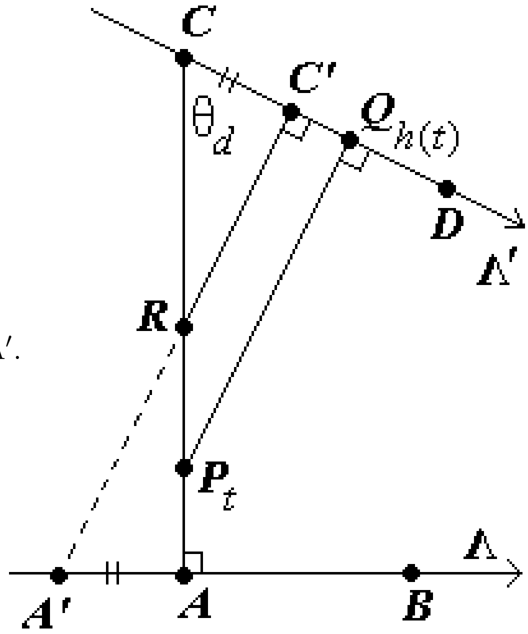



Figure 6.32:

- Find a ruler function $p(t) = P_t$ for $\overleftrightarrow{AC}$ with $P_0 = A$ and $P_d = C$. Also find a ruler function $q(t) = Q_t$ for $\Lambda'$.
- For each $t$ in the range $0 \le t < d$ let $h(t)$ be such that $Q_{h(t)}$ is the point on $\Lambda'$ with $P_tQ_{h(t)} \perp \Lambda'$.
- Then $h(t)$ is a continuous function and it isn't hard to extend this to the conclusion that the length $\left|P_tQ_{h(t)}\right|$ varies continuously with $t$. (We leave the verifications to the interested reader!)
- So $j(t) = t - \left|P_tQ_{h(t)}\right|$ is continuous, with $j(0) < 0$ and $j(t) > 0$ as $t$ approaches $d$.
- By the Intermediate Value Theorem, there is some $t_0$ between 0 and $d$ with $j(t_0) = 0$ – in other words, $\left|P_{t_0}Q_{h(t_0)}\right| = t_0$.
- Letting $R = P_{t_0}$ and $C' = Q_{h(t_0)}$ we have $|AR| = t_0 = |RC'|$.
- Let $A'$ be a point on $\Lambda$ with $|AA'| = |CC'|$, as in the figure.
- Then $\triangle ARA' \cong \triangle C'RC$ by Axiom SAS.
- This means $\angle ARA' \cong \angle C'RC$, so $C'$, $R$, and $A'$ are collinear by the Vertical Angles Theorem.

- Also, the above triangle congruence gives us $|RA'| = |RC|$, so $|A'C'| = |AC| = d$.

- Finally, since (again by the triangle congruence) $m\angle RA'A = m\angle RCC' = \theta_d$, we see that ${}_{(A,B)}\Lambda$ is asymptotically parallel to ${}_{(C,D)}\Lambda'$, as claimed. $\square$

**LEMMA 6.20.** *If both ${}_{(C,D)}\Lambda'$ and ${}_{(E,F)}\Lambda''$ are asymptotically parallel to ${}_{(A,B)}\Lambda$ then they are asymptotically parallel to each other.*

*Proof:* There are two cases to consider.

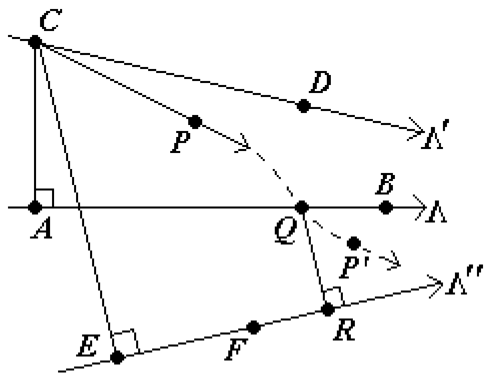



Figure 6.33:

**Case 1.** First suppose that $\Lambda'$ and $\Lambda''$ are in opposite halfplanes determined by $\Lambda$. We may assume that the points $A$, $B$, $C$, $D$, $E$, and $F$ are as in Figure 6.33 (so that $CA \perp \Lambda$ and $CE \perp \Lambda''$). By assumption we have $m\angle ACD = \theta_{|CA|}$. We will prove that $m\angle ECD = \theta_{|CE|}$.

Let $P$ be a point interior to $\angle ECD$. It will be enough for us to show that $\overrightarrow{CP}$ meets $\Lambda''$. Note that since $C$ and $E$ are on opposite sides of $\Lambda$, the segment $CE$ must meet $\Lambda$, and thus both $\Lambda$ and $\Lambda''$ are on the same side of $\Lambda'$. Thus the rays $\overrightarrow{CA}$ and $\overrightarrow{CE}$ lie on the same side of $\Lambda'$, so we may assume that $P$ is also interior to $\angle ACD$.

- Then $m\angle ACP < m\angle ACD = \theta_{|CA|}$, so $\overrightarrow{CP}$ must meet $\Lambda$ at a point $Q$. We may (by moving $B$ if necessary) assume that $(A, B)$ and $(Q, B)$ are equivalent orientations of $\Lambda$.

- Let $R$ be the point on $\Lambda''$ so that $QR \perp \Lambda''$. Then $m\angle RQB = \theta_{|QR|}$ (since ${}_{(E,F)}\Lambda''$ is asymptotically parallel to ${}_{(A,B)}\Lambda$).

- But then if $P'$ is a point on $\overrightarrow{CP}$ beyond $Q$ (as in the figure) we have $m\angle RQP' < m\angle RQB$, so $\overrightarrow{QP'}$ must meet $\Lambda''$.

- This shows that $\overrightarrow{CP}$ meets $\Lambda''$ and completes the proof in this case.

**Case 2.** Now suppose that $\Lambda'$ and $\Lambda''$ are on the same side of $\Lambda$. It seems intuitive to conjecture that given any three mutually parallel lines there must be one that is "between" the other two (in the sense that the other two are on opposite sides of it). However, this is not the case in hyperbolic geometry! But we have the additional hypothesis that ${}_{(C,D)}\Lambda'$ and ${}_{(E,F)}\Lambda''$ are both asymptotically parallel to ${}_{(A,B)}\Lambda$, and it isn't hard to justify (see Exercise 6.46) that in this case we may assume that $\Lambda$ and $\Lambda''$ are on opposite sides of $\Lambda'$. Assume also that $A$, $B$, $C$, $D$, $E$, and $F$ are as in Figure 6.34 (so that $EA \perp \Lambda$ and $EC \perp \Lambda'$). Let $P$ be any point interior to $\angle CEF$. As in the above case, we may assume that $P$ is also interior to $\angle AEF$. We will prove ${}_{(E,F)}\Lambda''$ is asymptotically parallel to ${}_{(C,D)}\Lambda'$ by showing that $\overrightarrow{EP}$ must meet $\Lambda'$. This is surprisingly easy.

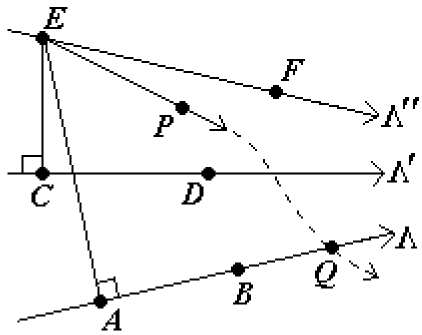



Figure 6.34:

- Since $m\angle AEF = \theta_{|EA|}$, the ray $\overrightarrow{EP}$ must meet $\Lambda$ at a point $Q$.
- But then $\overrightarrow{EP}$ contains points ($E$ and $Q$) on opposite sides of $\Lambda'$ and thus must meet $\Lambda'$. $\square$

Now let's add the convention that an oriented line should be considered to be asymptotically parallel to itself (so that the relation "asymptotically parallel" also has the reflexive property). We can then summarize what we've learned in the preceding lemmas as follows.

THEOREM **6.21.** *The relation of being asymptotically parallel is an equivalence relation on the set of oriented lines. If ${}_{(A,B)}\Lambda$ is an oriented line and $P$ is a point not on $\Lambda$ then there is exactly one oriented line through $P$ (namely one of the bounding parallels for $\mathcal{P}(\Lambda, P)$) that is asymptotically parallel to ${}_{(A,B)}\Lambda$.* $\square$

Equivalence relations, as you are probably aware, partition the set on which they are defined into *equivalence classes.* In this case, we have equivalence classes of asymptotically parallel oriented lines. We give these equivalence classes a special name.

**Definition.** An *ideal point* $\mathcal{X}$ is an equivalence class of oriented lines under the equivalence relation of being asymptotically parallel.[5]

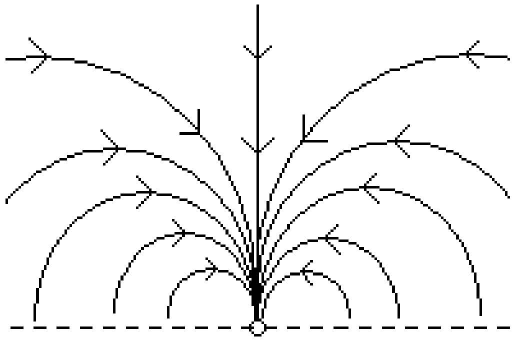


Figure 6.35: An ideal point in the halfplane model

The reason for the terminology "ideal point" is evident in Figure 6.35, which shows an ideal point in the halfplane model. It is easy to associate this equivalence class of oriented H-lines with the point on the $x$-axis to which all of the H-lines converge. It is an "ideal" point (as opposed to a "real" point) because it is not really a point in the geometry – it is more of a "direction" or "point at infinity". In the halfplane model, then, the $x$-axis acts as a sort of horizon. (There is actually one ideal point not corresponding to a point on the $x$-axis – see Exercise 6.47.)

Note that we have accomplished exactly what we set out to do in this section. We have found a meaning in pure axiomatic hyperbolic geometry for the phenomenon on display in Figure 6.28. In fact, we've done more than that – we've deduced from the axioms that something very much like the limiting horizon of the halfplane model's $x$-axis is built into the very nature of hyperbolic geometry. The $x$-axis is not a quirk of that particular model!

We'll close this section with two theorems that show how ideal points behave in some ways much like "real" points. If we think of a line $\Lambda$ as being "through" an ideal point $\mathcal{X}$ (or equivalently, the ideal point $\mathcal{X}$ as being "on" the line $\Lambda$) if an orientation of $\Lambda$ belongs to $\mathcal{X}$, then the second part of Theorem 6.21 above can be rephrased as "if $P$ is a point and $\mathcal{X}$ is an ideal point then there is a unique line through both $P$ and $\mathcal{X}$". Theorems 6.22 and 6.23 below could similarly be rephrased as

- each pair of ideal points determines a unique line.
- given a line $\Lambda$ and an ideal point $\mathcal{X}$ not on $\Lambda$, there is a unique line through $\mathcal{X}$ that is perpendicular to $\Lambda$.

**Theorem 6.22.** *If $\mathcal{X}_1$ and $\mathcal{X}_2$ are two distinct ideal points then there is a unique (unoriented) line $\Lambda = \overleftrightarrow{AB}$ such that ${}_{(A,B)}\Lambda \in \mathcal{X}_1$ and ${}_{(B,A)}\Lambda \in \mathcal{X}_2$.*

[5] Such an equivalence class is also sometimes called an *asymptotic pencil* of oriented lines.

*Proof:* Refer to Figure 6.36 in the following steps.

- Let $P$ be any point and let ${}_{(P,R_1)}\Lambda' \in \mathcal{X}_1$ and ${}_{(P,R_2)}\Lambda'' \in \mathcal{X}_2$ be the (unique) oriented lines through $P$ belonging to the two ideal points.
- Let $\overrightarrow{PQ}$ be the angle bisector of $\angle R_1PR_2$.
- By parts **(d)** and **(e)** of Theorem 6.16 (and the Intermediate Value Theorem) there must be a number $d > 0$ so that $\theta_d = \frac{1}{2}(m\angle R_1PR_2)$.

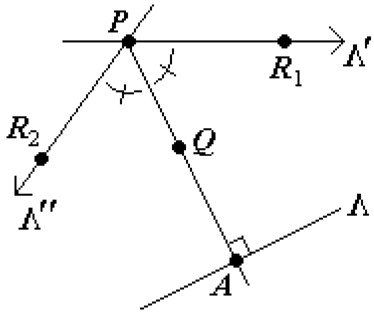



Figure 6.36:

- Let $A$ be the point on $\overrightarrow{PQ}$ with $|PA| = d$, and let $\Lambda$ be the line through $A$ with $\Lambda \perp PA$.
- It is clear that the two possible orientations of $\Lambda$ are asymptotically equivalent to ${}_{(P,R_1)}\Lambda'$ and ${}_{(P,R_2)}\Lambda''$ respectively.

We leave the proof that $\Lambda$ is unique as Exercise 6.49. $\square$

**THEOREM 6.23.** *If $\mathcal{X}$ is an ideal point and $\Lambda$ is a line neither orientation of which is in $\mathcal{X}$, then there is a unique ${}_{(A,B)}\Lambda' \in \mathcal{X}$ so that $\Lambda' \perp \Lambda$.*

*Proof:* See Exercise 6.50.

## Exercises

**6.44**. Suppose $\Lambda$ is a line through points $A$, $B$, $C$, and $D$. Suppose there is a ruler function for $\Lambda$ that is both of type $(A, B)$ and type $(C, D)$ (see p.233). Show that every ruler function for $\Lambda$ of type $(A, B)$ is also of type $(C, D)$. (Consider using the rays $\overrightarrow{AB}$ and $\overrightarrow{CD}$.)

**6.45**. Complete the proof of Lemma 6.18 in the case that point $P'$ is on ray $\overrightarrow{PR}$.

**6.46**. Let $\Lambda$, $\Lambda'$, and $\Lambda''$ be three mutually parallel lines.

**(a)** Show using the halfplane model that it is possible for none of these three to separate the remaining two lines into different halfplanes.

**(b)** Show (without using Lemma 6.20, which requires this fact in its proof) that if ${}_{(C,D)}\Lambda'$ and ${}_{(E,F)}\Lambda''$ are both asymptotically parallel to ${}_{(A,B)}\Lambda$ then one of the lines *does* separate the other two in different halfplanes. (*Hint:* assume that $\Lambda'$ and $\Lambda''$ are both on the same side of $\Lambda$ and that $C$ and $E$ are such that $AC \perp \Lambda'$ and $AE \perp \Lambda''$. Show that either $\overrightarrow{AC}$ or $\overrightarrow{AE}$ must hit both $\Lambda'$ and $\Lambda''$.)

**6.47**. Make a sketch similar to Figure 6.35 showing the ideal point consisting of all oriented H-lines asymptotically parallel to the H-line $\{(0, y) : y > 0\}$ oriented in the direction of *increasing* $y$-coordinate.

**6.48**. Prove that in hyperbolic geometry the corresponding angles formed by a line transversing two asymptotically parallel oriented lines are *not* congruent.

**6.49**. Prove the uniqueness portion of Theorem 6.22.

**6.50**. Prove Theorem 6.23 by filling in the details to the following steps:

- Let $A$ be any point of $\Lambda$ and let ${}_{(A,B)}\Lambda'$ be the member of $\mathcal{X}$ through $A$. (Why is it true that $\Lambda' \neq \Lambda$?)
- Let $C$ be a point on $\Lambda$ so that $\angle CAB$ is acute. (If $\Lambda' \perp \Lambda$ then we are done!)
- Find a point $P$ on $\overrightarrow{AC}$ so that $m\angle PAB = \theta_{|PA|}$, and consider an orientation of the perpendicular to $\Lambda$ through $P$.
- Prove uniqueness by showing that no other member of $\mathcal{X}$ can be perpendicular to $\Lambda$.

**6.51**. Part **(a)** of Theorem 6.16 is valid in neutral geometry, so one can define asymptotic parallelism for oriented lines in Euclidean geometry.

**(a)** Describe an ideal point in Euclidean geometry.

**(b)** Show that Theorems 6.22 and 6.23 are not valid in Euclidean geometry.

**(c)** Examine the proofs of these theorems (see Exercise 6.50) and determine where the Hyperbolic Parallel Axiom is used.

## H. Types of Parallel Lines

Given two parallel lines $\Lambda$ and $\Lambda'$ there are two possibilities. Either there are orientations of these two lines yielding asymptotically parallel oriented lines, or no such orientations can be found. We will (by a slight abuse of terminology) say that that (the *un*oriented lines) $\Lambda$ and $\Lambda'$ are *asymptotically parallel* in the former case. There is also a name for the latter possibility.

DEFINITION. Two lines are *ultraparallel* if they are parallel but not asymptotically parallel under any of their orientations.

Our task in this section is to characterize the behavior of these two types of parallel lines. The results we will prove are suggested by the informal diagrams in Figure 6.37:

- An ultraparallel pair of lines has a common perpendicular at which the minimum distance between the lines is achieved. The distance between the lines diverges to infinity toward either extremity.

- An asymptotically parallel pair of lines has no common perpendicular. The distance between the lines diverges to infinity toward one extremity and tends to zero toward the other.

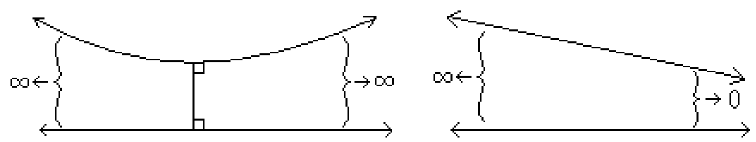



Figure 6.37: The behavior of ultraparallel lines (left) and asymptotically parallel lines (right)

As in the last section, we'll do our work in a sequence of lemmas. Our first lemma applies to the behavior at both extremities when $\Lambda$ and $\Lambda'$ are ultraparallel and to the behavior at one extremity when they are asymptotically parallel – see Exercise 6.52.

**Lemma 6.24.** *Let $\Lambda$ and $\Lambda'$ be parallel lines, and let $(A,B)$ be an orientation of $\Lambda$ so that neither orientation of $\Lambda'$ is asymptotically parallel to $_{(A,B)}\Lambda$. Let $p(t) = P_t$ be a ruler function for $\Lambda$ consistent with the orientation $(A,B)$, and for each $t$ let $Q_t$ be the point on $\Lambda'$ with $P_tQ_t \perp \Lambda'$. Then $\lim_{t\longrightarrow\infty} |P_tQ_t| = \infty$.*

*Proof:* Refer to Figure 6.38 in the following steps.

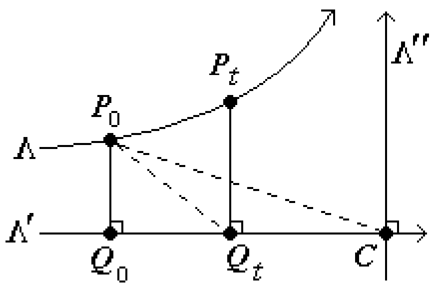



Figure 6.38:

- Let $\mathcal{X}$ be the ideal point containing $_{(A,B)}\Lambda$. By Theorem 6.23 there is an oriented line $_{(C,D)}\Lambda'' \in \mathcal{X}$ with $\Lambda'' \perp \Lambda'$. We may assume that $C$ is the point common to $\Lambda''$ and $\Lambda'$.
- Let $t > 0$ be any positive number. Then (since $\Lambda' \parallel \Lambda''$) the points $P_0$, $Q_0$, $P_t$, and $Q_t$ are all on the same side of $\Lambda''$. It is easy to check that $Q_0 * Q_t * C$.
- Applying Corollary 6.2 to triangle $P_0Q_tC$ we have $|P_0C| > |P_0Q_t|$ (since $\angle P_0Q_tC$ is clearly obtuse).
- But then using the triangle inequality we have

$$\begin{aligned} |P_tQ_t| + |P_0Q_t| &\geq |P_0P_t| = t \\ |P_tQ_t| &\geq t - |P_0Q_t| \\ &\geq t - |P_0C| \end{aligned}$$

- But $|P_0C|$ is constant, so $|P_tQ_t|$ will tend to infinity along with $t$. $\square$

**Lemma 6.25.** *Two parallel lines $\Lambda$ and $\Lambda'$ are ultraparallel if and only if they have a common perpendicular line.*

*Proof:* There are two parts to the proof. We prove here that if $\Lambda$ and $\Lambda'$ are ultraparallel then they have a common perpendicular, leaving the proof of the converse as an exercise. Assume, then, that $\Lambda$ and $\Lambda'$ are ultraparallel.

- Take a ruler function $p(t) = P_t$ for $\Lambda$ and for each $t$ let $Q_t$ be the point on $\Lambda'$ so that $P_tQ_t \perp \Lambda'$. Then by Lemma 6.24 we can say that $\lim_{t\longrightarrow\infty} |P_tQ_t| = \infty = \lim_{t\longrightarrow-\infty} |P_tQ_t|$.

- So, we may easily find $t_0$ and $t_1$ so that $|P_{t_0}| = |P_{t_t}|$.

- But then $Q_0Q_1P_1P_0$ is a Saccheri quadrilateral (with base $Q_0Q_1$).

- Let $A$ and $B$ be the midpoints of $P_0P_1$ and $Q_0Q_1$ respectively. By Theorem 6.6 the segment $AB$ is perpendicular to both $\Lambda = \overleftrightarrow{P_0P_1}$ and $\Lambda' = \overleftrightarrow{Q_0Q_1}$. □

All that remains is to show that the distance between asymptotically parallel lines tends to zero toward one extremity. The proof is both challenging and charming, drawing together many concepts from our work in this chapter.

**LEMMA 6.26.** *Let ${}_{(A,B)}\Lambda$ and ${}_{(C,D)}\Lambda'$ be two asymptotically parallel oriented lines. Let $p(t) = P_t$ be a ruler function for $\Lambda$ consistent with the orientation $(A, B)$, and for each $t$ let $Q_t$ be the point on $\Lambda'$ with $P_tQ_t \perp \Lambda'$. Then $\lim_{t\longrightarrow\infty} |P_tQ_t| = 0$.*

*Proof:* We separate the proof into a few major steps.

**Step 1.** We first claim that $|P_tQ_t|$ decreases as $t$ increases. For assume to reach a contradiction that $|P_aQ_a| < |P_bQ_b|$ where $a < b$.

- The length $|P_tQ_t|$ varies continuously with $t$ and we already know from Lemma 6.24 that $\lim_{t\longrightarrow-\infty} |P_tQ_t| = \infty$, so there must be $c < a$ with $|P_cQ_c| = |P_bQ_b|$.

- But then $Q_cQ_bP_bP_c$ is a Saccheri quadrilateral.

- By part **(b)** of Theorem 6.6, then, $\Lambda$ and $\Lambda'$ have a common perpendicular.

- But this contradicts Lemma 6.25 since $\Lambda$ and $\Lambda'$ are not ultraparallel.

**Step 2.** For each $t$ let $w(t)$ denote the distance $|Q_0Q_t|$ on $\Lambda'$. We claim that $\lim_{t\longrightarrow\infty} w(t) = \infty$. To see this, note that for all $t > 0$ we have

$$\begin{aligned} t &= |P_0P_t| \\ &< |P_0Q_0| + |Q_0Q_t| + |P_tQ_t| \\ &< 2|P_0Q_0| + w(t) \end{aligned}$$

(remember that $|P_tQt|$ decreases as $t$ increases). Clearly as $t$ tends to infinity, $w(t)$ must also tend to infinity.

**Step 3.** Since $|P_tQ_t|$ is a decreasing function of $t$ (obviously bounded below by zero), the limit $\lim_{t\longrightarrow\infty} |P_tQ_t|$ must exist. Assume to reach a contradiction that this limit is $d > 0$. Refer to Figure 6.39 in the following steps.

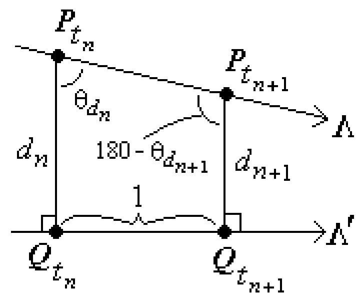



Figure 6.39:

- By step two, we may find values $t_1 < t_2 < t_3 < \cdots$ so that $w(t_n) = n$ for all $n \geq 1$.
- Let $d_n$ denote the length $|P_{t_n}Q_{t_n}|$. Then $\lim_{n\longrightarrow\infty} d_n = d$.
- Since $\Lambda$ and $\Lambda'$ are asymptotically parallel, $m\angle Q_{t_n}P_{t_n}P_{t_{n+1}} = \theta_{d_n}$ and $m\angle Q_{t_{n+1}}P_{t_{n+1}}P_{t_n} = 180 - \theta_{d_{n+1}}$.
- Since the angle of parallelism varies continuously with distance (part **(d)** of Theorem 6.16) we can say that $\lim_{n\longrightarrow\infty} \theta_{d_n} = \theta_d$.
- Now let $\varepsilon > 0$ be given, and choose $n$ large enough that $\theta_{d_n} > \theta_d - \varepsilon$.
- Then the defect of $Q_{t_n}Q_{t_{n+1}}P_{t_{n+1}}P_{t_n}$ is

$$\begin{aligned} d(Q_{t_n}Q_{t_{n+1}}P_{t_{n+1}}P_{t_n}) &= 360 - (90 + 90 + \alpha_n + \theta_{d_n}) \\ &= 180 - \alpha_n - \theta_{d_n} \\ &< 180 - (180 - \theta_{d_{n+1}}) - \theta_{d_n} \\ &= \theta_{d_{n+1}} - \theta_{d_n} \\ &< \theta_d - \theta_{d_n} \\ &< \theta_d - (\theta_d - \varepsilon) = \varepsilon \end{aligned}$$

- But let $x$ represent the defect of a Saccheri quadrilateral with base length 1 and sides of length $d$. By Theorem 6.9 we know that $x > 0$.

- But, as seen from Figure 6.40 and Theorem 6.4, $x < d(Q_{t_n}Q_{t_{n+1}}P_{t_{n+1}}P_{t_n}) < \varepsilon$.

- Since $\varepsilon > 0$ was arbitrary, we are forced to conclude $x = 0$ and we at last have our contradiction!

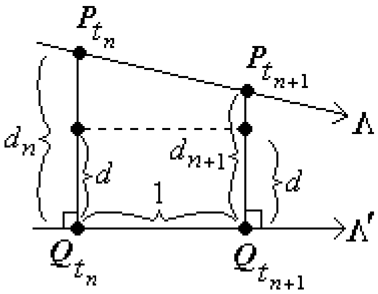



Figure 6.40:

□

We summarize the results of these lemmas in the following theorem.

**THEOREM 6.27.** *Let $\Lambda$ and $\Lambda'$ be two parallel lines. Let $p(t) = P_t$ be a ruler function for $\Lambda$ consistent the orientation $(A, B)$ of $\Lambda$, and for each $t$ let $Q_t$ be the point on $\Lambda'$ with $P_tQ_t \perp \Lambda'$.*

**(a)** *If there is a line perpendicular to both $\Lambda$ and $\Lambda'$ then $\Lambda$ and $\Lambda'$ are ultraparallel. In this case $\lim_{t\longrightarrow\infty} |P_tQ_t| = \lim_{t\longrightarrow-\infty} |P_tQ_t| = \infty$, and $|P_tQ_t|$ achieves its unique minimum when $P_tQ_t$ is the unique segment perpendicular to both $\Lambda$ and $\Lambda'$.*

**(b)** *If there is no line perpendicular to both $\Lambda$ and $\Lambda'$ then there is an orientation $(C, D)$ of $\Lambda'$ so that ${}_{(A,B)}\Lambda$ and ${}_{(C,D)}\Lambda'$ are asymptotically parallel. In this case $\lim_{t\longrightarrow\infty} |P_tQ_t| = 0$ while $\lim_{t\longrightarrow-\infty} |P_tQ_t| = \infty$.*

## Exercises

**6.52**. Show that if ${}_{(A,B)}\Lambda$ and ${}_{(C,D)}\Lambda'$ are distinct asymptotically parallel oriented lines then neither orientation of $\Lambda'$ is asymptotically parallel to ${}_{(B,A)}\Lambda$.

**6.53**. Prove the other half of Theorem 6.25 by showing that two lines with a common perpendicular must be ultraparallel.

**6.54**. Assume that $\Lambda$ and $\Lambda'$ are ultraparallel lines with common perpendicular $PQ$ (where $P \in \Lambda$ and $Q \in \Lambda'$). Prove that if $R$ is any point on $\Lambda$ other than $P$ then the distance from $R$ to $\Lambda'$ is greater than $|PQ|$.

**6.55**. In the halfplane model let $\Lambda$ be the H-line $x = 1$ and $\Lambda'$ the H-line $x = 0$. Then $\Lambda$ and $\Lambda'$ are asymptotically parallel. (It should be clear that they have no common perpendicular!)

**(a)** Let $P_y$ be the point $(1, y)$ on $\Lambda$. Find the coordinates for the point $Q_y$ on $\Lambda$ so that $P_yQ_y \perp \Lambda'$.

**(b)** Use the formula for H-distance (p.216) to find $d_y = d_H(P_y, Q_y)$.

**(c)** Show that $\lim_{y \longrightarrow 0} d_y = \infty$ and $\lim_{y \longrightarrow \infty} d_y = 0$.

# I. Afterword

We can't close this chapter without a few words about what we *haven't* covered. In these few pages we have barely explored the front porch to the universe that is hyperbolic geometry. You may wish to sometime venture further into the subject through a more advanced course of study, and there are many roads such an exploration could follow.

For instance, the models themselves (not just the halfplane model we have used, but also Beltrami's model and Poincaré's other model[6]) present a wealth of material for study. Computations here make heavy use of tools from calculus and the geometry of complex numbers.

In the geometry itself there are several interesting objects we have not introduced here. In fact, hyperbolic geometry abounds with surprisingly peculiar analogs to mostly mundane Euclidean constructs. For instance, in Euclidean geometry it is easy to define two families $\mathcal{X}$ and $\mathcal{Y}$ of lines that we can think of as the "horizontal lines" and the "vertical lines" – simply choose any two lines $\Lambda_x$ and $\Lambda_y$ with $\Lambda_x \perp \Lambda_y$ and let $\mathcal{X}$ be the family of all lines parallel to $\Lambda_x$ and $\mathcal{Y}$ be the family of all lines parallel to $\Lambda_y$. In hyperbolic geometry the situation is much more complex, and in many ways much more interesting. We can let $\mathcal{X}$ be the ideal point corresponding to a given oriented line, but then no line can be

[6]Poincaré's other model for hyperbolic geometry is known as the *disk model* because its underlying set is $\{(x, y) : x^2 + y^2 < 1\}$, an open disk in the plane.

perpendicular to more than one member of $\mathcal{X}$. However, there is a family $\mathcal{H}$ of *curves* such that:

- each point of the plane belongs to exactly one curve in $\mathcal{H}$, and
- each curve in $\mathcal{H}$ meets all lines in $\mathcal{X}$ at right angles.

These curves are called *horocycles* and play a critical role in advanced analysis of hyperbolic geometry. Figure 6.41 shows a half-plane model depiction of the family of horocycles corresponding to an ideal point.

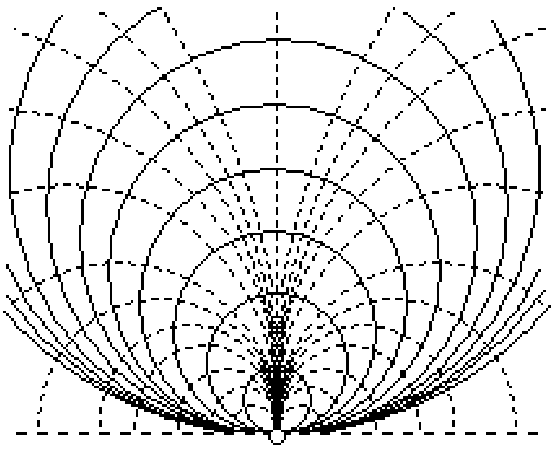


Figure 6.41: An ideal point $\mathcal{X}$ (dashed lines) and the corresponding family $\mathcal{H}$ of horcycles

One other topic deserves mention, and that is the concept of area. You might have been wondering if area can be defined in hyperbolic geometry. The answer is yes, but it certainly can't be the same as in our Area Axiom from Chapter 5. Indeed, in Euclidean geometry we use rectangles as the fundamental shape for computing areas. But rectangles don't even exist in hyperbolic geometry! The key, it turns out, is the hyperbolic concept of defect. The fundamental shape for measuring areas in hyperbolic geometry is the triangle, and the area of a triangle is proportional to its defect. The more a triangle's angle sum deviates from 180, the greater the hyperbolic area of the triangle. (Note how Theorem 6.4 shows that defect behaves the way we expect area to behave with regard to decomposing a shape into pieces.) This has the rather surprising consequence that, although the area of the hyperbolic plane itself is infinite, there is an upper bound on the area of triangles! In fact, this phenomenon characterizes hyperbolic geometry: the statement

*Given any number $m$ there is a triangle with area greater than $m$.*

could be added to our Theorem 6.7 list of assumptions equivalent to the parallel postulate.

Whether you pursue your study of hyperbolic geometry further or not, you should come away from your study of this chapter with three things:

- an appreciation for the long struggle that led to the discovery of hyperbolic geometry,
- an understanding of the relationship between hyperbolic and Euclidean geometry (in particular, why they are each "as real" as the other), and
- a familiarity with the basic terms and theorems of hyperbolic geometry as we have developed them here.

But now it's time to step back from the universe of hyperbolic geometry. Most of our proofs in Chapter 7 will be in the confines of neutral geometry as we explore the topic of geometric transformations. Then, having a working familiarity with both extensions of neutral geometry, we can enliven our work by comparing how the theory plays out in the Euclidean and hyperbolic realms.

# Chapter 7

# Transformations

Recall that in Chapter 3 we discussed some shortcomings of Euclid's axiom system in the *Elements*, and that one of them was his use of "superposition" to justify the SAS congruence criterion (see p.114). Superposition is the notion that two congruent figures will exactly coincide if one of them is moved so as to overlay the other. That this notion is not supported by Euclid's axioms (and thus was an inappropriate tool in that setting) does not detract from the fact that it really *is* the way we imagine congruence in our minds. We intuitively think of congruence as meaning that two objects "have the same shape", or are "copies of one another", and we routinely perform "mental superposition" on geometric figures to check for the appearance of congruence. It seems to be fundamental to the way we think.

You can view this chapter's material as an effort to put superposition on solid mathematical footing. We'll develop the theory of "motions of the plane" properly, from the axioms. The mathematical model for such a motion is a special type of function from the plane to itself called a *transformation*. We'll study the form these transformations can take and find a beautiful body of mathematics underlying their structure.

So let's start with a definition. How, exactly, can we give mathematical embodiment of the mental image we have for a "motion of the plane"? First, we want each point $P$ to be moved to a unique point $P'$, and the function $f(P) = P'$ describing this motion should be both one-to-one and onto. (No two different points should be moved to the same place, and each point of the plane should be the image of some other point under this motion.) The mathematical term for a one-to-one and onto function is a *bijection*. So our transformations should be bijections from the plane to itself. But certainly not *all* such bijections should qualify as transformations. Intuitively, we want to insist that the motion described by the bijection is "rigid" – though it may move a triangle to a new location, it shouldn't distort the triangle's shape or size. The following definition will meet our expectations:

**DEFINITION.** A *transformation of the plane* (or *transformation* for short) is a bijection of the plane to itself such that for every triangle $ABC$ we have $\triangle ABC \cong \triangle f(A)f(B)f(C)$.

Note that this definition has the following two immediate consequences. If $f$ is a transformation, then

- given two points $A$ and $B$, the distance $|f(A)f(B)|$ is equal to the distance $|AB|$, and
- given a third point $C$ we have $m\angle CAB$ equal to $m\angle f(C)f(A)f(B)$.

These properties are summarized by saying that $f$ "*preserves distances*" and "*preserves angle measures*". It's easy to see (by the SAS congruence criterion) that any bijection that preserves both distances and angle measures will carry triangles to congruent copies of themselves, so:

**FACT 7.1.** *A bijection from the plane to itself is a transformation if and only if it preserves distances and angle measures.* □

**Content of this chapter**

In this chapter we'll pursue the general goal of investigating the nature of transformations. What types of transformations are there? Is there a simple way to describe them? Is there a way to easily identify the type of a given transformation? What do transformations "look like" in the models? We'll try to answer all of these. Most of the time we'll be operating on two distinct tracks:

- First, each section will present a step in the development of the general theory of transformations. It's important to note that we'll be working in neutral geometry – our general theory will thus be applicable to transformations in both the Euclidean and hyperbolic planes.
- Second, whenever we introduce a new type of transformation, we'll pause to compute examples of them in both the standard Euclidean model and the upper halfplane hyperbolic model. These example calculations should

help solidify your understanding of the theory, but they also serve another purpose: they illustrate the importance of geometry to modern technology. Fields like computer graphics and computer-aided design make heavy use of transformations, and quick calculation of them is critical. If you've been wondering where the modern applications of geometry might be, this is your moment! Pay close attention in our examples to how the geometric *theory* is critical to making the calculations possible.

We'll end the chapter with two sections devoted to characterizing all of the possible transformations in both Euclidean and hyperbolic geometries. That's a good achievement on which to end our text – it's impressive to see the heights to which just a few axioms have brought us!

**A note on notation.**

Before we begin, a few comments are in order regarding the notation we'll use for our Euclidean model. You should recall (see p.96) that the underlying point set for that model is the Cartesian $xy$-coordinate plane, which we'll now denote by $\mathbb{R}^2$. A transformation of the Euclidean plane, then, will be a function $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$. We'll use matrix notation for these transformations.

You're probably accustomed to denoting a point of $\mathbb{R}^2$ by an ordered pair such as $(x, y)$. We'll want to instead use "column vectors" of the form $\begin{pmatrix} x \\ y \end{pmatrix}$ to denote points of $\mathbb{R}^2$. That way we can use $2 \times 2$ matrices, matrix multiplication, and vector algebra to write our functions. Our Euclidean transformations will be functions from $\mathbb{R}^2$ to $\mathbb{R}^2$ of the form

$$\begin{aligned} f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) &= \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} \\ &= \begin{pmatrix} ax+by \\ cx+dy \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} \\ &= \begin{pmatrix} ax+by+e \\ cx+dy+f \end{pmatrix}. \end{aligned}$$

For writing transformations in the hyperbolic (halfplane) model we'll use the familiar ordered pair notation. So transformations of the hyperbolic plane will take the form of functions $f(x, y) = (f_1(x, y), f_2(x, y))$ where $f_1$ and $f_2$ are functions from $\mathbb{R}^2$ to $\mathbb{R}$. Since the underlying set of the halfplane model includes only those points in $\mathbb{R}^2$ with positive $y$ coordinate, we will assume $y$ is always positive and insist that $f_2$ always return positive values.

# A. Basic Facts

We'll do three things in this section.

- First, we'll show a simpler description of transformations – namely, any function from the plane to itself that preserves distances is a transformation.
- We'll develop some basic facts about the algebra of transformations.
- Finally, we'll give a criterion for showing that two transformations are equal.

### Isometries

As just noted in the chapter introduction, a transformation is a bijection from the plane to itself that preserves distances and angle measures. In the general setting of any metric space (a set of points on which there is a distance function defined), a function from the space to itself that preserves distances is called an *isometry*. Thus, every transformation of the plane is an isometry. More remarkably, the converse holds: any isometry from the plane to itself is automatically a transformation.

It's easy to see that isometries are one-to-one (see Exercise 7.1). It remains, then, to prove that isometries of the plane are onto functions and that they preserve angle measure. We'll start with the latter of these two.

LEMMA **7.2.** *If $f$ is an isometry of the plane and $A$, $B$, and $C$ are points, then $\angle CAB \cong \angle f(C)f(A)f(B)$.*

*Proof:* This is actually very easy:

- Since $f$ is an isometry we have $\triangle f(A)f(B)f(C) \cong \triangle ABC$ by the SSS congruence criterion.
- So $\angle f(C)f(A)f(B)$ and $\angle CAB$ are corresponding angles in congruent triangles, and are thus congruent. □

A consequence of this lemma is that isometries preserve collinearity: if $C$ is on $\overleftrightarrow{AB}$ then $f(C)$ will be on $\overleftrightarrow{f(A)f(B)}$.

Now we'll work toward showing that isometries of the plane are onto functions. We'll begin by showing that isometries from $\mathbb{R}$ to $\mathbb{R}$ are onto.

**Lemma 7.3.** *Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be an isometry (so that $|f(s) - f(t)| = |s - t|$ for all $s$ and $t$) and let $a = f(0)$. Then either $f(t) = a + t$ or $f(t) = a - t$. In particular, $f$ is onto $\mathbb{R}$.*

*Proof:* Since $f$ is an isometry, for each $t$ we have

$$|f(t) - a| = |f(t) - f(0)| = |t - 0| = |t|,$$

so either $f(t) = a + t$ or $f(t) = a - t$. So $f(1)$ is either $a + 1$ or $a - 1$. We claim that if $f(1) = a + 1$ then $f(t) = a + t$ for all $t$. (Similar reasoning to the steps below show that if $f(1) = a - 1$ then $f(t) = a - t$ for all $t$.)

- Assume that $f(1) = a + 1$.
- Suppose $r$ is such that $f(r) = a - r$. Then:

$$|r - 1| = |f(r) - f(1)| = |(a - r) - (a + 1)| = |-r - 1| = |r + 1|.$$

- This clearly implies that $r = 0$, so we actually have $f(r) = a - r = a + r$.
- Thus $f(t) = a + t$ for all values of $t$. □

**Lemma 7.4.** *Let $f$ be an isometry of the plane to itself and let $A$ and $B$ be distinct points. Then $f$ carries the line $\overleftrightarrow{AB}$ onto the line $\overleftrightarrow{f(A)f(B)}$.*

*Proof:* We know from the observation following Lemma 7.2 that $f$ carries the points of $\overleftrightarrow{AB}$ *into* $\overleftrightarrow{f(A)f(B)}$. We need only show that for each point $S$ on $\overleftrightarrow{f(A)f(B)}$ there is a point $X$ on $\overleftrightarrow{AB}$ with $f(X) = S$.

- Find a ruler function $p(t) = P_t$ for $\overleftrightarrow{AB}$ so that $P_0 = A$ and $P_{|AB|} = B$.
- Likewise, find a ruler function $q(t) = Q_t$ for $\overleftrightarrow{f(A)f(B)}$ so that $Q_0 = f(A)$ and $Q_{|f(A)f(B)|} = Q_{|AB|} = f(B)$.
- We have functions $p : \mathbb{R} \longrightarrow \mathbb{R}^2$, $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, and $q^{-1} : \mathbb{R}^2 \longrightarrow \mathbb{R}$, so we can define the composition $j = q^{-1} \circ f \circ p : \mathbb{R} \longrightarrow \mathbb{R}$. This is an isometry from $\mathbb{R}$ to $\mathbb{R}$:

$$\begin{aligned}
|j(s) - j(t)| &= |q^{-1}(f(P_s)) - q^{-1}(f(P_t))| \\
&= |f(P_s)f(P_t)| && \text{(since } q \text{ is a ruler function)} \\
&= |P_sP_t| && \text{(since } f \text{ is an isometry)} \\
&= |s - t| && \text{(since } p \text{ is a ruler function)}
\end{aligned}$$

- So, by Lemma 7.3, either $j(t) = j(0) + t$ or $j(t) = j(0) - t$.
- But we can easily compute

$$j(0) = q^{-1}(f(P_0)) = q^{-1}(f(A)) = 0\,, \text{ and}$$

$$j(|AB|) = q^{-1}(f(P_{|AB|})) = q^{-1}(f(B)) = |AB|\,.$$

- So, $j(t) = t$ for every $t$. In other words, $f(P_t) = Q_t$ for all $t$.
- Since $S = Q_s$ for some number $s$, we have $S = f(P_s)$. □

LEMMA **7.5.** *Every isometry of the plane is an onto function.*

*Proof:* Let $f$ be an isometry from the plane to itself and let $S$ be a point of the plane. We must show there is some point $X$ with $f(X) = S$.

- Choose two points $A$ and $B$ and find their images $f(A)$ and $f(B)$.
- If $S$ is on $\overleftrightarrow{f(A)f(B)}$ then we are done by Lemma 7.4, so we may assume that $S$, $f(A)$, and $f(B)$ are noncollinear.

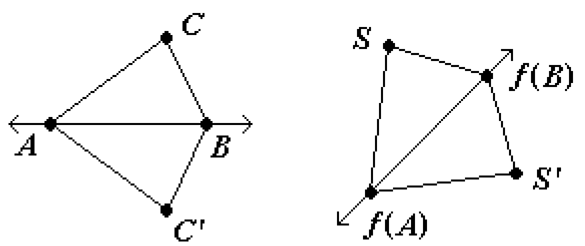



Figure 7.1:

- By Fact 4.9 there are two points $C$ and $C'$ (one on each side of $\overleftrightarrow{AB}$) so that $\triangle ABC \cong \triangle ABC' \cong \triangle f(A)f(B)S$ (see Figure 7.1).
- Since $f$ is an isometry, the SSS congruence criterion implies that $\triangle f(A)f(B)f(C) \cong \triangle ABC$.
- By transitivity of congruence, $\triangle f(A)f(B)f(C) \cong \triangle f(A)f(B)S$.
- Similarly, we may conclude that $\triangle f(A)f(B)f(C') \cong \triangle f(A)f(B)S$.

- But again by Fact 4.9, if $Y$ is a point so that $\triangle f(A)f(B)Y \cong \triangle f(A)f(B)S$ then either $Y = S$ or $Y$ is a uniquely determined point $S'$ on the opposite side of $\overleftrightarrow{f(A)f(B)}$ from $S$.
- So, either $f(C) = S$ or $f(C') = S$. $\square$

These lemmas together with Fact 7.1 accomplish our first goal.

**THEOREM 7.6.** *Every isometry of the plane is a transformation.* $\square$

### Combining transformations

The mental image of a transformation is that of a rigid motion of the plane. It seems clear from that mental image that the net result of following one rigid motion by another is itself a rigid motion of the plane. Thus, we would expect that the composition of two transformations is itself a transformation.

Another fact suggested by the mental image of plane motions relates to "undoing" such a motion. To each plane motion moving $Q$ to $P$ there is a reverse motion that moves $P$ back to $Q$. If the motion corresponds to the transformation $f$, then the reverse motion corresponds to the *inverse* for $f$ given by the rule

$$f^{-1}(P) = Q \text{ where } Q \text{ is the point such that } f(Q) = P.$$

This definition is perfectly valid because $f$ is a bijection – for each point $P$ there is exactly one point $Q$ with $f(Q) = P$. Our mental image suggests that $f^{-1}$ should itself be a transformation, and indeed it is.

**FACT 7.7.** *Let $f$ and $g$ be transformations. Then*

**(a)** *$f \circ g$ is a transformation, and*

**(b)** *$f^{-1}$ is also a transformation.*

*Proof:* We'll give the proof of part **(b)** here, leaving part **(a)** as an exercise. So, suppose $f$ is a transformation. To prove that $f^{-1}$ is a transformation we need only prove it is an isometry (by Theorem 7.6). Let $A$ and $B$ be any points. Then:

- Since $f$ is a bijection there are points $A'$ and $B'$ with $f(A') = A$ and $f(B') = B$.

- By the definition of $f^{-1}$ we have $f^{-1}(A) = A'$ and $f^{-1}(B) = B'$.
- Then since $f$ is a transformation (and thus an isometry) we have $|f^{-1}(A)f^{-1}(B)| = |A'B'| = |f(A')f(B')| = |AB|$. $\square$

Note that $f \circ f^{-1} = id$ is the *identity* transformation defined by $id(P) = P$ for all points $P$. If you've taken a course in modern abstract algebra, you probably recognize that the set of all transformations is really a *group* with composition of functions as its group operation and $id$ as its identity. Indeed, sets of geometric transformations (often called *symmetries*) were primary motivating examples in the development of group theory.

**Fixed points and fixed sets**

A productive approach to characterizing a transformation $f$ is to examine what points (if any) are not moved by $f$, and what sets (if any) are carried by $f$ to themselves.

NOTATION. If $f$ is a function from the plane to itself and $\Sigma$ is a set then we use the symbol $f(\Sigma)$ to denote the set $\{f(P) : P \in \Sigma\}$.

DEFINITIONS. If $f$ is a function from the plane to itself, the point $P$ is a *fixed point* for $f$ if $f(P) = P$. The set $\Sigma$ is a *fixed set* for $f$ if every point of $\Sigma$ is a fixed point for $f$. If $f(\Sigma) = \Sigma$ then we say that $\Sigma$ is a *preserved set* for $f$.

Note that every fixed set for $f$ is a preserved set for $f$ (but not vice versa). We'll say that $f$ *fixes* $P$ if $P$ is a fixed point for $f$. Similarly we might say that $f$ *fixes* $\Sigma$ or *preserves* $\Sigma$.

We'll start with a simple observation on preserved circles. It's proof is left as an exercise.

THEOREM **7.8.** *Let $\Gamma$ be a circle with center $P$ and let $f$ be a transformation. Then $f$ fixes $P$ if and only if it preserves $\Gamma$.*

Next we show that if $f$ fixes two points on a line then it fixes the entire line. Note that this means that a transformation will have either zero, one, or infinitely many fixed points.

**Theorem 7.9.** *If $f$ is a transformation fixing two distinct points $A$ and $B$, then $f$ fixes $\overleftrightarrow{AB}$.*

*Proof:* This follows immediately from the proof of Lemma 7.4. For if $f(A) = A$ and $f(B) = B$ then the two ruler functions $p(t) = P_t$ and $q(t) = Q_t$ must be identical by the uniqueness clause in Axiom L2. So $f(P_t) = Q_t = P_t$ for all values of $t$. $\square$

Our next lemma simply steps this up one level: if $f$ has three distinct fixed points, not all on a line, then $f$ fixes the entire plane.

**Theorem 7.10.** *If $f$ is a transformation that fixes a noncollinear set $\{A, B, C\}$ then $f = id$.*

*Proof:* Let $X$ be any point in the plane. We must show that $f(X) = X$ (see Figure 7.2).

- By Theorem 7.9 we see that $f$ fixes the union of the lines $\overleftrightarrow{AB}$, $\overleftrightarrow{BC}$, and $\overleftrightarrow{AC}$.
- Let $D$ be any point in the interior of triangle $ABC$.
- The line $\overleftrightarrow{DX}$ must meet two of these lines (see Fact 4.8), and so will contain two points fixed by $f$.
- But then Theorem 7.9 applied to $\overleftrightarrow{DX}$ shows that $\overleftrightarrow{DX}$ is entirely fixed by $f$.
- In particular, then, $X$ is fixed by $f$. $\square$

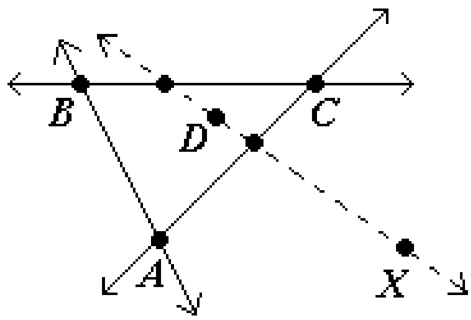



Figure 7.2:

Our concluding theorem is that a transformation is determined by its behavior on any noncollinear set: if $f$ and $g$ are transformations that agree on three noncollinear points, then they are in fact identical. Note that Theorem 7.10 is a special case of this theorem – simply take $g = id$. We leave the easy proof as Exercise 7.4.

**THEOREM 7.11.** *Suppose $f$ and $g$ are transformations and that there is a noncollinear set $\{A, B, C\}$ such that $f(A) = g(A)$, $f(B) = g(B)$, and $f(C) = g(C)$. Then $f = g$.*

## Exercises

**7.1**. Prove that every isometry is one-to-one.

**7.2**. Prove part **(a)** of Fact 7.7.

**7.3**. Prove Theorem 7.8.

**7.4**. Prove Theorem 7.11. (*Hint:* consider the transformation $f \circ g^{-1}$.)

**7.5**. (For readers who have had a course in modern abstract algebra.) Is the group of all transformations an Abelian (commutative) group?

**7.6**. Suppose $f$ is a transformation such that $f \circ f$ has a fixed point. Show that $f$ also has a fixed point.

**7.7**. Given a number $\theta$ and a vector $\begin{pmatrix} a \\ b \end{pmatrix}$ define the function $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ by the rule

$$f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix}.$$

Show that $f$ is a transformation in the usual model for Euclidean geometry. (Remember to use Theorem 7.6.)

**7.8**. Write the transformations described below in the form given in Exercise 7.7.

**(a)** $f$ carries the triangle with vertices $\left\{\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}\right\}$ to the triangle with vertices $\left\{\begin{pmatrix} -2 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix}\right\}$.

**(b)** $f$ carries the triangle with vertices $\left\{\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \end{pmatrix}\right\}$ to the triangle with vertices $\left\{\begin{pmatrix} 1 \\ -3 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix}\right\}$.

# B. Reflections

In this section we'll introduce a family of transformations that form the building blocks from which all transformations can be constructed.

**Definition.** Let $\Lambda$ be a line. We define a function $r_\Lambda$ from the plane to itself as follows:

- If $P$ is on $\Lambda$ then $r_\Lambda(P) = P$.
- Otherwise, $r_\Lambda(P)$ is the unique point such that $\Lambda$ is the perpendicular bisector of $r_\Lambda(P)P$.

This function $r_\Lambda$ is called the *reflection* across (or through) $\Lambda$.

The mental image for the transformation $r_\Lambda$ is a "flip" of the plane across the line $\Lambda$. (Imagine rotating the plane 180° in 3-space using $\Lambda$ as the axis of rotation, as in Figure 7.3.) We'll see in Section D that all transformations can be expressed as the composition of reflections, so in fact, all motions of the plane can be accomplished by a sequence of these "flips".

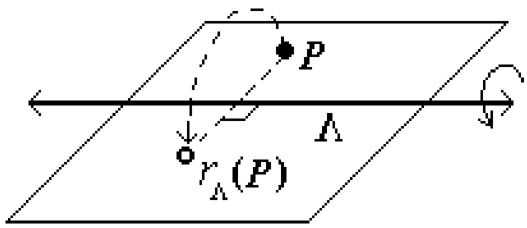



Figure 7.3: The mental image for a reflection across a line $\Lambda$.

We'll prove two theorems about reflections. The first simply states that $r_\Lambda$ really is a transformation. The second is more significant; it shows that $r_\Lambda$ is the only non-identity transformation fixing $\Lambda$.

**Theorem 7.12.** *The function $r_\Lambda$ is a transformation that fixes the line $\Lambda$ and preserves all lines perpendicular to $\Lambda$. Furthermore, $r_\Lambda$ is its own inverse – that is, $r_\Lambda \circ r_\Lambda = id$.*

*Proof:* It is clear from the definition that $r_\Lambda$ fixes $\Lambda$, preserves lines perpendicular to $\Lambda$, and that $r_\Lambda \circ r_\Lambda = id$. Thus we need only show that $r_\Lambda$ is a transformation. By Theorem 7.6 it is enough for us to show that $r_\Lambda$ is an isometry. So, let $A$ and $B$ be any two points. We must show that $|r_\Lambda(A)\, r_\Lambda(B)| = |AB|$. There are four cases to consider.

**Case 1:** If $A$ and $B$ are both on $\Lambda$ then the conclusion follows immediately because $r_\Lambda$ fixes all points on $\Lambda$.

**Case 2:** We leave the case where $A$ is on $\Lambda$ and $B$ is not as part of Exercise 7.10.

**Case 3:** Suppose that $A$ and $B$ are on opposite sides of $\Lambda$. Refer to Figure 7.4 in the following steps.

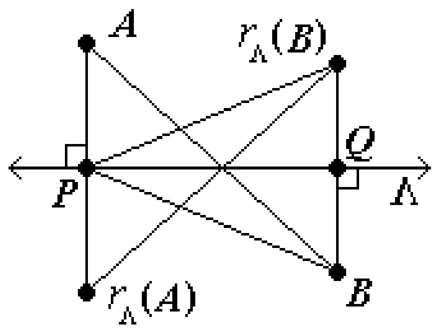



Figure 7.4: The mental image for the reflection $r_\Lambda$.

- Let $P$ and $Q$ be points of $\Lambda$ so that $AP \perp \Lambda$ and $BQ \perp \Lambda$.
- If $P = Q$ then the result follows easily since then $|r_\Lambda(A)\, r_\Lambda(B)| = |r_\Lambda(A)\, P| + |r_\Lambda(B)\, P| = |AP| + |BP| = |AB|$. So we may assume that $P \neq Q$.
- Since $\angle PQB \cong \angle PQ\, r_\Lambda(B)$ (both are right angles) and $|QB| = |Q\, r_\Lambda(B)|$, the triangles $\triangle PQB$ and $\triangle PQ\, r_\Lambda(B)$ are congruent by the SAS criterion.
- So, $|PB| = |P\, r_\Lambda(B)|$ and $\angle QPB \cong \angle QP\, r_\Lambda(B)$.
- $m\angle APB = 90 + m\angle QPB = 90 + m\angle QP\, r_\Lambda(B) = m\angle r_\Lambda(A)P\, r_\Lambda(B)$.
- So we have $\angle APB \cong \angle r_\Lambda(A)P\, r_\Lambda(P)$, $PB \cong P\, r_\Lambda(B)$, and $PA \cong P\, r_\Lambda(A)$, so $\triangle APB \cong \triangle r_\Lambda(A)\, P\, r_\Lambda(B)$ by the SAS criterion.
- Then $AB$ and $r_\Lambda(A)\, r_\Lambda(B)$ are congruent, since they are corresponding sides in these triangles.

**Case 4:** The proof is similar when $A$ and $B$ are on the same side of $\Lambda$, and is left to Exercise 7.10. $\square$

**Theorem 7.13.** *Let $f$ be any transformation fixing the line $\Lambda$. Then either $f = id$ or $f = r_\Lambda$.*

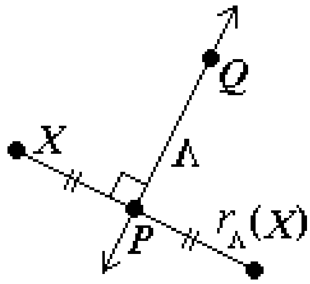



Figure 7.5:

*Proof:* Let $X$ be any point not on $\Lambda$, and let $P$ and $Q$ be points on $\Lambda$ with $XP \perp \Lambda$ (see Figure 7.5).

- Since $f$ is a transformation we must have $\triangle PQX \cong \triangle PQf(X)$.
- But $\angle XPQ$ is a right angle, so $\angle f(X)PQ$ must also be a right angle.
- So $f(X)$ is on the perpendicular line to $\Lambda$ at $P$. In other words, $f(X)$ is on $\overleftrightarrow{PX}$.
- Also we have $|XP| = |f(X)P|$, so either $f(X) = X$ or else $P$ is the midpoint of $f(X)X$. Rephrasing, either $f(X) = X$ or $f(X) = r_\Lambda(X)$.
- Now $f$, $id$, and $r_\Lambda$ all fix both $P$ and $Q$, so either $f$ agrees with $id$ on the noncollinear set $\{P, Q, X\}$ or it agrees with $r_\Lambda$ on that set.
- By Theorem 7.11 then, either $f = id$ or $f = r_\Lambda$. $\square$

**Reflections in the Euclidean model**

We'll now turn our attention to finding actual formulae for reflections of the Euclidean plane using the usual model. As outlined in the chapter introduction, we'll use matrix and vector notation for convenience.

First let's consider reflection across a line through the origin of $\mathbb{R}^2$. We'll handle the case of a vertical line in a moment, so for now consider a line $\Lambda$ with equation $y = mx$. Let $\begin{pmatrix} a \\ b \end{pmatrix}$ be any point in $\mathbb{R}^2$ – we'll find a formula for the point $\begin{pmatrix} a' \\ b' \end{pmatrix} = r_\Lambda\left(\begin{pmatrix} a \\ b \end{pmatrix}\right)$ obtained by reflecting $\begin{pmatrix} a \\ b \end{pmatrix}$ through $\Lambda$.

- As seen from Figure 7.6, this point will lie on the line through $\begin{pmatrix} a \\ b \end{pmatrix}$ with slope $-1/m$ (perpendicular to the slope of $\Lambda$).

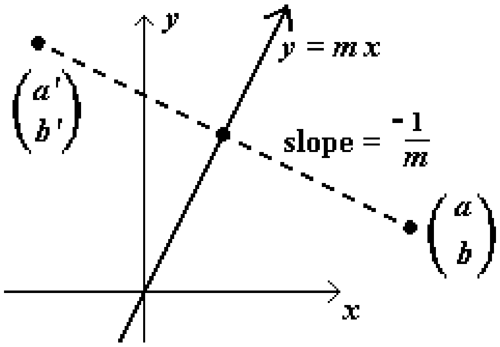



Figure 7.6: Finding coordinates for image of a point under $r_\Lambda$ when $\Lambda$ is $y = mx$.

- The equation of this line is

$$y = \frac{-1}{m}(x-a) + b\,.$$

- Using the two line equations we can find the $x$-coordinate of the point where the lines meet:

$$\begin{aligned} mx &= \frac{-1}{m}(x-a)+b \\ m^2x &= -x+a+mb \\ (m^2+1)x &= a+mb \\ x &= \frac{a+mb}{m^2+1} \end{aligned}$$

- Since it lies on the line $y = mx$, the $y$-coordinate of this point must be $m\frac{a+mb}{m^2+1}$.

- So,[1] $\frac{a+mb}{m^2+1}\begin{pmatrix} 1 \\ m \end{pmatrix}$ must be the midpoint between $\begin{pmatrix} a' \\ b' \end{pmatrix}$ and $\begin{pmatrix} a \\ b \end{pmatrix}$.

- We can use this to solve for $a'$ and $b'$:

$$\frac{a'+a}{2} = \frac{a+mb}{m^2+1} \quad \Longrightarrow \quad a' = \frac{1}{m^2+1}[2a+2mb-(m^2+1)a], \text{ and}$$

$$\frac{b'+b}{2} = m\frac{a+mb}{m^2+1} \quad \Longrightarrow \quad b' = \frac{1}{m^2+1}[2ma+2m^2b-(m^2+1)b]\,.$$

[1]Here, we've simplified the name of this point by factoring out the constant $\frac{a+mb}{m^2+1}$ from both $x$- and $y$-coordinates. The usual algebra rules for vectors apply: $c\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} cx \\ cy \end{pmatrix}$.

- This means

$$\begin{pmatrix} a' \\ b' \end{pmatrix} = \frac{1}{m^2+1}\begin{pmatrix} a - m^2a + 2mb \\ 2ma + m^2b - b \end{pmatrix}$$
$$= \frac{1}{m^2+1}\begin{pmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}$$

- Now the point $\begin{pmatrix} a \\ b \end{pmatrix}$ was arbitrary, so we can replace it with the generic $\begin{pmatrix} x \\ y \end{pmatrix}$ to obtain the formula

> If $\Lambda$ is the line $y = mx$ then
> $$r_\Lambda\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \frac{1}{m^2+1}\begin{pmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

Reflection across the vertical line $x = 0$ is easy – the point $\begin{pmatrix} x \\ y \end{pmatrix}$ is simply sent to the point $\begin{pmatrix} x \\ -y \end{pmatrix}$, giving us the following formula:[2]

> If $\Lambda$ is the line $x = 0$ then
> $$r_\Lambda\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} -x \\ y \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

Now what about lines *not* through the origin? We could determine the rule for such reflections by a direct calculation, similar to the way we did the $y = mx$ case above, but there's an easier way. Consider the line $\Lambda$ with equation $y = m(x-d)$ (with slope $m$ and $x$-intercept $d$). We'll accomplish reflection of the point $\begin{pmatrix} x \\ y \end{pmatrix}$ through $\Lambda$ in three steps[3], illustrated in Figure 7.7:

---

[2] It's interesting to note that the formula for reflection through the vertical line $x = 0$ can be viewed as the limit (as the slope $m$ tends to infinity) of the formula for reflection through $y = mx$: the limit of $\frac{1-m^2}{m^2+1}$ is $-1$, the limit of $\frac{m^2-1}{m^2+1}$ is 1, and the limit of $\frac{2m}{m^2+1}$ is 0.

[3] If you've had a course in modern abstract algebra, you might recognize these three steps as *conjugation.* If $f$ represents horizontal motion by distance $d$ then $r_\Lambda = f \circ r_{\Lambda'} \circ f^{-1}$, so $r_\Lambda$ is a conjugate of $r_{\Lambda'}$.

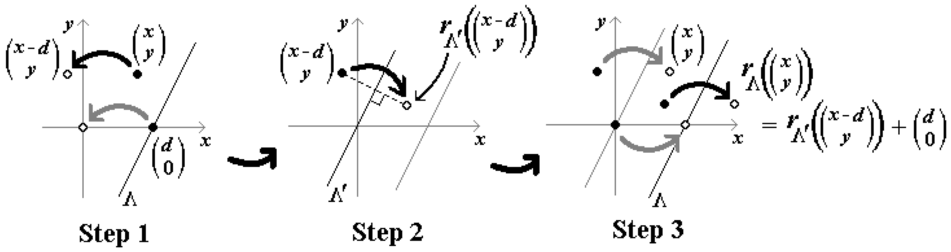



Figure 7.7: Finding $r_\Lambda\left(\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)\right)$ by the three steps explained below.

**Step 1:** First translate the entire plane $d$ units horizontally so that the point $\left(\begin{smallmatrix} d \\ 0 \end{smallmatrix}\right)$ is moved to the origin and $\Lambda$ is moved to the line $\Lambda'$ with equation $y = mx$. This moves $\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ to $\left(\begin{smallmatrix} x-d \\ y \end{smallmatrix}\right)$.

**Step 2:** Reflect through the line $\Lambda'$. The point $\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ has now been moved to $r_{\Lambda'}\left(\left(\begin{smallmatrix} x-d \\ y \end{smallmatrix}\right)\right)$ by these two successive motions.

**Step 3:** Now move the plane $d$ units horizontally in the other direction, so that the origin is moved back to $\left(\begin{smallmatrix} d \\ 0 \end{smallmatrix}\right)$ and the line of reflection is moved back to the original location of $\Lambda$. The end result of all three motions on $\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ is $r_{\Lambda'}\left(\left(\begin{smallmatrix} x-d \\ y \end{smallmatrix}\right)\right) + \left(\begin{smallmatrix} d \\ 0 \end{smallmatrix}\right)$.

Since $\Lambda'$ is a line $y = mx$, we can use the formula on p.265 to calculate

$$\begin{aligned} r_\Lambda\left(\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)\right) &= r_{\Lambda'}\left(\left(\begin{smallmatrix} x-d \\ y \end{smallmatrix}\right)\right) + \left(\begin{smallmatrix} d \\ 0 \end{smallmatrix}\right) \\ &= \frac{1}{m^2+1}\left(\begin{smallmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{smallmatrix}\right)\left(\begin{smallmatrix} x-d \\ y \end{smallmatrix}\right) + \left(\begin{smallmatrix} d \\ 0 \end{smallmatrix}\right) \\ &= \frac{1}{m^2+1}\left(\begin{smallmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{smallmatrix}\right)\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) + \frac{2md}{m^2+1}\left(\begin{smallmatrix} m \\ -1 \end{smallmatrix}\right) \end{aligned}$$

(Here we have omitted some routine algebra in combining vectors. Be sure to verify these steps yourself!)

If $\Lambda$ is the line $y = m(x-d)$ then

$$r_\Lambda\left(\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)\right) = \tfrac{1}{m^2+1}\left(\begin{smallmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{smallmatrix}\right)\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) + \tfrac{2md}{m^2+1}\left(\begin{smallmatrix} m \\ -1 \end{smallmatrix}\right).$$

For vertical lines $x = d$ a similar process works: first move $d$ units horizontally, second reflect through $x = 0$, and third move back $d$ units horizontally. Tracing the effect on a vector $\begin{pmatrix} x \\ y \end{pmatrix}$, we have:

$$\begin{pmatrix} x \\ y \end{pmatrix} \xrightarrow{\textbf{Step 1}} \begin{pmatrix} x-d \\ y \end{pmatrix} \xrightarrow{\textbf{Step 2}} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x-d \\ y \end{pmatrix}$$

$$\xrightarrow{\textbf{Step 3}} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x-d \\ y \end{pmatrix} + \begin{pmatrix} d \\ 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 2d \\ 0 \end{pmatrix}$$

So:[4]

> If $\Lambda$ is the line $x = d$ then
> $$r_\Lambda\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 2d \\ 0 \end{pmatrix} = \begin{pmatrix} -x+2d \\ y \end{pmatrix}.$$

EXAMPLE **7.14.** *Find the image of the triangle with vertices* $\left\{\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \end{pmatrix}\right\}$ *under reflection through the line* $\Lambda$ *with equation* $2x + 3y = 6$.

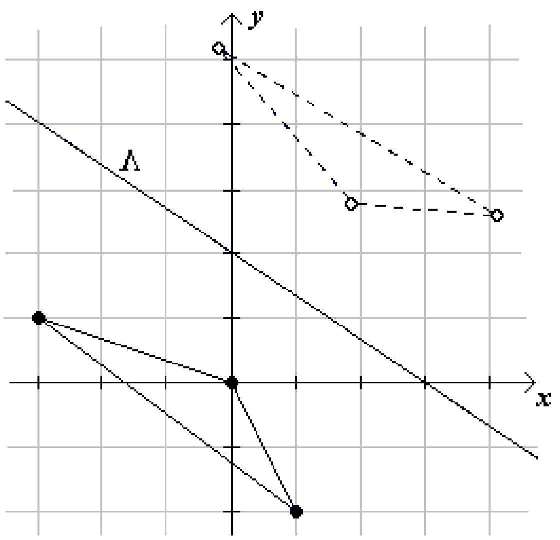



Figure 7.8:

[4]Again, notice that the formula for reflection through the vertical line $x = d$ is the limit as $m$ approaches infinity of the formula for the line $y = m(x - d)$.

*Solution:* First, write the equation of $\Lambda$ in the form $y = m(x-d)$:

$$y = -\frac{2}{3}x + 2 = -\frac{2}{3}(x-3)\,.$$

So, plugging in the values $m = -2/3$ and $d = 3$ to our formula, we have

$$\begin{aligned} r_\Lambda\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) &= \frac{9}{13}\begin{pmatrix} 5/9 & -4/3 \\ -4/3 & -5/9 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \frac{36}{13}\begin{pmatrix} 2/3 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 5/13 & -12/13 \\ -12/13 & -5/13 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 24/13 \\ 36/13 \end{pmatrix} \\ &= \frac{1}{13}\left[\begin{pmatrix} 5 & -12 \\ -12 & -5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 24 \\ 36 \end{pmatrix}\right] \end{aligned}$$

This allows us to quickly calculate:

$$r_\Lambda\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}\right) = \frac{1}{13}\left[\begin{pmatrix} 5 & -12 \\ -12 & -5 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 24 \\ 36 \end{pmatrix}\right] = \begin{pmatrix} 24/13 \\ 36/13 \end{pmatrix}$$

$$r_\Lambda\left(\begin{pmatrix} -3 \\ 1 \end{pmatrix}\right) = \frac{1}{13}\left[\begin{pmatrix} 5 & -12 \\ -12 & -5 \end{pmatrix}\begin{pmatrix} -3 \\ 1 \end{pmatrix} + \begin{pmatrix} 24 \\ 36 \end{pmatrix}\right] = \begin{pmatrix} -3/13 \\ 67/13 \end{pmatrix}$$

$$r_\Lambda\left(\begin{pmatrix} 1 \\ -2 \end{pmatrix}\right) = \frac{1}{13}\left[\begin{pmatrix} 5 & -12 \\ -12 & -5 \end{pmatrix}\begin{pmatrix} 1 \\ -2 \end{pmatrix} + \begin{pmatrix} 24 \\ 36 \end{pmatrix}\right] = \begin{pmatrix} 53/13 \\ 34/13 \end{pmatrix}$$

So, $r_\Lambda$ carries the given triangle to the triangle with the above listed points as vertices (see Figure 7.8). □

**EXAMPLE 7.15.** *A reflection $r_\Lambda$ carries the point $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ to the point $\begin{pmatrix} -4 \\ 3 \end{pmatrix}$. Where will it take the point $\begin{pmatrix} 0 \\ -2 \end{pmatrix}$?*

*Solution:* As is clear from Figure 7.9, the line of reflection $\Lambda$ must pass through $\begin{pmatrix} -1 \\ 2 \end{pmatrix}$ and have slope 3 (since it is the perpendicular bisector of the segment between $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} -4 \\ 3 \end{pmatrix}$). So $\Lambda$ has equation $y - 2 = 3(x+1)$, and setting $y = 0$ we find the $x$-intercept of $\Lambda$ to be $-5/3$. So we set $m = 3$ and $d = -5/3$ in the formula for $r_\Lambda$ on p.266, giving us

$$r_\Lambda\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \frac{1}{10}\begin{pmatrix} -8 & 6 \\ 6 & 8 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -3 \\ 1 \end{pmatrix}.$$

So,

$$r_\Lambda\left(\begin{pmatrix} 0 \\ -2 \end{pmatrix}\right) = \frac{1}{10}\begin{pmatrix} -8 & 6 \\ 6 & 8 \end{pmatrix}\begin{pmatrix} 0 \\ -2 \end{pmatrix} + \begin{pmatrix} -3 \\ 1 \end{pmatrix} = \begin{pmatrix} -21/5 \\ -3/5 \end{pmatrix}.$$

□

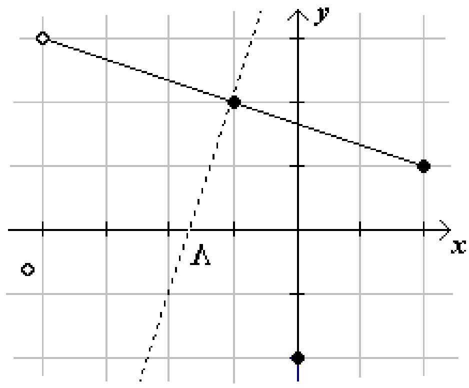



Figure 7.9:

**Reflections in the hyperbolic model**

Figure 7.10 depicts two cases of reflecting a point $P$ across an H-line. If the H-line is a vertical ray (as on the left) then the reflection agrees with the Euclidean reflection across that vertical line – something that is easily computed. However, if $\Lambda$ is one of the H-lines taking the form of a Euclidean semicircle, then it is not clear at all how to find $r_\Lambda(P)$. It will certainly be on the H-line through $P$ that is perpendicular to $\Lambda$, and its H-distance to $\Lambda$ should be equal to the H-distance from $P$ to $\Lambda$. But since H-distance is difficult to compute (recall the formula on p.216), its location on that H-line is not immediately obvious.

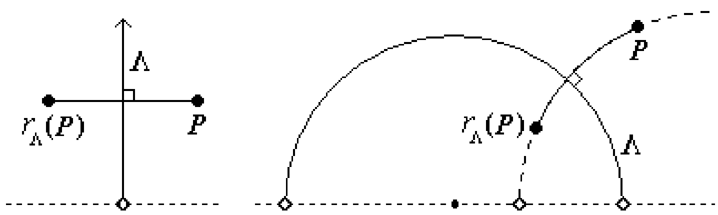



Figure 7.10: Reflecting through an H-line

Fortunately, $r_\Lambda(P)$ *is* easy to find. But to describe its location we'll need a detour to a seemingly unrelated Euclidean concept.

**Definition.** Let $\Gamma$ be a circle with center $C$ and radius $r$ and let $P$ be a point other than $C$. We define the *reflection of $P$ through* $\Gamma$ to be the unique point $P'$ on $\overrightarrow{CP}$ so that $|CP||CP'| = r^2$ (see Figure 7.11).

See Exercise 7.12 for a straightedge and compass construction of the reflection of a point through a circle. Note that points on the circle $\Gamma$ are their own reflections, and the reflection of a point inside the circle will always be a point outside the circle (and vice versa).

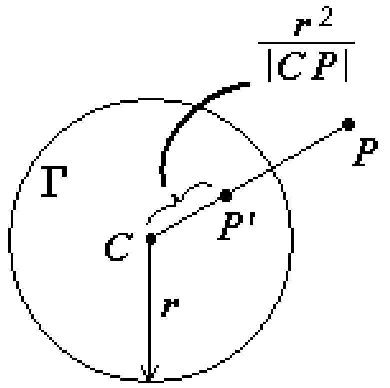



Figure 7.11: The reflection of $P$ through a circle $\Gamma$

Remarkably, this concept of reflection through a Euclidean circle corresponds exactly to hyperbolic reflection through an H-line. For if $\Gamma$ is a circle centered at a point $X$ on the $x$-axis and $\Lambda$ is the corresponding H-line in the halfplane model (so that $\Lambda$ is the portion of $\Gamma$ lying in the upper halfplane), then for any point $P$ in the upper halfplane, the point $r_\Lambda(P)$ is exactly the reflection of $P$ through the circle $\Gamma$. In other words:

| |
|---|
| If $\Lambda$ is an H-line given by a Euclidean circle with radius $r$ and center at the point $X$ on the $x$-axis, then $r_\Lambda(P)$ is the unique point on the (Euclidean) ray $\overrightarrow{XP}$ so that the product of the Euclidean distances $|XP|$ and $|Xr_\Lambda(P)|$ is $r^2$. |

We won't try to derive or justify this fact, depending as it does on the measurement of H-distance. Note, however, how convenient this makes the computation of $r_\Lambda$ – it isn't even necessary to find the H-line from $P$ that is perpendicular to $\Lambda$. (Though, of course, $r_\Lambda(P)$ *will* be on that H-line! In fact, it is the intersection of that H-line with the Euclidean segment $XP$, as depicted in Figure 7.12 – see Exercise 7.13.)

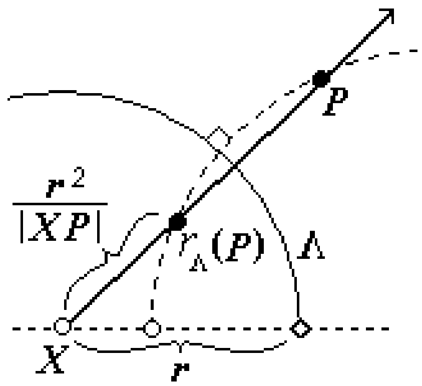



Figure 7.12: The reflection of $P$ through a circle $\Gamma$

What we *will* do below is use the above description of $r_\Lambda(P)$ to compute a formula for its coordinates in terms of the $x$- and $y$-coordinates of the point $P$. Then, we'll work an example using that formula, and verify (by computing angles and H-distances) that our answer actually meets the definition of a reflection.

So, let $(x, y)$ be a point in the upper halfplane, and let $\Lambda$ be an H-line. As mentioned above, if $\Lambda$ is a "vertical ray" H-line then reflection across $\Lambda$ is the same as in the Euclidean case:

If $\Lambda$ is the H-line $x = d$ then

$$r_\Lambda(x, y) = (-x + 2d, y).$$

Now suppose $\Lambda$ is the H-line with equation $y = \sqrt{r^2 - x^2}$ (the Euclidean semicircle with radius $r$ and center at the origin). Then we know that $r_\Lambda(x, y)$ will be on the ray from the origin through $(x, y)$ and so will have the form $(kx, ky)$ for some constant $k$. The distance from the origin to $r_\Lambda(x, y)$ times the distance from the origin to $(x, y)$ should be $r^2$, so:

$$\begin{aligned} \sqrt{(kx)^2 + (ky)^2}\sqrt{x^2 + y^2} &= r^2 \\ k(x^2 + y^2) &= r^2 \\ k &= \frac{r^2}{x^2 + y^2}. \end{aligned}$$

This means that $r_\Lambda(x, y) = (\frac{r^2 x}{x^2+y^2}, \frac{r^2 y}{x^2+y^2})$. To get a more general formula for the case when $\Lambda$ has equation $y = \sqrt{r^2 - (x-d)^2}$ (the Euclidean circle with radius $r$ and centered at $(d, 0)$) we imitate the three step procedure we used in the Euclidean case:

**Step 1:** Move the entire upper halfplane horizontally so that $\Lambda$ is moved to the H-line $\Lambda'$ with equation $y = \sqrt{r^2 - x^2}$.

**Step 2:** Reflect through $\Lambda'$ using the formula we derived above.

**Step 3:** Move the entire halfplane horizontally again so that the H-line of reflection is moved back to $\Lambda$.

Let's calculate the net effect of these steps on a generic point. Step 1 moves $(x, y)$ to $(x - d, y)$. For step 2 we use the formula we derived above to reflect $(x - d, y)$ through $y = \sqrt{r^2 - x^2}$, giving the result $\left(\frac{r^2(x-d)}{(x-d)^2+y^2}, \frac{r^2 y}{(x-d)^2+y^2}\right)$. Finally, step 3 merely shifts the $x$-coordinate by $d$, so the net result is that $(x, y)$ has been moved to $\left(\frac{r^2(x-d)}{(x-d)^2+y^2} + d, \frac{r^2 y}{(x-d)^2+y^2}\right)$.

If $\Lambda$ is the H-line $y = \sqrt{r^2 - (x-d)^2}$ then

$$r_\Lambda(x,y) = \left(\tfrac{r^2(x-d)}{(x-d)^2+y^2} + d, \tfrac{r^2 y}{(x-d)^2+y^2}\right).$$

EXAMPLE **7.16.** *Let $\Lambda$ be the H-line $y = \sqrt{9-(x+2)^2}$ and let $P$ be the point $(3,1)$. Find the coordinates of the point $P' = r_\Lambda(P)$, then verify that $\Lambda$ is the perpendicular H-bisector of the H-segment $PP'$.*

*Solution:* See Figure 7.13 for reference in the calculations that follow. We can find the coordinates of $P'$ easily using our formula with $r = 3$ and $d = -2$:

$$r_\Lambda(3,1) = \left(\frac{3^2(3+2)}{(3+2)^2+1^2} - 2, \frac{3^2(1)}{(3+2)^2+1^2}\right) = \left(\frac{-7}{26}, \frac{9}{26}\right).$$

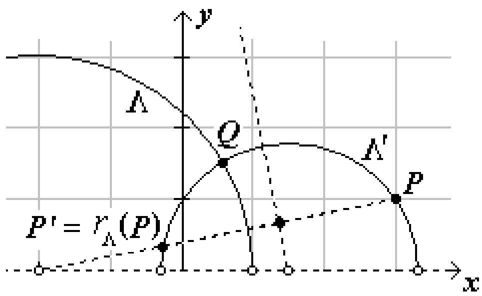



Figure 7.13:

To verify this answer we'll need to find the equation of the H-line through $P$ and $P'$. This will take the form of a semicircle through those two points with center where the $x$-axis meets the (Euclidean) perpendicular bisector of the Euclidean segment $PP'$. That perpendicular bisector will pass through the midpoint $(71/52, 35/52)$ and will have slope $-5$ (since the slope between $P$ and $P'$ is $1/5$), so its equation is $y - 35/52 = -5(x - 71/52)$. Solving for the $x$-intercept of this line we get $x = 3/2$. The (Euclidean) distance from $(3/2, 0)$ to $(3,1)$ is $\sqrt{13}/2$, so that is the radius of the circle defining the H-line $\Lambda'$ through $P$ and $P'$. Its equation is $y = \sqrt{13/4 - (x-3/2)^2}$.

Now, we need to check that $\Lambda'$ and $\Lambda$ really are perpendicular and that $P$ and $P'$ are H-equidistant from their point of intersection. We can use the equations of $\Lambda$ and $\Lambda'$ to find the coordinates of their point of intersection:

$$\begin{aligned} y = \sqrt{13/4 - (x-3/2)^2} &= \sqrt{9-(x+2)^2} \\ 13/4 - (x^2 - 3x + 9/4) &= 9 - (x^2 + 4x + 4) \\ 7x &= 4 \\ x &= 4/7, \quad y = \sqrt{9-(4/7+2)^2} = \frac{\sqrt{117}}{7} \end{aligned}$$

So, as in the figure, let $Q$ denote this point $\left(\frac{4}{7}, \frac{\sqrt{117}}{7}\right)$. Using the formula for H-distance (see p.216) we have

$$d_H(P, Q) = d_H\left((3,1), \left(\frac{4}{7}, \frac{\sqrt{117}}{7}\right)\right) = \text{Arccosh}\left(\frac{65}{2\sqrt{117}}\right)$$

and

$$d_H(P', Q) = d_H\left(\left(\frac{-7}{26}, \frac{9}{26}\right), \left(\frac{4}{7}, \frac{\sqrt{117}}{7}\right)\right) = \text{Arccosh}\left(\frac{65}{2\sqrt{117}}\right)$$

so $Q$ is the H-midpoint of the H-segment between $P$ and $P'$.

All that remains is to check that this H-segment is perpendicular to $\Lambda$. We do this by using derivatives to find the tangent slopes of $\Lambda$ and $\Lambda'$ at the point $Q$. For $\Lambda$ we have

$$\begin{aligned} y &= \sqrt{9-(x+2)^2} \\ \frac{dy}{dx} &= \frac{-x}{\sqrt{9-(x+2)^2}} \\ \left.\frac{dy}{dx}\right|_{x=4/7} &= \frac{-4/7}{\sqrt{9-(18/7)^2}} = \frac{-18}{\sqrt{117}} \end{aligned}$$

while for $\Lambda'$ we have

$$\begin{aligned} y &= \sqrt{\frac{13}{4} - \left(x - \left(\frac{3}{2}\right)^2\right)} \\ \frac{dy}{dx} &= \frac{-(x-3/2)}{\sqrt{\frac{13}{4} - \left(x - \left(\frac{3}{2}\right)^2\right)}} \\ \left.\frac{dy}{dx}\right|_{x=4/7} &= \frac{13/14}{\sqrt{\frac{13}{4} - \left(\frac{13}{14}\right)^2}} = \frac{13}{2\sqrt{117}} \end{aligned}$$

Since the product of these slopes is $\frac{-18}{\sqrt{117}} \cdot \frac{13}{2\sqrt{117}} = -1$, we have verified that the lines $\Lambda$ and $\Lambda'$ are perpendicular. $\square$

## Exercises

**7.9**. Prove the uniqueness of $r_\Lambda(P)$ claimed in the definition. That is, prove that given $\Lambda$ and $P$ there cannot be more than one point $Q$ with $\Lambda$ the perpendicular bisector of $PQ$.

**7.10**. Complete the proof of Theorem 7.12 by showing that $|r_\Lambda(A)r_\Lambda(B)| = |AB|$ in the remaining cases:

**(a)** Case 2: $A$ is on $\Lambda$ and $B$ is not.

**(b)** Case 4: $A$ and $B$ are on the same side of $\Lambda$.

**7.11**. Given a (Euclidean) line $\Lambda$ and a point $P$ not on $\Lambda$, give a straightedge and compass construction for the point $r_\Lambda(P)$.

**7.12**. Let $\Gamma$ be a Euclidean circle with center $C$ and radius $r$.

**(a)** Let $P$ be a point outside of $\Gamma$ (so $|CP| > r$). Let $Q$ be a point on $\Gamma$ so that $\overleftrightarrow{PQ}$ is tangent to $\Gamma$ (see Exercise 2.68). Let $P'$ be the point on $CP$ with $QP' \perp CP$. Prove that $P'$ is the reflection of $P$ through $\Gamma$.

**(b)** Now suppose that $P$ is a point *inside* $\Gamma$ (so $|CP| < r$). Keeping part **(a)** in mind, give a straightedge and compass construction for the reflection of $P$ through $\Gamma$.

**7.13**. Do the following with a dynamic geometry software package.

**(a)** Starting with a circle $\Gamma$ and a point $P$ not on $\Gamma$, construct the reflection $P'$ of $P$ through $\Gamma$. (Use the construction suggested by Exercise 7.12.)

**(b)** Verify using the software that all circles through both $P$ and $P'$ meet $\Gamma$ at right angles. (You can examine all such circles by defining the center to be a point on the perpendicular bisector of $PP'$, then simply moving this center along that perpendicular bisector.)

**7.14**. In each part below the equation of a line $\Lambda$ and a point $P$ is given. Find the coordinates of the point $r_\Lambda(P)$.

**(a)** $\Lambda$ is $2y = x$ and $P$ is $(7, 2)$

**(b)** $\Lambda$ is $y = 6 - x$ and $P$ is $(-3, -1)$

**(c)** $\Lambda$ is $3y + 4x = 3$ and $P$ is $(5, -3)$

**(d)** $\Lambda$ is $x + y = -2$ and $P$ is $(4, 4)$

**(e)** $\Lambda$ is $6x - y = 4$ and $P$ is $(0, 5)$

**7.15.** In each case find the image of the triangle with vertices $\left\{ \left( \begin{smallmatrix} 1 \\ 2 \end{smallmatrix} \right), \left( \begin{smallmatrix} -4 \\ 3 \end{smallmatrix} \right), \left( \begin{smallmatrix} -1 \\ 0 \end{smallmatrix} \right) \right\}$ under reflection through the given line.

**(a)** $y = 5$

**(b)** $x = -3$

**(c)** $x + 3y = 7$

**(d)** $3x - y = 1$

**7.16.** Suppose $\Lambda$ is a (Euclidean) line such that $r_\Lambda\left(\left( \begin{smallmatrix} 2 \\ -4 \end{smallmatrix} \right)\right) = \left(\left( \begin{smallmatrix} 0 \\ 1 \end{smallmatrix} \right)\right)$. Find $r_\Lambda\left(\left( \begin{smallmatrix} -1 \\ 3 \end{smallmatrix} \right)\right)$.

**7.17.** In each case there is an H-line $\Lambda$ and a point $P$ given. As we did in Example 7.16, find the coordinates of the point $P' = r_\Lambda(P)$ then verify that $\Lambda$ is the perpendicular H-bisector of the H-segment $PP'$.

**(a)** $\Lambda$ is $x = -2$ and $P$ is $(1, 5)$

**(b)** $\Lambda$ is $y = \sqrt{9 - x^2}$ and $P$ is $(2, 1)$

**(c)** $\Lambda$ is $y = \sqrt{1 - (x - 5)^2}$ and $P$ is $(3, 2)$

# C. Orientation

In Chapter 6 we introduced the notion of an oriented line (see p.233). We'll now extend that to an orientation for the entire plane. This corresponds to our experience with "clockwise" and "counterclockwise" orientations in the Cartesian coordinate plane. This isn't actually a big step, as we really need only add one assumption to what we have already established in our axioms.

Recall that Axiom AC (p.133) assumes a natural way to put "protractor functions" on circles. This actually implies that there is a natural orientation to each circle, and all we need to do to get an orientation for the entire plane is somehow assure that the orientations on circles are all in some way compatible. We do this by simply extending Axiom AC in the following way.

As in Figure 7.14, let $\Gamma_1$ and $\Gamma_2$ be circles with centers $C_1$ and $C_2$, respectively. Give the line $\Lambda = \overleftrightarrow{C_1C_2}$ the orientation $(C_1, C_2)$. (If $C_1 = C_2$ then choose any line $\Lambda$ through $C_1$ and give it either orientation.) Suppose that $\Gamma_1$ meets $\Lambda$ at points $A_1$ and $B_1$ and that $\Gamma_2$ meets $\Lambda$ at points $A_2$ and $B_2$, and suppose the orientations $(A_1, B_1)$ and $(A_2, B_2)$ for $\Lambda$ agree with $(C_1, C_2)$. Let $p(t) = P_t$ and $q(t) = Q_t$ be protractor functions for $\Gamma_1$ and $\Gamma_2$, respectively, with $P_0 = B_1$ and $Q_0 = B_2$.

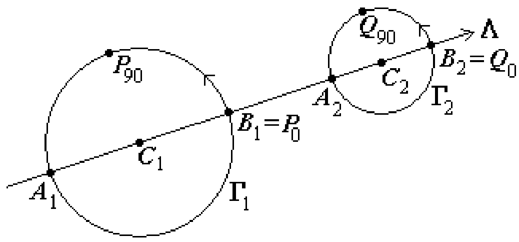



Figure 7.14:

> **Extension of Axiom AC.**
>
> *With objects as named in the above paragraph, it will always be the case that $P_{90}$ and $Q_{90}$ are on the same side of line $\Lambda$.*

What is actually going on here is this: there are, for each circle, two possible "orientations" for protractor functions, and Axiom AC stipulates that one has been selected for us. For while (according to the axiom) there is for each point

$A$ on $\Gamma$ a designated protractor function $p(t) = P_t$ that places $P_0$ at $A$, there is also a related "reverse" function $\hat{p}(t) = p(-t)$ that shares the properties of a protractor function – see Exercise 7.18. When the protractor functions for a circle are designated (as in Axiom AC), the circle is effectively assigned either a "clockwise" or "counterclockwise" orientation. This extension of Axiom AC now further stipulates that the choice of orientation for each circle is compatible with that of all other circles.

But what does all this have to do with transformations? To see the answer, suppose that $f$ is a transformation and that $\Gamma$ is a circle with protractor function $p(t) = P_t$ satisfying $P_0 = A$. Part (a) of Exercise 7.19 shows that $f(\Gamma)$ is a circle, and since $f(A)$ is a point of $f(\Gamma)$, there would (by Axiom CA) be a designated protractor function $q(t) = Q_t$ for $f(\Gamma)$ satisfying $Q_0 = f(A)$. Now $q(t)$ is, of course, a function from $\mathbb{R}$ to $f(\Gamma)$. But we can define another function from $\mathbb{R}$ to $f(\Gamma)$ by simply composing $f$ and $p$ – that is, by using the rule $f(p(t)) = f(P_t)$ – and this function also takes zero to $f(A)$ since $f(P_0) = f(A)$. Part (b) of Exercise 7.19 shows that either $f(P_t) = q(t)$ or $f(P_t) = q(-t)$. In the first case we say that $f$ *preserves orientation for* $\Gamma$ and in the second case we say that $f$ *reverses orientation for* $\Gamma$.

**FACT 7.17.** *If $f$ is a transformation then either $f$ preserves the orientation of every circle or $f$ reverses the orientation of every circle.*

*Proof:* Assume (to reach a contradiction) that $\Gamma_1$ and $\Gamma_2$ are circles and $f$ is a transformation that preserves the orientation of $\Gamma_1$ and reverses the orientation of $\Gamma_2$.

- We will assume that $p(t) = P_t$ and $q(t) = Q_t$ are protractor functions for $\Gamma_1$ and $\Gamma_2$, as in Figure 7.14.

- Then by our extension of Axiom AC, $P_{90}$ and $Q_{90}$ are in the same halfplane determined by $\Lambda = \overleftrightarrow{P_0Q_0}$.

- Let $r(t) = R_t$ and $s(t) = S_t$ be the protractor functions for $f(\Gamma_1)$ and $f(\Gamma_2)$ respectively, so that $R_0 = f(P_0)$ and $S_0 = f(Q_0)$.

- By assumption (since $f$ preserves orientation for $\Gamma_1$), we have $f(P_t) = R_t$ for all $t$.

- But also by assumption (since $f$ reverses orientation for $\Gamma_2$), $f(Q_t) = S_{-t}$ for all $t$.

- Thus, $f(P_{90}) = R_{90}$ and $f(Q_{90}) = S_{-90}$.
- Since $R_{90}$ and $S_{90}$ are (by our extension of Axiom AC) in the same halfplane determined by $\overleftrightarrow{R_0S_0}$ while $S_{90}$ and $S_{-90}$ are (by Axiom AC) in different halfplanes determined by that line, we conclude that $f(P_{90})$ and $f(Q_{90})$ are in different halfplanes determined by $\overleftrightarrow{R_0S_0} = f(\overleftrightarrow{P_0Q_0})$. This is a contradiction by Exercise 7.20. □

DEFINITION. The transformation $f$ is *orientation preserving* if it preserves orientation of circles and *orientation reversing* if it reverses orientation of circles.

Note that by Fact 7.17 we need only examine the effect of $f$ on *one* circle to determine if it is orientation preserving or orientation reversing.

We leave the proof of the following basic fact as an exercise.

THEOREM **7.18.** *The composition of two orientation preserving transformations or two orientation reversing transformations is orientation preserving. The composition of an orientation preserving transformation and an orientation reversing transformation (in either order) is orientation reversing.*

Our mental image of reflections as a "flip" of the plane suggests that they should reverse orientation. We'll see that this conjecture is correct by proving the following rule for the effect of a reflection on protractor functions.

FACT **7.19.** *Let $\Gamma$ be a circle with center $C$ and protractor function $q(t) = Q_t$. For each $\alpha$ let $\Lambda_\alpha$ be the line $\overleftrightarrow{CQ_\alpha}$. Then for each $t$ we have $r_{\Lambda_\alpha}(Q_{\alpha+t}) = Q_{\alpha-t}$.*

*Proof:* First note that since $r_{\Lambda_\alpha}$ fixes $C$ we know from Theorem 7.8 that $r_{\Lambda_\alpha}(Q_{\alpha+t})$ is a point of $\Gamma$.

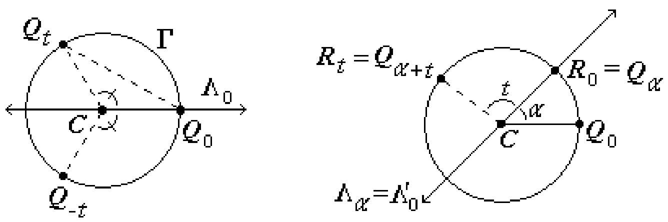



Figure 7.15:

- If $\alpha = 0$ and $0 < t < 180$ then we can easily see (refer to the left half of Figure 7.15) that $r_{\Lambda_0}(Q_t) = Q_{-t}$:
  - Since $r_{\Lambda_0}$ is a transformation fixing $C$ and $Q_0$ we have $\triangle Q_0CQ_t \cong \triangle Q_0C\,r_{\Lambda_0}(Q_t)$.
  - So $m\angle Q_0C\,r_{\Lambda_0}(Q_t) = m\angle Q_0CQ_t = t$, meaning that $r_{\Lambda_0}(Q_t)$ can be either $Q_t$ itself or else $Q_{-t}$.
  - But $r_{\Lambda_0}(Q_t) = Q_t$ would imply $r_{\Lambda_0} = id$ by Theorem 7.10 (since $r_{\Lambda_0}$ would fix $C$, $Q_0$, and $Q_t$), an impossibility!
- For negative values of $t$ ($-180 < t < 0$) the above case tells us $r_{\Lambda_0}(Q_{-t}) = Q_t$. So, $r_{\Lambda_0}(Q_t) = r_{\Lambda_0}(r_{\Lambda_0}(Q_{-t})) = r_{\Lambda_0} \circ r_{\Lambda_0}(Q_{-t}) = id(Q_{-t}) = Q_{-t}$.
- Clearly $r_{\Lambda_0}$ fixes $Q_{180} = Q_{-180}$, so we may now say $r_{\Lambda_0}(Q_t) = Q_{-t}$ for all $t$ in the range $-180 \leq t \leq 180$. This easily implies $r_{\Lambda_0}(Q_t) = Q_{-t}$ for all $t$.
- Now for the case $\alpha \neq 0$ we simply take a new protractor function $r(t) = R_t = Q_{t+\alpha}$ (using the last stipulation in Axiom CA) and let $\Lambda'_\theta$ denote $\overleftrightarrow{CR_\theta} = \overleftrightarrow{CQ_{\theta+\alpha}} = \Lambda_{\theta+\alpha}$ (as in the right half of Figure 7.15). Applying our previous case (with $\theta = 0$) gives

$$r_{\Lambda_\alpha}(Q_{\alpha+t}) = r_{\Lambda'_0}(R_t) = R_{-t} = Q_{\alpha-t}\,.$$

$\square$

**Corollary 7.20.** *Reflections are orientation reversing transformations.*

*Proof:* Consider a reflection $r_\Lambda$. By Theorem 7.17 it is enough to show that $r_\Lambda$ reverses the orientation of one circle.

- Let $\Gamma$ be a circle with center on $\Lambda$.
- We may assume that the protractor function $q(t) = Q_t$ for $\Gamma$ is such that $Q_0$ is on $\Lambda$.
- By Fact 7.19 we have $r_\Lambda(Q_t) = Q_{-t}$.
- So, $r_\Lambda(\Gamma) = \Gamma$ and $r_\Lambda(q(t))$ is the reverse of the protractor function $q(t)$. Thus $r_\Lambda$ reverses orientation. $\square$

## Exercises

**7.18**. Let $\Gamma$ be a circle and $p(t) = P_t$ the protractor function for $\Gamma$ with $P_0 = A$.

**(a)** Show that $\hat{p}(t) = p(-t) = \hat{P}_t$ satisfies the following properties as in Axiom CA:

- $\hat{P}_0 = A$.
- $\hat{p}$ is one-to-one and onto from $(-180, 180]$ to $\Gamma$.
- $\hat{p}$ is periodic, with $\hat{P}_{\alpha+360} = \hat{P}_\alpha$ for all $\alpha$.
- $\hat{P}_\alpha * C * \hat{P}_{\alpha+180}$ for all $\alpha$.
- The portions of $\Gamma$ contained in the two halfplanes determined by $\overleftrightarrow{CA}$ are $\{\hat{P}_\alpha : -180 < \alpha < 0\}$ and $\{\hat{P}_\alpha : 0 < \alpha < 180\}$.
- If $0 < \beta < 180$ then the portion of $\Gamma$ in the interior of $\angle \hat{P}_\beta CA$ is $\{P_\alpha : 0 < \alpha < \beta\}$.

**(b)** According to Axiom AC, there is one designated protractor function for $\Gamma$ that takes zero to point $A$. If we wanted to use $\hat{p}(t)$ as this designated protractor function instead of $p(t)$, what else would we have to change?

**7.19**. Let $f$ be a transformation and $\Gamma$ be a circle with center $C$ and protractor function $p(t) = P_t$ with $P_0 = A$.

**(a)** Prove that $f(\Gamma)$ is a circle with center $f(C)$.

**(b)** Let $q(t) = Q_t$ be the protractor function for $f(\Gamma)$ with $Q_0 = f(A)$. Define $r(t) = f(P_t)$. Show that either $r(t) = q(t)$ or $r(t) = q(-t)$ for all $t$ (see the previous exercise).

**7.20**. Complete the proof of Theorem 7.17 by proving that if $\Lambda$ is a line, $f$ is a transformation, and $A$ and $B$ are points on the same side of $\Lambda$, then $f(A)$ and $f(B)$ are on the same side of $f(\Lambda)$.

**7.21**. Prove Theorem 7.18.

# D. Rotations

From Theorems 7.9 and 7.10 we know that a transformation fixes either nothing, one point, one line, or all points. Theorem 7.13 tells us that all transformations fixing a line (but not the entire plane) are reflections. In this section we'll study the transformations that fix a single point.

Such transformations are called *rotations*, and our first goal will be to prove that this name is deserved. We'll show that transformations fixing a single point of necessity act in a manner compatible with our mental image of "rotating the plane about a point".

**Theorem 7.21.** *Let $f$ be a transformation fixing exactly one point $A$, and let $\Gamma$ be a circle centered at $A$ (with protractor function $q(t) = Q_t$). Suppose that $f(Q_0) = Q_\theta$ with $-180 < \theta \leq 180$. Then for each $t$ we have $f(Q_t) = Q_{t+\theta}$.*

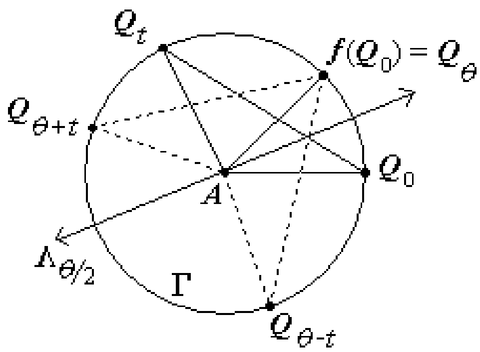



Figure 7.16:

*Proof:* Assume that $Q_t$ is a point of $\Gamma$ other than $Q_0$ (with $-180 < t \leq 180$). We will show that $f(Q_t) = Q_{t+\theta}$. Refer to Figure 7.16 in the following steps.

- Clearly $f(Q_t) = Q_{t+\theta}$ holds for $t = 0$. For $t = 180$ note that since $Q_{180}$ is collinear with $A$ and $Q_0$ then $f(Q_{180})$ must be collinear with $f(A) = A$ and $f(Q_0) = Q_\theta$. So either $f(Q_{180}) = Q_\theta$ (which is impossible since $f$ is one-to-one) or $f(Q_{180}) = Q_{180+\theta}$. Having verified the result for $t = 0$ and $t = 180$, we will assume henceforth that $Q_t$ is not collinear with $A$ and $Q_0$.

- Since $f$ is a transformation we have $\triangle AQ_0Q_t \cong \triangle f(A)f(Q_0)f(Q_t) = \triangle AQ_\theta f(Q_t)$.

- So $m\angle Q_\theta Af(Q_t) = m\angle Q_0AQ_t = |t|$, so $f(Q_t)$ must be either $Q_{\theta+t}$ or $Q_{\theta-t}$.

- But we claim $f(Q_t)$ cannot be $Q_{\theta-t}$:
  - By Fact 7.19 we have $r_{\Lambda_{\theta/2}}(Q_\theta) = r_{\Lambda_{\theta/2}}(Q_{(\theta/2+\theta/2)}) = Q_{(\theta/2-\theta/2)} = Q_0$ and $r_{\Lambda_{\theta/2}}(Q_t) = r_{\Lambda_{\theta/2}}(Q_{\theta/2+(t-\theta/2)}) = Q_{\theta/2-(t-\theta/2)} = Q_{\theta-t}$.
  - So if $f(Q_t) = Q_{\theta-t}$ we would have $f(A) = A = r_{\Lambda_{\theta/2}}(A)$, $f(Q_0) = Q_\theta = r_{\Lambda_{\theta/2}}(Q_0)$, and $f(Q_t) = Q_{\theta-t} = r_{\Lambda_{\theta/2}}(Q_t)$.
  - By Theorem 7.11 we would then have $f = r_{\Lambda_{\theta/2}}$, which is impossible since $f$ has only one fixed point. □

So, the transformation $f$ in the above theorem has the effect of "rotating" the points on each circle centered at $A$ by a fixed amount. This justifies the following definition.

**Definition.** A transformation $f$ fixing a single point $A$ is called a *rotation* about $A$ by $\theta$ where $\theta$ is as in Theorem 7.21 above.

Note that $\theta$ in the above definition can be any number in the range $-180 < \theta \leq 180$. We can think of rotations with positive $\theta$ as being "counterclockwise" and rotations with negative $\theta$ as being "clockwise". In either case, the circle $\Gamma$ is preserved by $f$ and $f(Q_t) = Q_{t+\theta}$ is (by Axiom CA) just another protractor function for $\Gamma$. Thus, $f$ is an orientation preserving transformation.

We mentioned at the outset of Section B that reflections are the fundamental building blocks for all transformations. We'll see now how each rotation can be expressed in terms of reflections. The process is really quite simple. Since reflection across each line through $A$ fixes $A$, compositions of those reflections will also fix $A$. Our next theorem proves that every rotation about $A$ can be written as the composition of two such reflections.

**Theorem 7.22.** *Let $f$ be a rotation about $A$. Then given any line $\Lambda$ through $A$ there is a unique line $\Lambda'$ through $A$ so that $f = r_{\Lambda'} \circ r_\Lambda$.*

*Proof:* First we show the existence of a line $\Lambda'$ so that $f = r_{\Lambda'} \circ r_\Lambda$. Choose a point $P$ on $\Lambda$ (other than $A$) and let $\Lambda'$ be the line bisecting the angle $\angle PAf(P)$ (as in Figure 7.17).

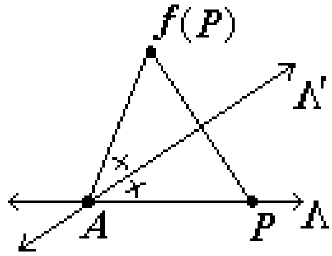



Figure 7.17:

- It's easy to show using our neutral geometry facts from Chapter 4 that $\Lambda'$ is the perpendicular bisector of $Pf(P)$ (see Exercise 7.22), so $f(P) = r_{\Lambda'}(P)$.

- Then we have $r_{\Lambda'} \circ f(P) = P$ and $r_{\Lambda'} \circ f(A) = A$, so $r_{\Lambda'} \circ f$ is a transformation fixing the entire line $\overleftrightarrow{PA} = \Lambda$ by Theorem 7.9.
- But $r_{\Lambda'} \circ f$ is clearly not the identity transformation (since $r_{\Lambda'}^{-1} = r_{\Lambda'} \neq f$), so $r_{\Lambda'} \circ f = r_\Lambda$ by Theorem 7.13.
- So:

$$f = id \circ f = (r_{\Lambda'} \circ r_{\Lambda'}) \circ f = r_{\Lambda'} \circ (r_{\Lambda'} \circ f) = r_{\Lambda'} \circ r_\Lambda \,.$$

Now we must show that this line $\Lambda'$ is unique. So suppose $\Lambda''$ is any line such that $f = r_{\Lambda''} \circ r_\Lambda$. Then:

$$\begin{aligned} r_{\Lambda''} &= r_{\Lambda''} \circ id = r_{\Lambda''} \circ (r_\Lambda \circ r_\Lambda) = (r_{\Lambda''} \circ r_\Lambda) \circ r_\Lambda = f \circ r_\Lambda \\ &= (r_{\Lambda'} \circ r_\Lambda) \circ r_\Lambda = r_{\Lambda'} \circ (r_\Lambda \circ r_\Lambda) = r_{\Lambda'} \circ id = r_{\Lambda'} \,. \end{aligned}$$

So clearly $\Lambda''$ must equal $\Lambda'$. □

COROLLARY **7.23.** *Every transformation fixing exactly one point is a rotation and can be written as the composition of two reflections. In fact, if $A$ is the fixed point of $f$ and $P$ is any point other than $A$ then $f = r_{\Lambda'} \circ r_\Lambda$ where $\Lambda = \overleftrightarrow{PA}$ and $\Lambda'$ is the line bisecting $\angle PAf(P)$.* □

Now suppose $f$ is a transformation fixing exactly one point $A$. By Theorem 7.21 $f$ is a rotation about $A$ by some $\theta$ in the range $-180 < \theta \leq 180$. What is $\theta$? If objects are named as in the proof of Theorem 7.22 (see Figure 7.17) and the acute (or possibly right) angle between $\Lambda$ and $\Lambda'$ has measure $\alpha$, then it is clear that $m\angle PAf(P) = 2\alpha$. So, if $\Gamma$ is centered at $A$ and passes through $P$ and if we let $q(t) = Q_t$ be the protractor function for $\Gamma$ with $Q_0 = P$, then $f(P)$ will be either $Q_{2\alpha}$ or $Q_{-2\alpha}$. Thus, $\theta$ is either $2\alpha$ or $-2\alpha$.

We'll now shift gears to see how rotations play out in the models for Euclidean and hyperbolic geometry. We'll start with the Euclidean case.

**Rotations in the Euclidean model**

In Section B we were able to derive formulae for reflections in vector notation. We'll now do the same for rotations, starting with rotations about the origin. So, let $f$ be the rotation by $\theta$ about the origin, where $-180 < \theta \leq 180$.[5] Let $O$

[5]As is customary, we will adopt the orientation for the plane in which positively oriented circles have counterclockwise sense. So, a negative value of $\theta$ will correspond to a clockwise rotation while a positive value of $\theta$ will correspond to a counterclockwise rotation.

denote the origin $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $U$ be the point $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. From trigonometry we know that a protractor function of the unit circle $x^2 + y^2 = 1$ is $Q_t = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$. From Theorem 7.21 we have $f(U) = f(Q_0) = Q_\theta = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$.

By Corollary 7.23 we can express $f$ as the composition of two reflections, $f = r_{\Lambda'} \circ r_\Lambda$, where $\Lambda$ is the $x$-axis and $\Lambda'$ is the bisector of angle $\angle UO\,f(U)$ – see Figure 7.18. Now the slope of line $\Lambda'$ is evidently $\sin(\frac{\theta}{2})/\cos(\frac{\theta}{2}) = \tan(\frac{\theta}{2})$, so applying our formula for reflection across lines through the origin (p.265) we have[6]

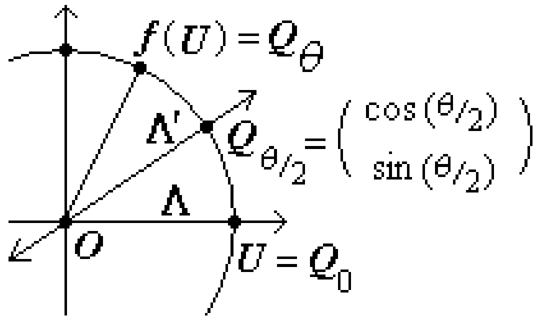



Figure 7.18:

$$\begin{aligned}
f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) &= r_{\Lambda'} \circ r_\Lambda\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = r_{\Lambda'}\left(\begin{pmatrix} x \\ -y \end{pmatrix}\right) \\
&= \frac{1}{\tan^2(\frac{\theta}{2}) + 1}\begin{pmatrix} 1 - \tan^2(\frac{\theta}{2}) & 2\tan(\frac{\theta}{2}) \\ 2\tan(\frac{\theta}{2}) & \tan^2(\frac{\theta}{2}) - 1 \end{pmatrix}\begin{pmatrix} x \\ -y \end{pmatrix} \\
&= \cos^2(\theta/2)\begin{pmatrix} 1 - \tan^2(\frac{\theta}{2}) & 2\tan(\frac{\theta}{2}) \\ 2\tan(\frac{\theta}{2}) & \tan^2(\frac{\theta}{2}) - 1 \end{pmatrix}\begin{pmatrix} x \\ -y \end{pmatrix} \\
&= \begin{pmatrix} \cos^2(\frac{\theta}{2}) - \sin^2(\frac{\theta}{2}) & 2\sin(\frac{\theta}{2})\cos(\frac{\theta}{2}) \\ 2\sin(\frac{\theta}{2})\cos(\frac{\theta}{2}) & \sin^2(\frac{\theta}{2}) - \cos^2(\frac{\theta}{2}) \end{pmatrix} \\
&= \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}\begin{pmatrix} x \\ -y \end{pmatrix} \\
&= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}
\end{aligned}$$

If $f$ is rotation by $\theta$ about the origin then

$$f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

For rotation about a point $A = \begin{pmatrix} a \\ b \end{pmatrix}$ other than the origin we use a three step procedure similar to what we did for reflections. The process is illustrated in Figure 7.19.

[6]The calculations that follow will use the “double angle” identities from trigonometry: $\sin(2\alpha) = 2\sin\alpha\cos\alpha$, and $\cos(2\alpha) = \cos^2\alpha - \sin^2\alpha$.

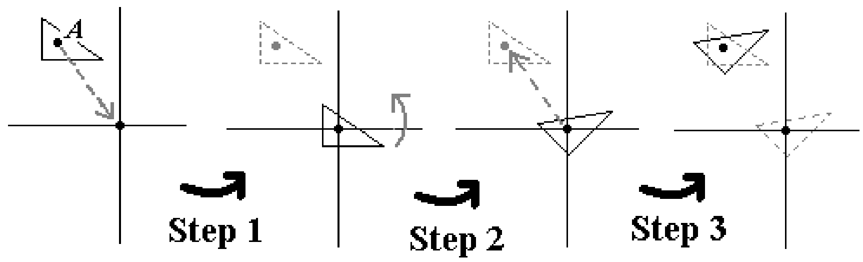



Figure 7.19: Accomplishing a rotation about $A$ in three steps: **(1)** translate $A$ to the origin, **(2)** rotate about the origin, and **(3)** translate the center of rotation back to $A$.

Calculating the net effect on a point $\begin{pmatrix} x \\ y \end{pmatrix}$ is easy:

$$\begin{pmatrix} x \\ y \end{pmatrix} \xrightarrow{\textbf{Step 1}} \begin{pmatrix} x-a \\ y-b \end{pmatrix} \xrightarrow{\textbf{Step 2}} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x-a \\ y-b \end{pmatrix}$$

$$\xrightarrow{\textbf{Step 3}} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x-a \\ y-b \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix}$$

> If $f$ is rotation by $\theta$ about $\begin{pmatrix} a \\ b \end{pmatrix}$ then
>
> $$f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x-a \\ y-b \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix}$$

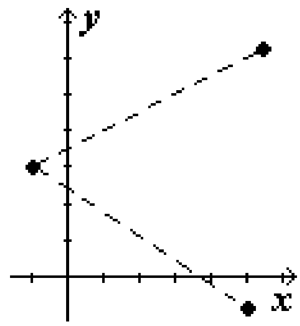



Figure 7.20:

EXAMPLE **7.24.** *Find* $f\left(\begin{pmatrix} 5 \\ -1 \end{pmatrix}\right)$ *where* $f$ *is rotation by* 60 *degrees about* $\begin{pmatrix} -1 \\ 3 \end{pmatrix}$ *(see Figure 7.20).*

*Solution:* By the formula above we have

$$\begin{aligned} f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) &= \begin{pmatrix} \cos(60) & -\sin(60) \\ \sin(60) & \cos(60) \end{pmatrix} \begin{pmatrix} x+1 \\ y-3 \end{pmatrix} + \begin{pmatrix} -1 \\ 3 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x+1 \\ y-3 \end{pmatrix} + \begin{pmatrix} -1 \\ 3 \end{pmatrix}, \end{aligned}$$

so

$$f\left(\begin{pmatrix} 5 \\ -1 \end{pmatrix}\right) = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 6 \\ -4 \end{pmatrix} + \begin{pmatrix} -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 2+2\sqrt{3} \\ 3\sqrt{3}+1 \end{pmatrix}. \quad \square$$

**EXAMPLE 7.25.** *Find the formula for the rotation $f$ carrying triangle $PQR$ in Figure 7.21 to triangle $P'Q'R'$.*

*Solution:* The center of rotation must be on each of the perpendicular bisectors of segments $PP'$, $QQ'$, and $RR'$. The perpendicular bisector of $PP'$ is clearly the vertical line $x = 1$, and the perpendicular bisector of $QQ'$ is easily computed as $y = 2x + 2$. These lines meet at the point $A = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$, so that must be the point about which the rotation occurs. But now the slope of $AP$ is $-1$ while the slope of $AP'$ is 1. So $AP \perp AP'$ and it is clear that the rotation is clockwise by 90 degrees. Thus, our formula is

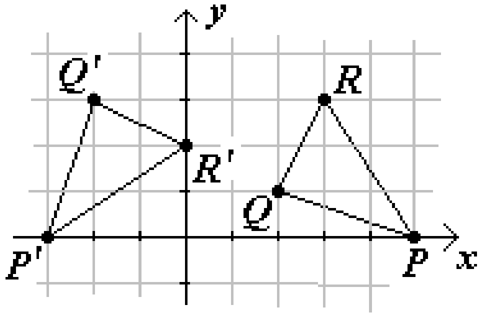



Figure 7.21:

$$\begin{aligned} f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) &= \begin{pmatrix} \cos(-90) & -\sin(-90) \\ \sin(-90) & \cos(-90) \end{pmatrix}\begin{pmatrix} x-1 \\ y-4 \end{pmatrix} + \begin{pmatrix} 1 \\ 4 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} x-1 \\ y-4 \end{pmatrix} + \begin{pmatrix} 1 \\ 4 \end{pmatrix} = \begin{pmatrix} y-3 \\ -x+5 \end{pmatrix}. \end{aligned}$$

□

**EXAMPLE 7.26.** *Find a formula for the rotation $f$ in the Euclidean model that carries $P = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ to $f(P) = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$ and carries $R = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$ to $f(R) = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$.*

*Solution:* Notice that the fixed point of this rotation must be equidistant from $P$ and $f(P)$ and also equidistant from $R$ and $f(R)$. Thus we can locate the fixed point by intersecting the perpendicular bisectors of $Pf(P)$ and $Rf(R)$. These are evidently the lines $y = 1$ and $y = x-1$, so the fixed point must be $A = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ (see Figure 7.22). It is also evident that the rotation is clockwise by the angle measure $2\alpha$ where $\alpha$ is as in Figure 7.23. Using simple triangle trigonometry we can calculate $\sin\alpha = \frac{1}{\sqrt{5}}$ and $\cos\alpha = \frac{2}{\sqrt{5}}$, so using the usual "double angle" identities we have

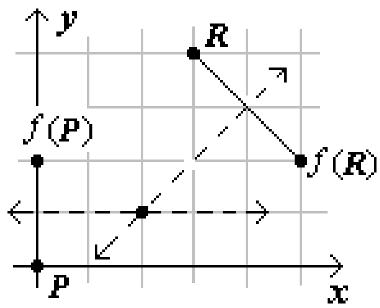



Figure 7.22:

$$\sin(-2\alpha) = -2\sin\alpha\cos\alpha = -4/5\,, \text{ and}$$

$$\cos(-2\alpha) = \cos^2\alpha \; - \sin^2\alpha = 3/5\,.$$

So using the formula on p.285 we have

$$f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} 3/5 & 4/5 \\ -4/5 & 3/5 \end{pmatrix}\begin{pmatrix} x-2 \\ y-1 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

It's easy to check using this formula that
$f\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$ and $f\left(\begin{pmatrix} 3 \\ 4 \end{pmatrix}\right) = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$.

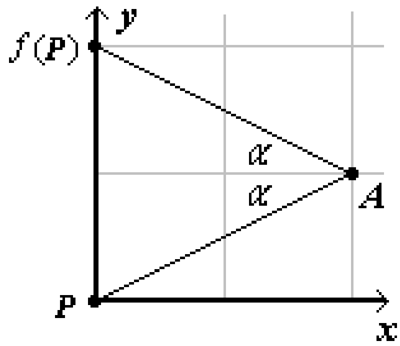



Figure 7.23: □

**Rotations in the hyperbolic model**

We will not here compute a general formula for rotations by $\theta$ about a point in the halfplane model for hyperbolic geometry. However, we can use the recipe given by Theorem 7.22 and Corollary 7.23 to compute specific rotations in that model by successive reflections across H-lines.

Example 7.27. *Compute the result of rotating the point $A = (2, 3)$ by 90 degrees about the point $P = (0, 3)$ in the halfplane model for hyperbolic geometry.*

*Solution:* Let $f$ be rotation by 90 degrees about the point $P = (0, 3)$ and let $\Lambda$ be the H-line $x = 0$. According to Theorem 7.22 there is an H-line $\Lambda'$ through $(0, 3)$ so that $f = r_{\Lambda'} \circ r_\Lambda$ (see Figure 7.24). By Corollary 7.23 we know that $\Lambda'$ will be the H-line through $(0, 3)$ making an angle of 45 degrees with $\Lambda$. Clearly then $\Lambda'$ will be given by a (Euclidean) semicircle passing through $(0, 3)$ and with its radius segment forming an angle of 45 degrees with the $y$-axis. (The radius segment of a circle is perpendicular to the circle's tangent, so if the tangent makes an angle of 45 degrees with the $y$-axis, then so will the radius segment!) It's easy to see that the center of this semicircle must be at $(-3, 0)$, so

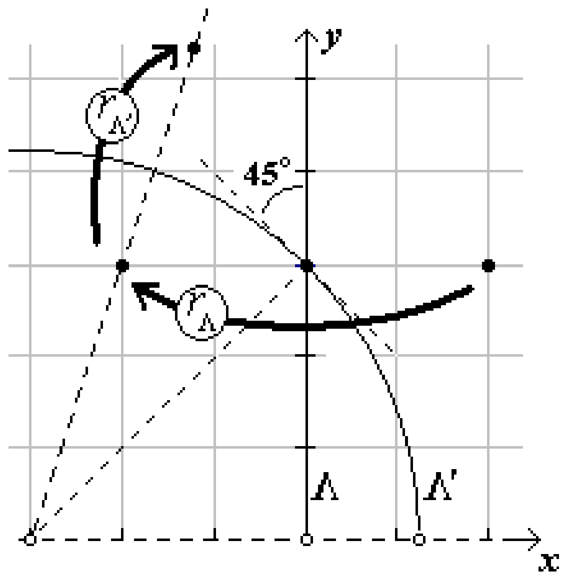



Figure 7.24:

the equation of $\Lambda'$ is $(x+3)^2 + y^2 = 18$. The formula for $r_\Lambda$ is clearly

$$r_\Lambda(x,y) = (-x,y)\,,$$

and the formula on p.272 gives us

$$r_{\Lambda'}(x,y) = \left(\frac{18(x+3)}{(x+3)^2+y^2} - 3, \frac{18y}{(x+3)^2+y^2}\right)\,.$$

So:

$$\begin{aligned} f(A) = f(2,3) &= r_{\Lambda'} \circ r_\Lambda(2,3) \\ &= r_{\Lambda'}(-2,3) \\ &= \left(\frac{18}{10} - 3, \frac{54}{10}\right) = \left(-\frac{6}{5}, \frac{27}{5}\right)\,. \end{aligned}$$

We'll leave it to you (see Exercise 7.28) to check that the H-distances $d_H(P,A)$ and $d_H(P,f(A))$ are equal and that the H-angle $\angle AP\,f(A)$ measures 90-degrees.
□

## Exercises

**7.22**. In the proof of Theorem 7.22 show that $\Lambda'$ is the perpendicular bisector of $Af(A)$ using only neutral geometry results.

**7.23**. Suppose that $r_{\Lambda'} \circ r_\Lambda$ is rotation by $\theta$ about the point $P$. What is $r_\Lambda \circ r_{\Lambda'}$? Justify your answer.

**7.24**. Find the result of rotating the given point of the Euclidean plane as indicated.

**(a)** $\left(\begin{smallmatrix} 2 \\ -3 \end{smallmatrix}\right)$ rotated by 45 degrees about the origin.

**(b)** $\left(\begin{smallmatrix} 4 \\ 1 \end{smallmatrix}\right)$ rotated by $-120$ degrees about $\left(\begin{smallmatrix} 2 \\ 0 \end{smallmatrix}\right)$.

**(c)** The origin rotated by 30 degrees about $\left(\begin{smallmatrix} 5 \\ 7 \end{smallmatrix}\right)$.

**(d)** $\left(\begin{smallmatrix} 6 \\ 1 \end{smallmatrix}\right)$ rotated by $-90$ degrees about $\left(\begin{smallmatrix} 3 \\ -4 \end{smallmatrix}\right)$.

**(e)** $\left(\begin{smallmatrix} -2 \\ -1 \end{smallmatrix}\right)$ rotated by 150 degrees about $\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$.

**(f)** $\begin{pmatrix} 0 \\ 4 \end{pmatrix}$ rotated by $-135$ degrees about $\begin{pmatrix} 3 \\ -4 \end{pmatrix}$.

**7.25**. Find the formula for the rotation $f$ in the Euclidean model that carries the triangle with vertices $\left\{\begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \end{pmatrix}\right\}$ to the triangle with vertices $\left\{\begin{pmatrix} 2 \\ -4 \end{pmatrix}, \begin{pmatrix} -1 \\ -6 \end{pmatrix}, \begin{pmatrix} -2 \\ -5 \end{pmatrix}\right\}$.

**7.26**. Find the formula for the rotation in the Euclidean model that carries $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ to $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$ and carries $\begin{pmatrix} 2 \\ 2 \end{pmatrix}$ to $\begin{pmatrix} 2+2\sqrt{2} \\ 0 \end{pmatrix}$.

**7.27**. Find the formula for the rotation in the Euclidean model that carries $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ to $\begin{pmatrix} 0 \\ 4 \end{pmatrix}$ and carries $\begin{pmatrix} 12 \\ 5 \end{pmatrix}$ to $\begin{pmatrix} 0 \\ 17 \end{pmatrix}$.

**7.28**. Verify the answer from Example 7.27 by showing that $d_H(P, A) = d_H(P, f(A))$ (see the formula for H-distance on p.216) and that the H-angle $\angle AP\,f(A)$ measures 90-degrees.

**7.29**. Using the formulae in Example 7.27, compute the result of rotating each of the following points by 90 degrees about $(0, 3)$ in the halfplane model for hyperbolic geometry.

**(a)** $(0, 4)$

**(b)** $(3, 1)$

**(c)** $(-5, 2)$

**7.30**. Let $f$ be a clockwise rotation by 90 degrees about the point $(4, 3)$ in the halfplane model for hyperbolic geometry. Find a formula for $f(x, y)$.

**7.31**. Let $f$ be a counterclockwise rotation by 60 degrees about the point $(4, 3)$ in the halfplane model for hyperbolic geometry. Find a formula for $f(x, y)$. (You can still imitate the solution to Example 7.27, but this time the angle that $\Lambda'$ must make with the vertical H-line $\Lambda$ is 30 degrees.)

## E. Rank

The goal of this section is a single important theorem. We've mentioned that reflections are the building blocks from which all transformations can be made. It's now time to make that statement exact, and prove it.

**Theorem 7.28.** *Every transformation is either the identity, a reflection, or the composition of two or three reflections.*

*Proof:* Let $f$ be a transformation. If $f$ fixes two points then it fixes a line $\Lambda$ (by Theorem 7.9) and thus is either $r_\Lambda$ or the identity (by Theorem 7.13). On the other hand, if $f$ fixes exactly one point, then $f$ is the composition of two reflections (by Corollary 7.23). So, we may assume that $f$ has *no* fixed points. Then:

- Choose a point $A$ and let $\Lambda$ be the perpendicular bisector of $Af(A)$.
- Then $r_\Lambda \circ f(A) = r_\Lambda(f(A)) = A$, so $r_\Lambda \circ f$ fixes $A$ but is not the identity (since $f$ is not a reflection).
- So, as noted above, $r_\Lambda \circ f = g$ where $g$ is either a reflection or a rotation (composition of two reflections).
- But $r_\Lambda$ is its own inverse, so in either case $f = id \circ f = (r_\Lambda \circ r_\Lambda) \circ f = r_\Lambda \circ (r_\Lambda \circ f) = r_\Lambda \circ g$, so $f$ is the composition of no more than three reflections. □

Based on this theorem we introduce an important classification tool for transformations – the *rank* of the transformation.

**Definition.** Let $f$ be a transformation. We define the *rank* of $f$ as follows:

- If $f = id$ then $\text{rank}(f) = 0$.
- If $f$ is a reflection then $\text{rank}(f) = 1$.
- If $f$ is a composition of two reflections (but is not the identity) then $\text{rank}(f) = 2$.
- If $f$ is a composition of three reflections (but is not itself a reflection) then $\text{rank}(f) = 3$.

From Theorem 7.18 we can see that even-rank transformations are orientation preserving, while odd-rank transformations are orientation reversing.

Let's take a quick inventory of where we stand in our quest to classify all transformations. Using fixed points as a classification tool, we have the following facts:

- The only transformation fixing three noncollinear points is the identity (Theorem 7.10).
- If $f$ is a transformation with two fixed points and is not the identity then $f$ is a reflection (Theorems 7.9 and 7.13).
- If $f$ has exactly one fixed point then $f$ is a (rank 2) rotation (Corollary 7.23).

What is left to investigate? As seen in the table below, the only transformations we have yet to classify have no fixed points and have rank either 2 or 3. We'll turn our attention to these two remaining cases in the next two sections.

| **Rank** | **Fixed Points** | **Type** |
|---|---|---|
| 0 | entire plane | identity |
| 1 | one line | reflection |
| 2 | one point | rotation |
| | none | ?? |
| 3 | none | ?? |

## Exercises

**7.32**. True or False? (Questions for discussion)

**(a)** The composition of two orientation preserving transformations will always be orientation preserving.

**(b)** It is possible for the composition of a reflection and a rotation to be the identity.

**(c)** No rank 3 transformation can have a fixed point.

**(d)** The composition of any three reflections will be a rank 3 transformation.

**(e)** All transformations can be expressed as the composition of rotations.

**7.33**. (For readers who have had a course in modern abstract algebra.) Let $\mathcal{G}$ be the group of all transformations of the plane. Prove that $\mathcal{E} = \{f \in \mathcal{G} \mid \text{rank}(f) \text{ is even}\}$ is a subgroup of $\mathcal{G}$. Is the same true of the set of transformations with odd rank? Explain.

**7.34**. Suppose $f$ is a rank 3 transformation. Prove that $f^{-1}$ is also rank 3.

## F. Translations

In this section we'll investigate one of the incomplete rows from the table of transformations on p.291. Specifically, we'll consider transformations of rank 2 that have no fixed points.

So let $f$ be such a transformation. Then (since it has rank 2) we know we can write $f$ as the composition of two reflections, say $f = r_{\Lambda'} \circ r_\Lambda$. Now if the lines $\Lambda$ and $\Lambda'$ were not parallel then their point of intersection would be fixed by both $r_\Lambda$ and $r_{\Lambda'}$ and thus would be fixed by $f$. But $f$ has no fixed points, so it must be that $\Lambda$ and $\Lambda'$ are parallel.

Thus there is only one possible recipe for making a rank 2 transformation with no fixed points. Our first theorem will prove that the recipe works.

**THEOREM 7.29.** *Let $\Lambda$ and $\Lambda'$ be distinct parallel lines, and let $f$ be the transformation $f = r_{\Lambda'} \circ r_\Lambda$. Then $f$ has no fixed points, but preserves all lines perpendicular to both $\Lambda$ and $\Lambda'$.*

*Proof:* Assume to reach a contradiction that $f$ has a fixed point $A$.

- Clearly $A$ cannot be on $\Lambda$, for otherwise we would have $f(A) = r_{\Lambda'} \circ r_\Lambda(A) = r_{\Lambda'}(A) \neq A$.
- Let $A'$ denote the point $r_\Lambda(A)$. Then by the definition of $r_\Lambda$, the line $\Lambda$ is the perpendicular bisector of segment $AA'$.
- But since $f(A) = A$ we must have $r_{\Lambda'}(A') = A$, so $\Lambda'$ is also the perpendicular bisector of $AA'$.
- This implies $\Lambda = \Lambda'$, a contradiction.

We leave it to the reader (see Exercise 7.35) to complete the proof by showing that $f$ preserves lines that are perpendicular to both $\Lambda$ and $\Lambda'$. $\square$

We can easily describe the action of $r_{\Lambda'} \circ r_\Lambda$ on points of a preserved line, as shown by our next theorem. We leave the elementary proof as an exercise.

**THEOREM 7.30.** *Let $\Lambda$ and $\Lambda'$ be parallel and let $f$ be the transformation $f = r_{\Lambda'} \circ r_\Lambda$. Suppose $A$ a point of $\Lambda$ and $D$ is a point of $\Lambda'$ such that $AD \perp \Lambda$ and $AD \perp \Lambda'$ (see Figure 7.25). Let $p(t) = P_t$ be the ruler function for $\overleftrightarrow{AD}$ so that $P_0 = A$ and $P_{|AD|} = D$. Then for all points $P_t$ on $\overleftrightarrow{AD}$ we have $f(P_t) = P_{t+2|AD|}$.*

This theorem justifies the name "translation" for such transformations, since it states that points of a preserved line are moved by a fixed distance along that line. In fact, suppose now that $A$ and $B$ are two arbitrary points. Choose $\Lambda$ to be the perpendicular to $\overleftrightarrow{AB}$ through $A$ and $\Lambda'$ to be the perpendicular bisector of $AB$, then clearly $f = r_{\Lambda'} \circ r_\Lambda$ moves $A$ to $B$ and preserves the line $\overleftrightarrow{AB}$. Accordingly, we give the following definition.

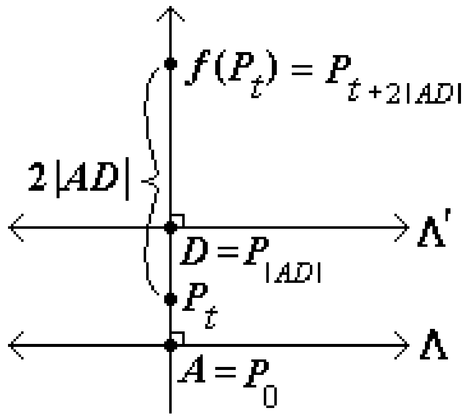



Figure 7.25:

**DEFINITION.** Given two points $A$ and $B$, the *translation from $A$ to $B$* is the transformation $f = r_{\Lambda'} \circ r_\Lambda$ where $\Lambda$ is the line through $A$ perpendicular to $AB$ and $\Lambda'$ is the perpendicular bisector of $AB$.

The theorems above show that the translation from $A$ to $B$ preserves the line $\overleftrightarrow{AB}$ and moves each of its points a distance of $|AB|$ along the line in the direction of the orientation $(A, B)$.

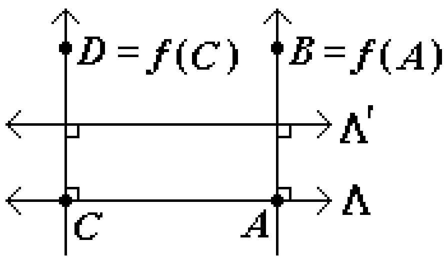



Figure 7.26: Euclidean translation from $A$ to $B$ (or from $C$ to $D$)

Here there is a distinct difference in the Euclidean and hyperbolic cases. For in the Euclidean case, to each point $C$ on $\Lambda$ there is a point $D$ on the other side of $\Lambda'$ so that

$CD \perp \Lambda$, $CD \perp \Lambda'$, and $|CD| = |AB|$. In this case, then, the translation $f = r_{\Lambda'} \circ r_\Lambda$ also preserves $\overleftrightarrow{CD}$ and might also be called the translation from $C$ to $D$ (see Figure 7.26). In fact, given any point $P$ and its image $Q = f(P)$, $f$ is the only translation moving $P$ to $Q$.

On the other hand, in the hyperbolic case of the above definition the lines $\Lambda$ and $\Lambda'$ are ultraparallel by Theorem 6.27 and $\overleftrightarrow{AB}$ is the *only* line perpendicular to both, and the only line preserved by $f = r_{\Lambda'} \circ r_\Lambda$ (see Exercise 7.37). And given any point $P$ not on $\overleftrightarrow{AB}$ and its image $Q = f(P)$, while it is true that $f$ is *a* translation moving $P$ to $Q$, $f$ is definitely not *the* translation from $P$ to $Q$, for it does not preserve the line $\overleftrightarrow{PQ}$ (see Exercise 7.40).

We'll learn more about these differences as we look at examples in the Euclidean and hyperbolic models.

### Translations in the Euclidean model

Translations are the easiest of all transformations to describe in the Euclidean model. For let $f = r_{\Lambda'} \circ r_\Lambda$ be the translation from $A$ to $B$ as in the definition. Then each point of the plane belongs to a line preserved by $f$ (see Exercise 7.39), so Theorem 7.30 tells us that for each point $\begin{pmatrix} x \\ y \end{pmatrix}$ in the model, $f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right)$ will be $\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} x+a \\ y+b \end{pmatrix}$ where $\begin{pmatrix} a \\ b \end{pmatrix}$ is the vector in the direction of the oriented line ${}_{(A,B)}\overleftrightarrow{AB}$ (see Figure 7.27) and with length $\sqrt{a^2+b^2}$ equal to $|AB|$ (which is equal to twice the distance between $\Lambda$ and $\Lambda'$).

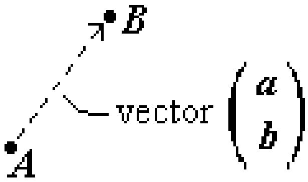



Figure 7.27:

If $f$ is translation from $A$ to $B$ then in the Euclidean model we can write $f$ as

$$f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x+a \\ y+b \end{pmatrix}$$

where $\begin{pmatrix} a \\ b \end{pmatrix}$ is the vector directed from $A$ to $B$.

**EXAMPLE 7.31.** *Describe all possible orientation preserving transformations that move the point* $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ *to the point* $\begin{pmatrix} -1 \\ -1 \end{pmatrix}$.

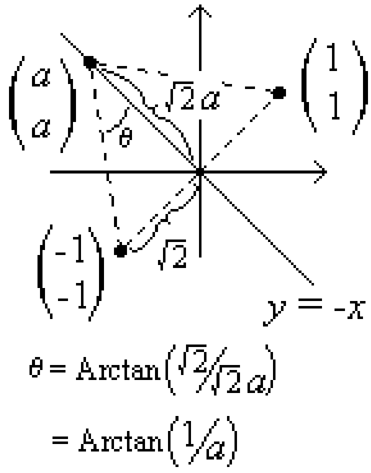



Figure 7.28:

*Solution:* Orientation preserving transformations have even rank, so a quick reference to the table on p.291 shows that the only candidates are rotations and translations.

- The only Euclidean translation that moves $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ to $\begin{pmatrix} -1 \\ -1 \end{pmatrix}$ is

$$f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x-2 \\ y-2 \end{pmatrix}.$$

- The rotations that move $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ to $\begin{pmatrix} -1 \\ -1 \end{pmatrix}$ must be about points on the line $y = -x$, as seen from Figure 7.28.
  - In fact, we can see from that figure that when $a \neq 0$, the rotation about $\begin{pmatrix} a \\ a \end{pmatrix}$ should be by angle measure $\pm 2\text{Arctan}(1/a)$.
  - Finally, we also need to include 180 degree rotation about the origin.

These rotations could easily be written out in vector notation using the formulae from Section D. □

**Translations in the hyperbolic model**

Now consider the hyperbolic case of a translation $f = r_{\Lambda'} \circ r_\Lambda$ from $A$ to $B$. As we noted above (see p.294), the lines $\Lambda$ and $\Lambda'$ must be ultraparallel (since they have a common perpendicular $\overleftrightarrow{AB}$) and $f$ has only one preserved line, namely $\overleftrightarrow{AB}$. There are two interesting observations to make regarding this:

**1.** Hyperbolic translations don't "look like" Euclidean ones! The image of points sliding along an infinite family of parallel preserved lines simply does not hold in the hyperbolic case.

**2.** There is yet another possible type of rank 2 hyperbolic transformations: those resulting from successive reflections across asymptotically parallel lines. We'll see in Section I that this does in fact lead to a new and different variety of transformations (called the *ideal point rotations*).

We won't compute a general formula for an arbitrary hyperbolic translation, but we will show how our formulae for hyperbolic reflections can be used to find translation formulae in specific instances.

EXAMPLE **7.32.** *Find two different hyperbolic translations that move the point* $(0,2)$ *in the upper halfplane model to the point* $(0,10)$.

*Solution:* Our two translations will be $f_1 = r_{\Lambda'} \circ r_{\Lambda_1}$ and $f_2 = r_{\Lambda'} \circ r_{\Lambda_2}$ where $\Lambda'$ is the perpendicular H-bisector of the H-segment between $(0,2)$ and $(0,10)$ and $\Lambda_1$ and $\Lambda_2$ are different H-lines parallel to $\Lambda'$ and passing through $(0,2)$. That way, $(0,2)$ will be fixed by both $r_{\Lambda_1}$ and $r_{\Lambda_2}$, so $f_1(0,2) = f_2(0,2) = r_{\Lambda'}(0,2) = (0,10)$.

Clearly $\Lambda'$ will be an H-line taking the form of a Euclidean circle centered at the origin. We want $(0,10)$ to be the reflection of $(0,2)$ through this circle, so if its radius is $r$ then we should have $r^2 = (2)(10)$, or $r = \sqrt{20}$. Thus $\Lambda'$ is the H-line with equation $x^2 + y^2 = 20$. Using our results from Section B we see that the formula for $r_{\Lambda'}$ is

$$r_{\Lambda'}(x,y) = \left(\frac{20x}{x^2+y^2}, \frac{20y}{x^2+y^2}\right).$$

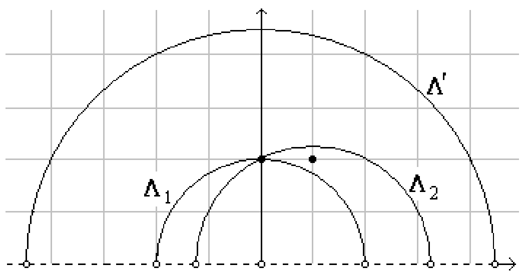



Figure 7.29:

For $\Lambda_1$ we'll use the H-line $x^2 + y^2 = 4$ and for $\Lambda_2$ we'll use $(x-1)^2 + y^2 = 5$ (see Figure 7.29). The formulae for the reflections across these H-lines are

$$\begin{aligned} r_{\Lambda_1}(x,y) &= \left(\frac{4x}{x^2+y^2}, \frac{4y}{x^2+y^2}\right), \text{ and} \\ r_{\Lambda_2}(x,y) &= \left(\frac{5(x-1)}{(x-1)^2+y^2} + 1, \frac{5y}{(x-1)^2+y^2}\right). \end{aligned}$$

It's now easy to check that $f_1(0,2) = r_{\Lambda'} \circ r_{\Lambda_1}(0,2) = r_{\Lambda'}(0,2) = (0,10)$ and similarly $f_2(0,2) = (0,10)$. However, $f_1$ and $f_2$ are definitely different transformations, as can be seen by evaluating them at a point not on the $y$-axis:

$$\begin{aligned} f_1(1,2) &= r_{\Lambda'} \circ r_{\Lambda_1}(1,2) = r_{\Lambda'}(4/5, 8/5) = (5, 10) \\ f_2(1,2) &= r_{\Lambda'} \circ r_{\Lambda_2}(1,2) = r_{\Lambda'}(1, 5/2) = (80/29, 200/29) \end{aligned}$$

□

## Exercises

**7.35**. Complete the proof of Theorem 7.29 by showing that if $\Lambda$ and $\Lambda'$ are parallel lines and $\Lambda''$ is perpendicular to both of them, then $f = r_{\Lambda'} \circ r_\Lambda$ preserves $\Lambda''$.

**7.36**. Prove Theorem 7.30.

**7.37**. Let $\Lambda$ and $\Lambda'$ be ultraparallel lines in the hyperbolic plane, so that $f = r_{\Lambda'} \circ r_\Lambda$ is a translation of the hyperbolic plane. Let $\Lambda''$ be the unique line perpendicular to both $\Lambda$ and $\Lambda'$ (see Theorem 6.27). Prove that $\Lambda''$ is the only line preserved by $f$. (*Hint:* let $P$ be a point not on $\Lambda''$ and consider the points $\{P, f(P), f(f(P)), f(f(f(P))), \ldots\}$. For any line through $P$ to be preserved, this set would have to be collinear. Show that it is not by also considering the points $\{Q, f(Q), f(f(Q)), f(f(f(Q))), \ldots\}$ where $Q$ is the point on $\Lambda''$ with $PQ \perp \Lambda''$.)

**7.38**. Suppose that $r_{\Lambda'} \circ r_\Lambda$ is translation from $A$ to $B$. What is $r_\Lambda \circ r_{\Lambda'}$? Justify your answer.

**7.39**. Let $f$ be translation from $A$ to $B$ in the Euclidean plane. Prove that each point of the plane belongs to a line preserved by $f$.

**7.40**. Let $A$ and $B$ be points in the hyperbolic plane. Show how to find infinitely many distinct hyperbolic translations all of which move $A$ to $B$. (*Hint:* all of them can be written as $r_{\Lambda'} \circ r_\Lambda$ where $\Lambda'$ be the perpendicular bisector of $AB$.)

**7.41**. Let $\Lambda$ and $\Lambda'$ be the lines $y = mx$ and $y = m(x - d)$ in the Euclidean model. Use our formulae for reflections in Section B to compute a formula for $r_{\Lambda'} \circ r_\Lambda$. Your result should reduce to $\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix}$ where $\begin{pmatrix} a \\ b \end{pmatrix}$ is a vector with slope $-1/m$ (assuming that $m \neq 0$).

**7.42**. (For readers who have had a course in modern abstract algebra.) Let $\Lambda$ be a line and let $\mathcal{P}^+(\Lambda)$ be the set of all orientation preserving transformations that preserve $\Lambda$. Describe the contents of $\mathcal{P}^+(\Lambda)$. Is $\mathcal{P}^+(\Lambda)$ a subgroup of the group $\mathcal{G}$ of all transformations?

**7.43**. Give formulae for three different orientation preserving transformations (in the Euclidean model) that move the point $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ to the point $\begin{pmatrix} 4 \\ -2 \end{pmatrix}$.

**7.44**. As in Example 7.32, compute $f_1(4,5)$ and $f_2(4,5)$ where $f_1$ and $f_2$ are two different hyperbolic translations moving the point $(0,1)$ of the upper halfplane model to the point $(0,8)$.

**7.45**. As in Example 7.32, compute $f_1(1,2)$ and $f_2(1,2)$ where $f_1$ and $f_2$ are two different hyperbolic translations moving the point $(2,3)$ of the upper halfplane model to the point $(-2,3)$.

## G. Glide Reflections

Refer again to the table on p.291. The only line of that table we have yet to investigate is the rank 3 case. That will be our goal in this section.

There are several combinations of our previous transformation types that might give us different kinds of rank 3 transformations. In particular, there is no immediate reason to suspect that the following are not all distinctly different varieties:

**(a)** Reflection across a line through $A$ followed by a rotation about $A$.

**(b)** Reflection across a line followed by a rotation about a point not on that line.

**(c)** Reflection across a line followed by a translation parallel to that line.

**(d)** Reflection across a line followed by a translation not parallel to that line.

**(e)-(h)** Any of the above with the reflection following the rotation or translation.

However, somewhat remarkably, item **(c)** in this list includes all of the others. That is, *every* rank 3 transformation can be written in the form of item **(c)**.

**DEFINITION.** Suppose $g$ is the translation from $A$ to $B$ and $\Lambda$ is the line $\overleftrightarrow{AB}$. Then the transformation $g \circ r_\Lambda$ is called a *glide reflection*.

Note that this is definitely a rank 3 transformation. We can see that its rank is odd because it is orientation reversing (being the composition of an orientation reversing reflection and an orientation preserving translation). But it is clearly *not* a rank 1 reflection, for it has no fixed points (see Exercise 7.46). This is consistent with what we've previously deduced: all transformations with fixed points are either reflections or rotations (or the identity) and so have rank less than 3.

But while a glide reflection has no fixed points, it clearly does have one (and only one) preserved line, namely the line $\Lambda$ that defines the reflection and direction of translation. We can specify a glide reflection by merely identifying its preserved line as well as the distance along that line by which the translation acts.

Another important note on the above definition concerns the order in which the component motions of a glide reflection are performed. Composition of transformations is generally not a commutative operation: $f \circ g$ is generally not the same as $g \circ f$. However, the translation $g$ from $A$ to $B$ *does* commute with the reflection across $\Lambda = \overleftrightarrow{AB}$ (see Exercise 7.47). So, while we've defined the glide reflection to be $g \circ r_\Lambda$, it can also be taken to be $r_\Lambda \circ g$.

We'll now proceed with our principal theorem for this section. As you work through its proof, note how it outlines a procedure for identifying the preserved line and the translation in any glide reflection. We'll use that procedure in the examples that follow the proof.

**THEOREM 7.33.** *Every rank 3 transformation is a glide reflection.*

*Proof:* Let $f$ be a rank 3 transformation. We will show that $f$ is a glide reflection. We know that $f$ has no fixed points, for any transformation fixing more than one point is either the identity or a reflection (Theorem 7.13) and any transformation fixing exactly one point is a rotation (Corollary 7.23). Let $P$ be any point and let $\Lambda$ be the perpendicular bisector of segment $Pf(P)$.

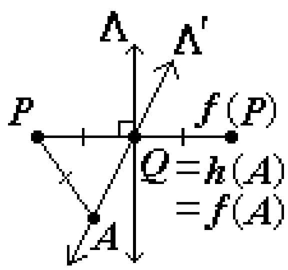



Figure 7.30:

- Then $h = r_\Lambda \circ f$ fixes the point $P$ and so must have rank 2. (Note that $h$ cannot have rank 1: since $f$ is rank 3 and thus is orientation reversing, $h$ must be orientation preserving.)

- Thus $h$ must be a rotation about the point $P$.

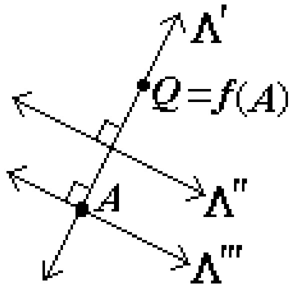



Figure 7.31:

- As in Figure 7.30, let $Q$ be the midpoint of $Pf(P)$ (so that $Q$ is on $\Lambda$), and let $A$ be the point so that $h(A) = Q$.

- Then since $f$ can be written as $r_\Lambda \circ h$, we have $f(A) = r_\Lambda(h(A)) = r_\Lambda(Q) = Q$.

- But it is easy to check (see Exercise 7.48) that $f(Q)$ is on $\overleftrightarrow{QA}$, so $f$ must preserve the line $\Lambda'$ $= \overleftrightarrow{QA}$.

- Now let $\Lambda''$ be the perpendicular bisector of $QA$. Then $r_{\Lambda''} \circ f$ both preserves $\Lambda'$ and fixes $A$ (see Figure 7.31).

- But $r_{\Lambda''} \circ f$ must have rank 2 (as before, it can't be rank 1) and so must have only the one fixed point $A$. Clearly, then, $r_{\Lambda''} \circ f$ is a rotation by angle measure 180 about $A$.

- So if $\Lambda'''$ is the line through $A$ perpendicular to $\Lambda'$, then $r_{\Lambda'''} \circ r_{\Lambda''} \circ f$ must fix all of $\Lambda'$. Since this composition is orientation reversing, it must be that $r_{\Lambda'''} \circ r_{\Lambda''} \circ f = r_{\Lambda'}$ (see Theorem 7.13).

- Thus we have $f = (r_{\Lambda''} \circ r_{\Lambda'''}) \circ r_{\Lambda'}$, where, of course, $r_{\Lambda''} \circ r_{\Lambda'''}$ is the translation from $A$ to $Q$. This shows that $f$ is a glide reflection, as claimed. □

The procedure implicit in this proof is so useful that we'll pause here to outline it explicitly, then illustrate it in some examples.

**Procedure to identify the preserved line and translation vector of a glide reflection $f$:**

- Choose a point $P$ and let $\Lambda$ be the perpendicular bisector of $Pf(P)$.
- $r_\Lambda \circ f$ will be a rotation $h$ about $P$.
- Let $Q$ be the midpoint of $Pf(P)$ and find the point $A$ so that $h(A) = Q$.
- Then $f$ is reflection across the preserved line $\overleftrightarrow{QA}$ followed by translation from $A$ to $Q$.

EXAMPLE **7.34.** *Let $f$ be the transformation carrying the triangle with vertices $\left\{\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ -2 \end{pmatrix}\right\}$ in the Euclidean model to the triangle with vertices $\left\{\begin{pmatrix} -2 \\ 1 \end{pmatrix}, \begin{pmatrix} -4 \\ 2 \end{pmatrix}, \begin{pmatrix} -5 \\ 1 \end{pmatrix}\right\}$. Show that $f$ is a glide reflection by identifying its preserved line and translation vector.*

*Solution:* First note that by Theorem 7.11 there is only one such transformation. We'll use the above procedure in identifying its action.

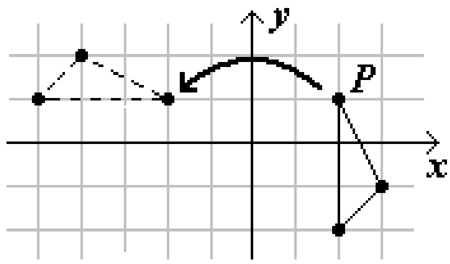



Figure 7.32:

- First, since $f$ is a transformation that carries one of the triangles in Figure 7.32 to the other, it is clear that the point $P = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ must be carried to $f(P) = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$.
- Now the perpendicular bisector of $Pf(P)$ is clearly the vertical line $x = 0$. Accordingly, we'll call this line $\Lambda$. As predicted by our procedure, following $f$ by reflection across $\Lambda$ yields a rotation $h$ about $P$ – in this case a counterclockwise rotation by 90 degrees (see the leftmost part of Figure 7.33).
- The preserved line should be $\Lambda' = \overleftrightarrow{QA}$ where $Q$ is the midpoint $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ of $Pf(P)$ and $A$ is the point so that $h(A) = Q$. Clearly, $A$ is the point $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ (obtained by rotating $Q = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ by 90 degrees *clockwise* about $P = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ – see the middle portion of Figure 7.33). So $\Lambda'$ has equation $y = 1 + x$.

- Our glide reflection consists of a reflection across this line $\Lambda'$ followed by a translation from $A$ to $Q$ (translation vector $\begin{pmatrix} -2 \\ -2 \end{pmatrix}$). This is illustrated in the rightmost portion of Figure 7.33. $\square$

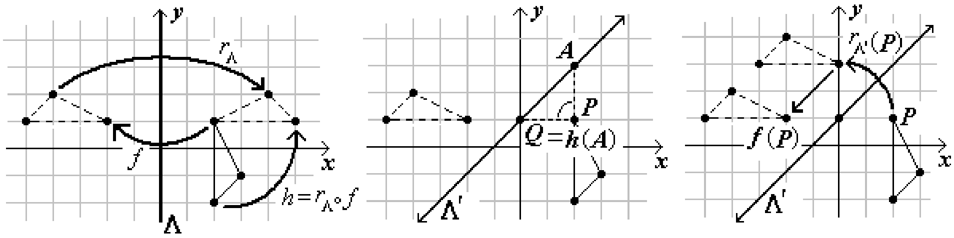



Figure 7.33:

EXAMPLE **7.35.** *Find all transformations in the Euclidean model that carry* $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ *to* $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ *and carry* $\begin{pmatrix} 3 \\ 4 \end{pmatrix}$ *to* $\begin{pmatrix} 5 \\ 2 \end{pmatrix}$.

*Solution:* We should expect to find two such transformations – one orientation preserving and one orientation reversing. For, as in Figure 7.34, let $P$ and $R$ denote the points $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 3 \\ 4 \end{pmatrix}$, and let $S$ be the point $\begin{pmatrix} 3 \\ 3 \end{pmatrix}$. If $f$ is such a transformation then the triangle $\triangle f(P)f(R)f(S)$ must be congruent to $\triangle PRS$, so $f(S)$ must be one of the two points shown in the figure. But Theorem 7.11 tells us there can only be one transformation for each of these possibilities.

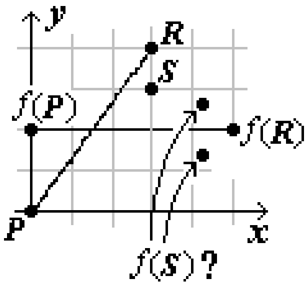



Figure 7.34:

Example 7.26 in Section D gives us one of these two transformations – a clockwise rotation about the point $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$. The other transformation must then be orientation reversing and thus of odd rank. It seems evident that no reflection will work (for the perpendicular bisectors of $Pf(P)$ and $Rf(R)$ are different), so the transformation we are looking for must be a glide reflection. We'll find its preserved line and translation vector using our procedure.

- Let $\Lambda$ be the perpendicular bisector $y = 1$ of $Pf(P)$, as in the leftmost portion of Figure 7.35.

- Clearly, $h = r_\Lambda \circ f$ is the clockwise rotation $h$ by angle measure 45 about the point $P$.

- So (see the middle portion of Figure 7.35), the line preserved by $f$ is $\Lambda' = \overleftrightarrow{QA}$ where $Q$ is the midpoint $\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$ of $Pf(P)$ and $A$ is the point so that $h(A) = Q$.

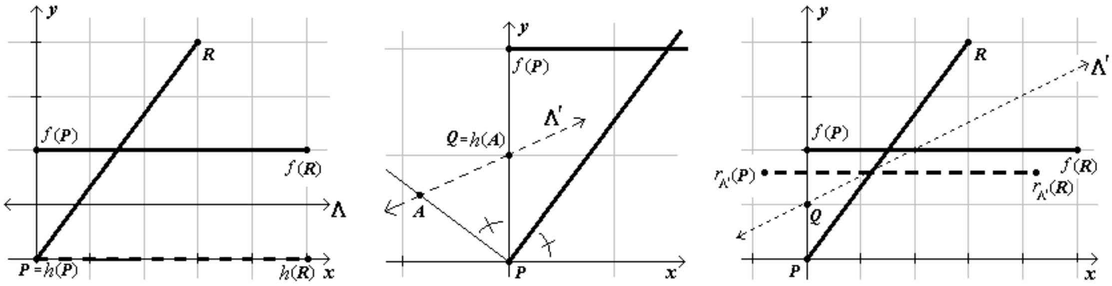



Figure 7.35:

- It is clear that since the slope of $PR$ is $4/3$, the slope of $PA$ will be $-3/4$ (refer again to the middle portion of Figure 7.35), so $A$ is on the line $y = \frac{-3}{4}x$ and has distance $|QP| = 1$ from $P = \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$. From this it's easy to deduce that $A$ must be the point $\left(\begin{smallmatrix} -4/5 \\ 3/5 \end{smallmatrix}\right)$.

- Then $\Lambda' = \overleftrightarrow{QA}$ has slope $1/2$ and so has equation $y = \frac{1}{2}x + 1 = \frac{1}{2}(x+2)$.

- Using our formula for reflections in the Euclidean model (see p.266) we have

$$\begin{aligned} r_{\Lambda'}\left(\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)\right) &= \frac{4}{5}\left(\begin{smallmatrix} 3/4 & 1 \\ 1 & -3/4 \end{smallmatrix}\right)\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) - \frac{8}{5}\left(\begin{smallmatrix} 1/2 \\ -1 \end{smallmatrix}\right) \\ &= \left(\begin{smallmatrix} 3/5 & 4/5 \\ 4/5 & -3/5 \end{smallmatrix}\right)\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) + \left(\begin{smallmatrix} -4/5 \\ 8/5 \end{smallmatrix}\right). \end{aligned}$$

- The translation vector (from $A$ to $Q$) is $\left(\begin{smallmatrix} 4/5 \\ 2/5 \end{smallmatrix}\right)$.

The rightmost diagram in Figure 7.35 illustrates the relationship between $r_{\Lambda'}(PR)$ and $f(PR)$. Combining the formula for $r_{\Lambda'}$ above with translation by $\left(\begin{smallmatrix} 4/5 \\ 2/5 \end{smallmatrix}\right)$ we get the following formula for the glide reflection $f$:

$$f\left(\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)\right) = r_{\Lambda'}\left(\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)\right) + \left(\begin{smallmatrix} 4/5 \\ 2/5 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 3/5 & 4/5 \\ 4/5 & -3/5 \end{smallmatrix}\right)\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) + \left(\begin{smallmatrix} 0 \\ 2 \end{smallmatrix}\right).$$

We'll leave it to you to check that with this formula we have $f\left(\left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)\right) = \left(\begin{smallmatrix} 0 \\ 2 \end{smallmatrix}\right)$ and $f\left(\left(\begin{smallmatrix} 3 \\ 4 \end{smallmatrix}\right)\right) = \left(\begin{smallmatrix} 5 \\ 2 \end{smallmatrix}\right)$. □

EXAMPLE **7.36.** *Let $f$ be the glide reflection in the halfplane model for hyperbolic geometry given by reflecting across the H-line $x = 0$ then translating from $(0,1)$ to $(0,4)$. Calculate the image under $f$ of the H-triangle with vertices* $\{(2,1),(2,3),(3,1)\}$.

*Solution:* Reflecting across $x = 0$ is easy, so let's first think about the formula for translation from $(0,1)$ to $(0,4)$. By definition, this translation can be written as $g = r_{\Lambda'} \circ r_\Lambda$ where $\Lambda$ is the H-line through $(0,1)$ perpendicular to $x = 0$ and $\Lambda'$ is the perpendicular H-bisector between $(0,1)$ and $(0,4)$. Clearly $\Lambda$ is given by $x^2 + y^2 = 1$, and since $(0,4)$ is the reflection of $(0,1)$ through $x^2 + y^2 = 4$, this last equation gives us the H-line $\Lambda'$. Using the formulae from Section F, the translation $g$ is then given by

$$\begin{aligned} g(x,y) &= r_{\Lambda'} \circ r_\Lambda(x,y) \\ &= r_{\Lambda'}\left(\frac{x}{x^2+y^2}, \frac{y}{x^2+y^2}\right) \\ &= \left(\frac{4\left(\frac{x}{x^2+y^2}\right)}{\left(\frac{x}{x^2+y^2}\right)^2 + \left(\frac{y}{x^2+y^2}\right)^2}, \frac{4\left(\frac{y}{x^2+y^2}\right)}{\left(\frac{x}{x^2+y^2}\right)^2 + \left(\frac{y}{x^2+y^2}\right)^2}\right) \\ &= (4x, 4y)\,. \end{aligned}$$

So if $\Lambda''$ is the H-line $x = 0$ then the desired glide reflection is

$$f(x,y) = g \circ r_{\Lambda''}(x,y) = g(-x,y) = (-4x, 4y)\,.$$

So, the vertices $(2,1)$, $(2,3)$, and $(3,1)$ are carried to $(-8,4)$, $(-8,12)$, and $(-12,4)$ respectively (see Figure 7.36). □

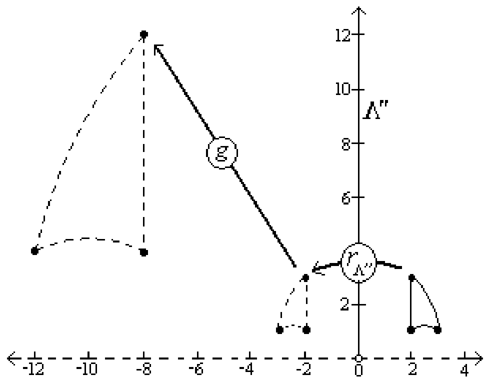



Figure 7.36:

## Exercises

**7.46**. Show that a glide reflection has no fixed points.

**7.47**. Let $g$ be translation from $P$ to $Q$ and let $\Lambda$ be the line $\overleftrightarrow{PQ}$. Show that $g \circ r_\Lambda = r_\Lambda \circ g$. Your proof should use only neutral geometry. (By definition, $g$ is given by $r_{\Lambda''} \circ r_{\Lambda'}$ where $\Lambda'$ is the line through $P$ perpendicular to $\Lambda$ and $\Lambda''$ is the perpendicular bisector of $PQ$.)

**7.48**. In the proof of Theorem 7.33 show that $f(Q)$ is on $\Lambda'$ (and thus that $f$ preserves $\Lambda' = \overleftrightarrow{QA}$).

**7.49**. In each case, identify the glide reflection $f$ by identifying its preserved line and translation vector.

**(a)** $f\left(\begin{pmatrix} 1 \\ 2 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$, $f\left(\begin{pmatrix} 2 \\ -1 \end{pmatrix}\right) = \begin{pmatrix} 4 \\ -3 \end{pmatrix}$, and $f\left(\begin{pmatrix} 4 \\ -2 \end{pmatrix}\right) = \begin{pmatrix} 5 \\ -5 \end{pmatrix}$.

**(b)** $f\left(\begin{pmatrix} 3 \\ -2 \end{pmatrix}\right) = \begin{pmatrix} 3 \\ -3 \end{pmatrix}$, $f\left(\begin{pmatrix} 1 \\ 4 \end{pmatrix}\right) = \begin{pmatrix} -3 \\ -1 \end{pmatrix}$, and $f\left(\begin{pmatrix} -1 \\ -3 \end{pmatrix}\right) = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$.

**7.50**. Find all transformations in the Euclidean model that carry $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ to $\begin{pmatrix} 0 \\ 4 \end{pmatrix}$ and carry $\begin{pmatrix} 2\sqrt{2} \\ 0 \end{pmatrix}$ to $\begin{pmatrix} 2 \\ 2 \end{pmatrix}$.

**7.51**. Find all transformations in the Euclidean model that carry the segment $AB$ to the segment $CD$. (You should find *four* in each case.)

**(a)** $A = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $B = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$, $C = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $D = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$.

**(b)** $A = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $B = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$, $C = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $D = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$.

**7.52**. Let $f$ be the glide reflection in the halfplane model for hyperbolic geometry given by reflecting across the H-line $x = 3$ then translating from $(3, 1)$ to $(3, 3)$. Calculate the image under $f$ of the H-triangle with vertices $\{(0, 2), (0, 4), (2, 2)\}$.

## H. Euclidean Transformations

Our work in the previous sections has given us a complete description of all transformations of the Euclidean plane. In this section we'll consider the task of identifying a Euclidean transformation by using what we've learned. Let's begin by summarizing the classification scheme. We found that every transformation can be assigned a rank between 0 and 3, with the following possibilities in each case.

**Rank 0:** Only the identity has rank 0.

**Rank 1:** The rank 1 transformations are the reflections across lines.

**Rank 2:** There are two possibilities for a rank 2 transformation $r_{\Lambda'} \circ r_\Lambda$:

- **(a)** If $\Lambda$ and $\Lambda'$ are not parallel, then $r_{\Lambda'} \circ r_\Lambda$ is a rotation about their point of intersection.
- **(b)** But if $\Lambda$ and $\Lambda'$ *are* parallel then $r_{\Lambda'} \circ r_\Lambda$ is a translation.

**Rank 3:** All rank 3 transformations are glide reflections.

This is summarized in the following table.

**Table of Euclidean Transformations**

| Rank | Fixed Points | Type |
|---|---|---|
| 0 | entire plane | identity |
| 1 | one line | reflection |
| 2 | one point | rotation |
| | none | translation |
| 3 | none | glide reflection |

Keeping in mind that even rank transformations are orientation preserving while odd rank transformations are orientation reversing, we can use the information in this table to identify a transformation. The examples below illustrate the process.

EXAMPLE **7.37.** *Suppose $f$ is counterclockwise rotation by angle measure 90 about the point $A = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $g$ is translation from $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ to $\begin{pmatrix} 0 \\ 3 \end{pmatrix}$. What is the transformation $g \circ f$?*

*Solution:* Since both $f$ and $g$ are orientation preserving, their composition will also be orientation preserving. But $f$ and $g$ are clearly not inverses of each other, so $g \circ f$ cannot be the identity. So $g \circ f$ must have rank 2 and is thus either a rotation or a translation. We can distinguish between these two possibilities by using fixed points.

Does $g \circ f$ have a fixed point? Indeed it does! As shown in Figure 7.37, the point $B = \begin{pmatrix} -1/2 \\ 7/2 \end{pmatrix}$ is taken by $f$ to the point $\begin{pmatrix} -1/2 \\ 1/2 \end{pmatrix}$ which in turn is returned by $g$ to $\begin{pmatrix} -1/2 \\ 7/2 \end{pmatrix}$. So, $g \circ f$ must be a rotation about $B$. Furthermore, we can easily tell that this rotation is (like $f$) counterclockwise by angle measure 90 by examining its action on $A$: $g \circ f(A) = g(A) = \begin{pmatrix} 1 \\ 5 \end{pmatrix}$, so clearly $m\angle ABg(f(A)) = 90$. $\square$

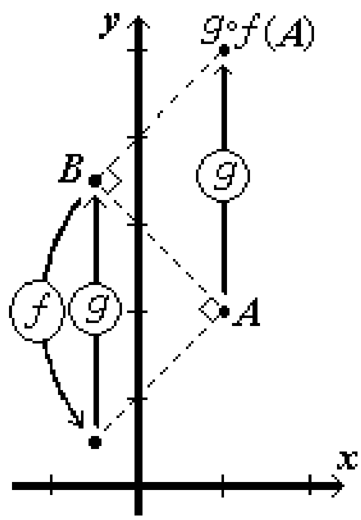



Figure 7.37:

EXAMPLE **7.38.** *Suppose $f$ is clockwise rotation by angle measure 60 about the point $A = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $g$ is clockwise rotation by angle measure 90 about the point $B = \begin{pmatrix} 4 \\ 0 \end{pmatrix}$. What is the transformation $g \circ f$?*

*Solution:* Similar to Example 7.37 above, the composition $g \circ f$ must be orientation preserving and thus be either a rotation or a translation, depending on whether or not it has a fixed point. But as indicated in Figure 7.38, $g \circ f$ clearly *does* have a fixed point – the point labeled $C$ in that figure. To find the coordinates of $C$, note that triangle $BDC$ is isosceles (since $m\angle DBC = m\angle DCB = 45$) so we may let $x$ be the length $|BD| = |DC|$. Then the length of $AD$ is $\sqrt{3}\,x$ since

$ACD$ is a 30-60-90 triangle (recall Lemma 2.8). Thus, $(1+\sqrt{3})x = |AB| = 4$, so $x = \frac{4}{1+\sqrt{3}}$. So $C$ must be the point $\frac{1}{1+\sqrt{3}}\begin{pmatrix} 4\sqrt{3} \\ 4 \end{pmatrix}$.

To calculate the amount of rotation, consider the point $E$ as in the figure. Here, $ABE$ is an equilateral triangle so that $f$ carries $E$ to $B$, then subsequently $g$ leaves $B$ fixed. In other words, $g \circ f(E) = B$, so $m\angle ECB$ will give the amount of (clockwise) rotation about $C$. But triangle $ECB$ is isosceles and $m\angle CBE$ is $60 - 45 = 15$. So (by the 180° Sum Theorem) $m\angle ECB$ must be 150. Thus $g \circ f$ is a clockwise rotation about $C$ by angle measure 150. $\square$

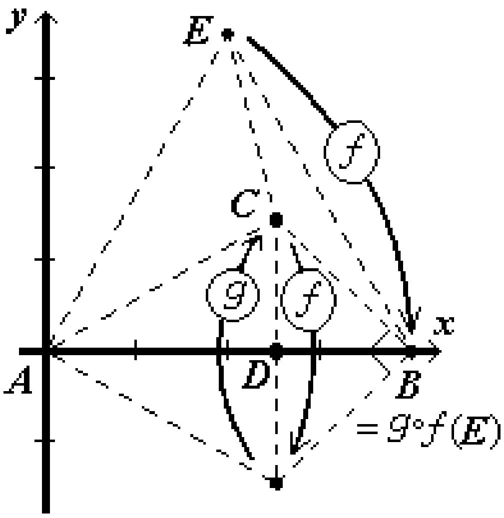



Figure 7.38:

EXAMPLE **7.39.** *Suppose $g$ is clockwise rotation by angle measure 90 about the point $P = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $\Lambda$ is the line $x = 2$. What is the transformation $f = r_\Lambda \circ g$?*

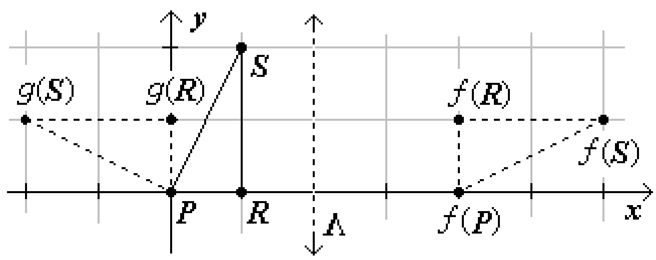



Figure 7.39:

*Solution:* Note that $f$ is the composition of an orientation preserving rotation and an orientation reversing reflection. So $f$ will itself be orientation reversing and thus must be a reflection or a glide reflection. Let $R$ and $S$ be the points $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ (as in Figure 7.39). It's clear that the segments $Pf(P)$, $Rf(R)$, and $Sf(S)$ do not all have the same perpendicular bisector, so $f$ is not a reflection. The only possibility, then, is that $f$ is a glide reflection, and the procedure from the proof of Theorem 7.33 (see p.301) will identify it. (Refer to Figure 7.40 in the analysis below.)

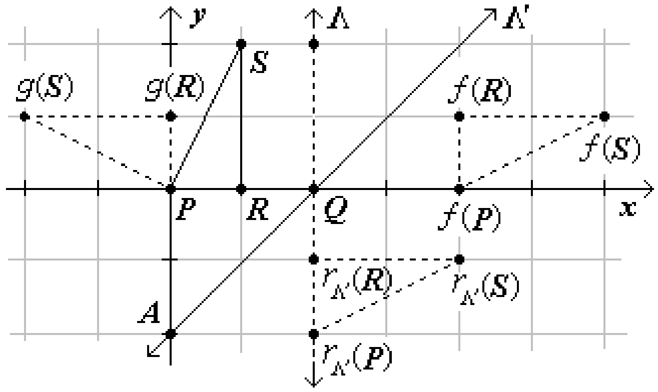



Figure 7.40:

- The perpendicular bisector of $Pf(P)$ is $\Lambda$ itself and $r_\Lambda \circ f = r_\Lambda \circ (r_\Lambda \circ g) = g$.
- So the preserved line for $f$ is $\Lambda' = \overleftrightarrow{QA}$ where $Q$ is the midpoint $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$ of $Pf(P)$ and $A = \begin{pmatrix} 0 \\ -2 \end{pmatrix}$ is the point such that $g(A) = Q$. This is clearly the line $y = x - 2$.
- So $f$ is reflection across $y = x - 2$ followed by translation from $A$ to $Q$ (translation vector $\begin{pmatrix} 2 \\ 2 \end{pmatrix}$). $\square$

## Exercises

**7.53**. With $f$ and $g$ as in Example 7.37, what is the transformation $f \circ g$. Justify your answer.

**7.54**. Suppose $g$ and $\Lambda$ are as in Example 7.39. What is the transformation $h = g \circ r_\Lambda$?

**7.55**. In each case, identify the transformation $g \circ f$.

**(a)** $f$ is clockwise rotation by angle measure 90 about $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $g$ is translation from $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ to $\begin{pmatrix} 3 \\ 3 \end{pmatrix}$.

**(b)** $f$ is counterclockwise rotation by angle measure 180 about $\binom{0}{0}$ and $g$ is translation from $\binom{0}{0}$ to $\binom{2}{5}$.

**(c)** $f$ is clockwise rotation by angle measure 90 about $\binom{0}{0}$ and $g$ is clockwise rotation by angle measure 60 about $\binom{0}{4}$.

**(d)** $f$ is clockwise rotation by angle measure 90 about $\binom{0}{0}$ and $g$ is counterclockwise rotation by angle measure 60 about $\binom{0}{4}$.

**(e)** $f$ is clockwise rotation by angle measure 90 about $\binom{0}{0}$ and $g$ is clockwise rotation by angle measure 180 about $\binom{0}{4}$.

**(f)** $f$ is translation from $\binom{0}{0}$ to $\binom{4}{0}$ and $g$ is reflection across $x = 0$.

**(g)** $f$ is reflection across $x = 0$ and $g$ is counterclockwise rotation by angle measure 90 about $\binom{0}{0}$.

**(h)** $f$ is counterclockwise rotation about $\binom{0}{0}$ by angle measure 180 and $g$ is reflection across $y = 1$.

**(i)** $f$ is reflection across $x = 0$ and $g$ is reflection across $y = 2$.

**(j)** $f$ is reflection across $y = 0$ and $g$ is translation from $\binom{0}{0}$ to $\binom{-3}{-4}$.

**7.56**. Can the composition of two rotations ever be a translation? Explain and justify your answer.

**7.57**. Can the composition of two rotations ever be a glide reflection? Explain and justify your answer.

**7.58**. Suppose $g$ is a glide reflection with preserved line $\Lambda$ and $\Lambda'$ is a line perpendicular to $\Lambda$. What type of transformation will $r_{\Lambda'} \circ g$ be? Justify your answer.

**7.59**. Suppose $g$ is a rotation with fixed point $A$ and $\Lambda$ is a line through $A$. What type of transformation will $r_\Lambda \circ g$ be? Justify your answer.

**7.60**. (For readers who have had a course in modern abstract algebra.) Let $\mathcal{T}$ be the set of all Euclidean translations together with the identity. Prove that $\mathcal{T}$ is an Abelian (commutative) subgroup of the group of all Euclidean transformations.

## I. Hyperbolic Transformations

The classification of Euclidean transformations (as at the start of Section H) needs only one modification to become a classification of hyperbolic transformations. The rank 2 case is slightly more complicated, because we must account for both ultraparallel and asymptotically parallel subcases:

**Rank 2:** Given a rank 2 transformation $f = r_{\Lambda'} \circ r_{\Lambda}$,

- **(a)** if $\Lambda$ and $\Lambda'$ are not parallel then $f$ is a rotation about their point of intersection.
- **(b)** if $\Lambda$ and $\Lambda'$ are ultraparallel then $f$ is a translation between points on their common perpendicular line.
- **(c)** if $\Lambda$ and $\Lambda'$ are asymptotically parallel then ...

Then *what*? We've yet to investigate that case, but will do so now.

DEFINITION. Suppose $\Lambda$ and $\Lambda'$ are asymptotically parallel lines with orientations belonging to the ideal point $\mathcal{X}$. Then the transformation $f = r_{\Lambda'} \circ r_{\Lambda}$ is called an *ideal point rotation* about $\mathcal{X}$.

Ideal point rotations, like translations, have no fixed points, for Theorem 7.29 still applies. (However, in this case that theorem does not tell us if $f$ has preserved lines, for $\Lambda$ and $\Lambda'$ have no common perpendiculars according to Theorem 6.27.) Since they have no fixed points, it may seem odd to name these as any kind of rotation – to understand the reason for that name we'll need to understand something about the action of these transformations.

We'll look to a more familiar case for inspiration: when $\Lambda$ and $\Lambda'$ are not parallel and intersect at a point $A$ then $r_{\Lambda'} \circ r_{\Lambda}$ is a rotation about $A$. Both $r_{\Lambda}$ and $r_{\Lambda'}$ fix $A$ in this case, so their composition also fixes $A$. What is the corresponding behavior when $\Lambda$ and $\Lambda'$ are asymptotically parallel? The fixed points of $r_{\Lambda}$ are simply the points on $\Lambda$ and the fixed points of $r_{\Lambda'}$ are in a like manner the points of $\Lambda'$. So (since $\Lambda$ and $\Lambda'$ share no points in common) there is no point fixed by both reflections. There is, however, *something* that both $r_{\Lambda}$ and $r_{\Lambda'}$ leave fixed. We'll need some preliminaries in order to specify what it is.

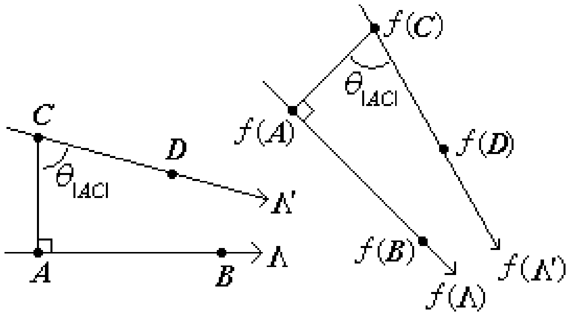



Figure 7.41: The conclusion of Lemma 7.40: a transformation carries asymptotically parallel lines to asymptotically parallel lines.

Let ${}_{(A,B)}\Lambda$ be an orientation of the line $\overleftrightarrow{AB} = \Lambda$. If $f$ is any transformation then we know that $f(\Lambda)$ will be the line $\overleftrightarrow{f(A)f(B)}$ (see Lemma 7.4) and we may give that line the orientation ${}_{(f(A),f(B))}f(\Lambda)$. Now suppose ${}_{(C,D)}\Lambda'$ is an oriented line that is asymptotically parallel to ${}_{(A,B)}\Lambda$ with $CA \perp \Lambda$ and $m\angle ACD = \theta_{|AC|}$ (see Figure 7.41). Then what is the relationship between ${}_{(f(A),f(B))}f(\Lambda)$ and ${}_{(f(C),f(D))}f(\Lambda')$? Since $f$ preserves distances and angles we have

- $f(C)f(A) \perp f(\Lambda)$,
- $|f(A)f(C)| = |AC|$, and
- $m\angle f(A)f(C)f(D) = m\angle ACD = \theta_{|AC|} = \theta_{|f(A)f(C)|}$.

So ${}_{(f(A),f(B))}f(\Lambda)$ and ${}_{(f(C),f(D))}f(\Lambda')$ are asymptotically parallel! We have just proved:

**LEMMA 7.40.** *If ${}_{(A,B)}\Lambda$ and ${}_{(C,D)}\Lambda'$ asymptotically parallel oriented lines and $f$ is any transformation, then ${}_{(f(A),f(B))}f(\Lambda)$ and ${}_{(f(C),f(D))}f(\Lambda')$ are also asymptotically parallel.* □

Now suppose $\mathcal{X}$ is the ideal point consisting of all oriented lines asymptotically parallel to ${}_{(A,B)}\Lambda$ while $\mathcal{X}'$ is the ideal point consisting of all oriented lines asymptotically parallel to ${}_{(f(A),f(B))}f(\Lambda)$. By this lemma we know that the image under $f$ of each member of $\mathcal{X}$ will be a member of $\mathcal{X}'$. We'll say that $\mathcal{X}'$ is the image of $\mathcal{X}$ under $f$, and write $f(\mathcal{X}) = \mathcal{X}'$. Note that if $f$ fixes $\Lambda$ then $\mathcal{X} = \mathcal{X}'$, so in this case we can say that $f$ fixes $\mathcal{X}$.

That is what is fixed by both $r_\Lambda$ and $r_{\Lambda'}$ when $\Lambda$ and $\Lambda'$ are asymptotically parallel: these two lines have orientations belonging to the same ideal point $\mathcal{X}$, and (by the above reasoning) both $r_\Lambda$ and $r_{\Lambda'}$ will fix $\mathcal{X}$. So, $f(\mathcal{X}) = r_{\Lambda'} \circ r_\Lambda(\mathcal{X}) =$

$r_{\Lambda'}(\mathcal{X}) = \mathcal{X}$ – that is, $f$ also fixes the ideal point $\mathcal{X}$. This explains the name *ideal point rotation*: $f = r_{\Lambda'} \circ r_\Lambda$ has no fixed points, but it does have a fixed ideal point!

**THEOREM 7.41.** *Let $f = r_{\Lambda'} \circ r_\Lambda$ be an ideal point rotation where ${}_{(A,B)}\Lambda$ and ${}_{(C,D)}\Lambda'$ are asymptotically parallel oriented lines. Let $\mathcal{X}$ be the ideal point containing both ${}_{(A,B)}\Lambda$ and ${}_{(C,D)}\Lambda'$. Then $f$ has no fixed points, no preserved lines, and $\mathcal{X}$ as its only fixed ideal point.*

*Proof:* We have already observed that $f$ has no fixed points (because of Theorem 7.29) and that $f$ fixes the ideal point $\mathcal{X}$. We need to show that $f$ has no preserved lines and no other fixed ideal points.

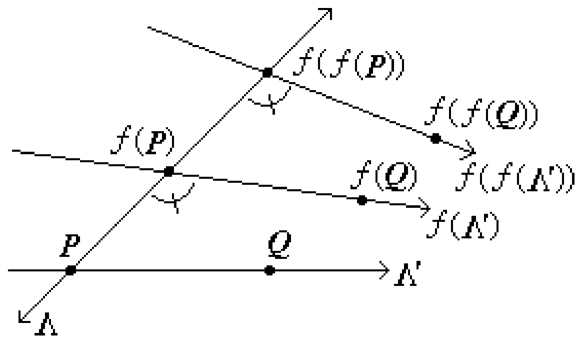



Figure 7.42:

Assume to reach a contradiction that $f$ has a preserved line $\Lambda$, and let $P$ be a point of $\Lambda$.

- Then $P$, $f(P)$, and $f(f(P))$ all lie on $\Lambda$.
- Furthermore, these three points are distinct since if two of them were equal then $f$ would have a fixed point (see Exercise 7.6).
- Since $|Pf(P)|$ must equal $|f(P)f(f(P))|$ it must be that $f(P)$ is between $P$ and $f(f(P))$.
- The ideal point $\mathcal{X}$ contains an oriented line through $P$ – call it ${}_{(P,Q)}\Lambda'$.
- Now $f$ fixes the ideal point $\mathcal{X}$, so ${}_{(f(P),f(Q))}f(\Lambda')$ and ${}_{(f(f(P)),f(f(Q)))}f(f(\Lambda'))$ also belong to $\mathcal{X}$.
- But then (since $f$ is a transformation) we have $m\angle Pf(P)f(Q) = m\angle f(P)f(f(P))f(f(Q))$ (see Figure 7.42).

- This means the transversal of $f(\Lambda')$ and $f(f(\Lambda'))$ by $\Lambda$ has congruent corresponding angles, an impossibility since $f(\Lambda')$ and $f(f(\Lambda))$ are asymptotically parallel (see Exercise 6.48).

This contradiction shows that $f$ cannot have a preserved line.

It remains to show that $f$ cannot have a fixed ideal point besides $\mathcal{X}$. So, assume to reach a contradiction that $\mathcal{X}'$ is a second fixed ideal point for $f$.

- Then by Theorem 6.22 there is a unique line $\Lambda = \overleftrightarrow{AB}$ such that ${}_{(A,B)}\Lambda \in \mathcal{X}$ and ${}_{(B,A)}\Lambda \in \mathcal{X}'$.

- Since $f(\mathcal{X}) = \mathcal{X}$ we must have ${}_{(f(A),f(B))}f(\Lambda) \in \mathcal{X}$.

- Similarly, since $f(\mathcal{X}') = \mathcal{X}'$ we must have ${}_{(f(B),f(A))}f(\Lambda) \in \mathcal{X}'$.

- Thus $f(\Lambda)$ has orientations in each of $\mathcal{X}$ and $\mathcal{X}'$, so by the uniqueness of $\Lambda$ (from Theorem 6.22) we must have $f(\Lambda) = \Lambda$.

- But this means $\Lambda$ is preserved by $f$, contrary to what we proved in the paragraph above.

This last contradiction shows that $f$ has no second fixed ideal point, so the proof of the theorem is complete. $\square$

We've already observed (from Lemma 7.40) that if a transformation $f$ fixes a line $\Lambda$ then it fixes the ideal points to which the orientations of $\Lambda$ belong. But actually more is true. For Lemma 7.40 implies that if $f$ preserves a line $\Lambda$ and also preserves its orientation (in the sense that $(f(A), f(B))$ gives the same orientation as $(A, B)$) then $f$ fixes both ideal points to which the orientations of the line belong. Thus, reflection across $\Lambda$ will fix both ideal points determined by $\Lambda$, and similarly translations and glide reflections (each of which have a preserved line) will fix two ideal points. The ideal point rotations are unique in fixing a single ideal point. Summing up, we have the following table for hyperbolic transformations.

| Table of Hyperbolic Transformations | | |
|---|---|---|
| **Rank** | **Fixes ...** | **Type** |
| 0 | entire plane | identity |
| 1 | one line, two ideal points | reflection |
| 2 | one point | rotation |
| | two ideal points | translation |
| | one ideal point | ideal point rotation |
| 3 | two ideal points | glide reflection |

## Exercises

**7.61**. Let $f$ be a rotation by angle measure 180 about a point $P$. Does $f$ preserve any ideal points? Explain and justify your answer.

**7.62**. We'll say that a transformation $f$ preserves the oriented line ${}_{(A,B)}\Lambda$ if ${}_{(f(A),f(B))}f(\Lambda) = {}_{(A,B)}\Lambda$. Prove that $f$ preserves an oriented line if and only if it fixes two distinct ideal points.

**7.63**. Prove that the inverse of a translation is always a translation, the inverse of a rotation is always a rotation, and the inverse of an ideal point rotation is always an ideal point rotation.

**7.64**. In Euclidean geometry the composition of translations is always a translation (see Exercise 7.60). Show that the composition of three hyperbolic translations may be a rotation by completing the steps below.

- Let $ABC$ be a triangle with $m\angle A = \alpha$, $m\angle B = \beta$, and $m\angle C = \gamma$.
- Let $f_1$ be the translation from $C$ to $A$, $f_2$ the translation from $A$ to $B$, and $f_3$ the translation from $B$ to $C$.
- Clearly $f_3 \circ f_2 \circ f_1$ fixes $C$, so is either the identity or a rotation about $C$.
- Show that $f_3 \circ f_2 \circ f_1(A) \neq A$ and that $f_3 \circ f_2 \circ f_1$ is a rotation about $C$ by the defect of triangle $ABC$.

# Index