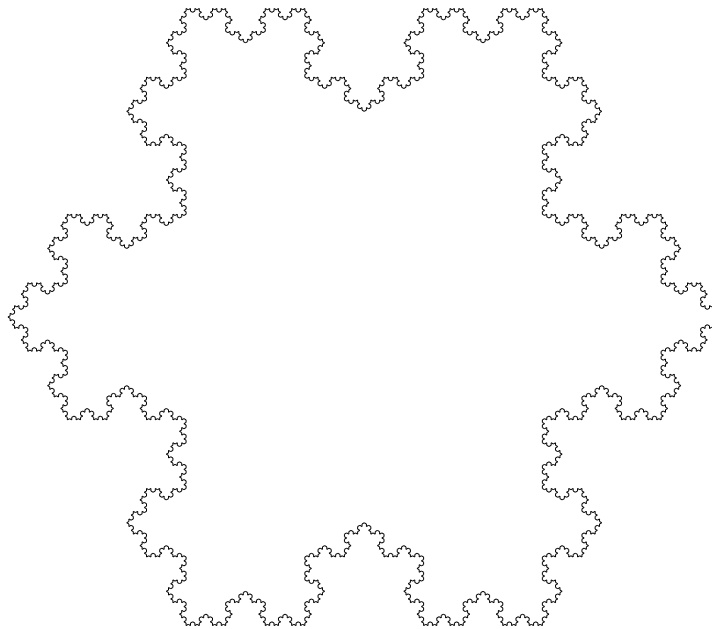# Geometry with Computers

Computer-Based Techniques to
Learn and Teach
Euclidean Geometry

Tom Davis

Draft Date: May 17, 2006

Chapter 0

Preface

> Mathematics must be written into the mind, not read into it. "No head
> for mathematics" nearly always means "Will not use a pencil."
> *Arthur Latham Baker*

0.1 Why Another Geometry Book?

Euclidean Geometry is ancient, and thousands of books and articles have been written on the subject. Why write another?

Here are the differences between this book and the others:

- Most mathematics books are written to communicate results: theorems, calculation methods and algorithms. This book concentrates on computer-based techniques for solving geometric problems.

- Although it may not be obvious, mathematics is highly geometrical—virtually every formula has an associated picture. Often it is much easier to obtain understanding by mentally manipulating the picture than by manipulating a formula algebraically. What better way to practice mental manipulation of pictures than to study geometry?

  Mental manipulation of geometry is difficult for many people, but today almost any relatively new personal computer has graphical capabilities that would have been unimaginably powerful 20 years ago. These machines can help visualize geometric results in a dynamic manner that is far more compelling than the fixed drawings appearing in standard mathematics books or that can be produced with pencil and paper. This book comes with a computer program called **Geometer** that allows you to manipulate existing geometric diagrams and to create your own.

- **Geometer** can also be used as an experimental tool to run geometry "experiments" that can aid greatly in both understanding and proving results.

- Although Euclidean geometry is a huge field, it is not a big part of the general mathematics curriculum in the United States. I have a Ph.D. in mathematics, and the only formal class I had in Euclidean geometry was in high school. Your geometry class may have been different, but in mine we somehow managed to avoid learning most of the truly beautiful results.

- High school geometry is often the first introduction students have to constructing mathematical proofs. It is obviously better to begin with simple examples when teaching a student to write proofs, so the proof construction exercises in typical high school texts can almost all be completed in three or four steps.

  Unfortunately, if one is limited to proofs of only a few steps, a huge proportion of geometry is inaccessible. For people who can follow a ten or twenty step proof, it is amazing what results are accessible—some of the most beautiful theorems in mathematics.

  One's technique of geometric proof can be honed with artificial drills, or one can work on interesting problems. This is similar to the difference between running on a treadmill or on mountain trails—the treadmill is completely predictable while the trails can be tricky but the trail runner will not be bored out of his mind. This book is aimed at those who did enough treadmill geometry in high school and want to see some beauty while they learn advanced techniques.

  The material here is accessible to brighter high school students, and can be used as supplementary material by teachers or by the students themselves.

  The **Geometer** program can also be used as a presentation device in geometry classes. Although most of the material is beyond what is taught in high school, many of the more important introductory results are available on the enclosed CD in **Geometer** format.

  There is additional material on the CD to help—some more elementary theorems, and a tutorial on the construction of sophisticated **Geometer** diagrams, for example. See the index file on the CD for up to date information.

- Finally, with computer graphics increasing in power every year, Euclidean geometry (and projective geometry as well) may be destined for a comeback. More and more computer applications are being designed to help people visualize problems geometrically, but to make the computers do a good job of that, more knowledge of geometry is required.

0.2 Audience

This book is for teachers of Euclidean geometry but it contains topics of interest for motivated students or anyone who loves geometry. Students and coaches of students who compete in high level mathematics competitions should find the advanced material in this text valuable as training aids.

0.3 Origins of Geometer

The I worked at Silicon Graphics for 16 years doing graphics programming and helping design computer graphics hardware and so I have a fairly solid grounding in both the theoretical and practical details of geometry.

Although I was paid money to be a computer engineer, I have never ceased to be a mathematician as well. From time to time both in and out of work I ran into interesting geometric theorems. A couple of them seemed so surprising and non-intuitive that I wrote computer programs to help me visualize them. After going through that exercise a few times, I realized that it would not be hard to write a general-purpose program where arbitrary geometric diagrams could be entered with a graphical user interface (GUI) and dynamically altered. I hacked together the original version of **Geometer** on Silicon Graphics UNIX machines in my spare time. (I was wrong when I thought it "would not be hard.")

The version of **Geometer** included with this book is vastly modified and improved (and it now runs on Windows, Macintosh OS X and Linux machines), but much of the underlying philosophy is the same.

As I worked on this book, I tried to practice what I preach. I knew of many interesting results that I had never actually proved, and rather than just look up the standard proofs, I struggled to figure them out myself. Of course from time to time I had to look in various books for "hints" and even when I did solve a problem completely without hints, I later read the literature to see what the "standard" methods are.

In my struggles to find the proofs, I used **Geometer** extensively in exactly the same ways I expect you the reader to use it. It may not be bug-free, but I can assure you that it is quite robust, and since it has been used to create thousands of real dynamic diagrams, it is fairly well streamlined: when I got annoyed with something, I just fixed it.

0.4 Acknowledgements

> A mathematician is a machine for turning coffee into theorems. [1]
> *Alfréd Rényi*

Zvezdelina Stankova showed me many wonderful geometric ideas. In particular, I learned from her the power of the technique of inversion in a circle.

I would like to thank Joshua Zucker and his students in Palo Alto's Gunn High School Math Circle for all the stimulating discussions of mathematics problems and for serving as unwitting guinea pigs for some of the presentation techniques used here. I also learned a lot in Zvezdelina's UC Berkeley Math Circle, and in Tatiana Shubin's Math Circle at San Jose State University.

Special thanks go to Gerald Alexanderson and Arthur Benjamin for their help in converting what began as a disorganized collection of half-baked ideas into the completely-baked version you hold in your hands. Rudolf Fritsch provided many detailed corrections and suggestions.

I received many wonderful suggestions for improvements and enhancements to the **Geometer** program from Paul Haeberli and Henry Moreton.

This book was produced using LaTeX, the Emacs editor, and the PostScript files produced by **Geometer**. After an infuriating attempt to use the most popular word processor in the world for the text, Emacs and LaTeX were a total pleasure to use, so I must thank Donald Knuth for writing TeX and Leslie Lamport for writing LaTeX.

My wife, Ellyn Bush, although not a mathematician, read and re-read many parts of the text, and provided many valuable insights on how to make things simpler and clearer. She put up with all of my complaints with only a few of her own.

Finally, as the quotation at the beginning of this section indicates, I would like to thank whoever it was who discovered that beans of the shrub *Coffea arabica* could be roasted, ground and steeped in boiling water to produce the nectar of life. I converted that nectar not only into theorems, but also into text and into the computer code that produced **Geometer**.

Tom Davis
May, 2004

---

[1]This result was sharpened by Pal Turan who added that weak coffee produces only lemmas.

Contents

Chapter 1

Introduction

This book describes techniques and strategies that use computer software to enhance the study of geometry. If you are itching to try the **Geometer** program on the enclosed CD, that is as good a way to start as any. Install the program and go directly to the tutorial which you will find under the *Help* pulldown menu.

The **Geometer** program upon which this book is based is freely available and in the public domain. Anyone may obtain a copy at:

        http://www.geometer.org/geometer

You may already be proficient in the use of one of the commercial computer geometry programs such as Geometer's Sketchpad®, Cabri Geometry®, Cinderella® or others, and if so, almost all the techniques described here can be applied using those packages. The main advantage of **Geometer** is that the included CD contains computer-readable versions of all the examples and figures in this book in **Geometer** format.

This book is for anyone who wants to learn about computer methods to enhance the study of geometry, be they students or teachers of the subject. The term "student" refers to anyone who is studying geometry—either as part of a formal course or simply for the love of the subject. The text does assume a high school mathematics background, but Appendix A reviews all the important prerequisites.

The purpose of this book is to demonstrate techniques, not theorems. Although as far as possible interesting theorems have been chosen for the demonstrations, no attempt has been made to assure that the list of theorems presented is complete in any sense. While it is true that the text material combined with the additional examples on the CD does cover a large number of theorems, please do not feel slighted if your favorite theorem does not appear.

1.1   Computer Geometry Programs

A computer geometry program is a computer program with a graphical user interface that allows you to draw geometric objects and to adjust them dynamically. Such systems are constraint-based in that most of the objects in a figure are not placed absolutely, but are defined in terms of the positions of other objects on the screen. For example, your drawing may contain two freely movable points and a third point that is their midpoint. In most computer geometry programs you can freely adjust the positions of the first two points, but not the third. If you do adjust either of the first points, however, the position of third point will change so that it remains at their midpoint.

The constraints in such a system can be nested arbitrarily deeply. For example, a line passing through the midpoint described in the previous paragraph will also move when the midpoint moves, but the midpoint can only move if one of the two original points is moved, and so on.

This book demonstrates many things that computer geometry programs can do, but it is equally important to be aware of what they cannot do.

Such programs do provide accurate drawings. They allow you to visualize not just one example of a drawing, but thousands of variations. They allow you to test geometric conjectures, and can present dynamic illustrations of relationships or theorems.

What they cannot do (at least as of the publication date) is to generate proofs of theorems. In order for a theorem to be accepted as true, a formal proof is required, and although computer geometry programs can help you to find that proof, it is you who will have to provide the completed version.

The more geometry you know, the more helpful a computer can be. If you have drawn a diagram and see in it a relationship that you would like to prove, it may be that the computer can discover other relationships in the diagram that you did not notice and that may help you construct a proof. But it will be up to you to discover how to use that information to produce a formal proof.

For example, imagine that you are trying to show that three circles in a diagram meet at a point. When you allow the computer to examine the diagram, suppose it notices that a certain set of four seemingly unrelated points are concyclic ("concyclic"

means that they all lie on the same circle). This unlikely occurrence is easy for a computer to recognize, but usually difficult for a human, and it could easily be related to your three circle concurrence problem. But even if knowing that the points are concyclic makes your result obvious, you still need to show that the four points are, in fact, concyclic. It is clear that the more ways you know to prove that four points lie on the same circle, the easier your task will be. (If you're interested in what such a list might look like, see item 15 in Section 1.3.)

Another obvious limitation of computers is due to round-off error in their calculations. It should not surprise you if a computer indicates that a certain right triangle has sides of lengths 3, 4, and 5.0000001, although any student of high school geometry knows that this cannot be true. Thus some common sense is required. If you are asked to find an unknown length, and the computer measures it as 7.000013, you will probably have a lot more luck if you try to prove that it is 7, not 7.000013. If, however, the computer says the length is 7.071907, it is probably *not* equal to 7. If you are very good, you may notice how close 7.071907 is to $\sqrt{50}$—knowledge is power.

1.2 The Contents of the CD

The enclosed CD contains a copy of the public-domain computer geometry program called "**Geometer**" that runs on Windows and Linux machines as well as on Macintosh systems running OS X. All the illustrations in this text were prepared with **Geometer**. See Section 1.5 for details about how to interpret the captions on the figures. The **Geometer** CD includes:

- The **Geometer** program itself.

- A tutorial and user's guide for the program.

- The source files for all the examples in the user's guide and tutorial.

- The source files for all the illustrations in this book.

- Solutions to all the exercises in this book in the form of **Geometer** files (called **Geometer** diagrams).

- Many additional files of interesting geometric results not covered in this book.

The **Geometer** program can be used at many levels, depending on how much you want to learn about it. See Chapter 2 for a much more detailed description of what **Geometer** (and programs like it) can do.

- To view a diagram in this text simply double-click the corresponding **Geometer** file on the CD.

- To manipulate the figure, place the cursor over a point and press down the mouse button. Then until you release the button, the point will be dragged with the cursor, and the figure will be modified appropriately.

- Some **Geometer** diagrams are "proofs". The author of the proof organized the diagram so that you can step forward and back through the proof using the *Start*, *Next* and *Prev* (previous) buttons. At each stage during the proof, the figure can be manipulated as described above.

- Still other **Geometer** diagrams are scripts, in which case the *Script* button will be solidly drawn. If you click on this button with the mouse, the diagram will pass through a prearranged script. Sometimes you can manipulate the figure before pressing the *Script* button to observe the consequences of the script with different initial configurations.

- Using the mouse buttons and menus you can construct and save your own simple **Geometer** diagrams to test conjectures and to search for patterns and geometric relationships.

- In addition to manipulation with the mouse using a graphical user interface, **Geometer** diagrams have a textual form that can be edited either within **Geometer** or using your favorite text editor. Using this technique, all the features are available, including designing your own scripts or proofs.

1.3   Organization of the Book

Here is the basic organization of the rest of the book:

1. **Introduction**

   What you are reading now.

2. **Computer-Assisted Geometry**

   This chapter consists of a quick survey (with examples) of almost every possible use for a computer geometry program.

3. **Geometric Constructions**

   To be able to use a computer geometry program, you have to know how to construct diagrams that correspond to the problems you are investigating. This chapter describes with many examples methods to construct computer diagrams. Classic straightedge and compass constructions are described and discussed, but most of the emphasis is on how to use the more powerful techniques available in computer geometry programs, **Geometer** in particular.

   At the end of the chapter is a list of construction exercises whose solutions appear in **Geometer** files on the CD.

4. **Computer-Aided Proof**

   This chapter describes how to use your computer to help you search for a proof of a geometric theorem.

5. **More Useful Theorems**

   Here are a few useful theorems and their proofs that are not typically covered in a high school course. When possible, **Geometer** will be used as an aid to find those proofs. The results in this chapter plus those in the review appendix provide all the classical tools needed to solve the rest of the problems in the book.

6. **Locus of Points**

   Computer geometry programs make the search for "locus of a point" problems almost trivial. In this chapter we will examine some nice examples.

7. **Triangle Centers**

   The centroid, circumcenter, incenter and orthocenter are four well-known classical triangle centers, but there are hundreds of others. This chapter examines some of them.

8. **Inversion in a Circle**

   The technique of inversion in a circle is described, with and without computer geometry programs. Applications of this technique to geometric construction are addressed.

9. **Projective Geometry**

   Geometer and probably most other computer geometry programs do many of their calculations using projective geometry rather than Euclidean geometry. In addition, much of computer graphics is based upon projective geometry. This chapter presents the fundamental concepts of projective geometry, but from a Euclidean point of view. Calculations in homogeneous coordinates are described.

10. **Harmonic Point Sets**

    This is a continuation of the previous chapter on projective geometry, but again, from a Euclidean point of view. In many geometric configurations, sets of harmonic points appear, and if that is the case, many methods can be applied to help solve problems. **Geometer** is very good at finding harmonic point sets.

11. **Geometric Presentations**

    This chapter is primarily for teachers of geometry or for anyone who would like to present geometric results to others using **Geometer**. Various techniques are discussed that will improve the quality of **Geometer** diagrams making them easier to understand and more beautiful. Techniques for the presentation of proofs, constructions, and animated scripts are discussed with examples.

12. **Geometer Proofs**

    Even if **Geometer** is not used to find a proof, it can be used to present a proof step by step in a very intuitive way. It requires some effort to construct such proofs, but if they are already constructed, such proofs are easy to use, both by a student or a teacher. At the beginning of the chapter is a list of problems with diagrams; at the end is a list of problems described with text only. Complete solutions for all the problems appear as **Geometer** diagrams on the CD.

13. **Appendix: Mathematics Review**

   Here we review parts of high school mathematics. Most of the topics are from
   geometry, but some trigonometry and analytic geometry are included. The im-
   portant results are listed, where "important" means that the result is useful for
   deriving other results. The results are standard, and most can be found in high
   school mathematics texts.

14. **Appendix: Geometer Art** At the beginning of each chapter is a **Geometer** il-
   lustration that is at least partly artistic. This appendix describes how each one is
   made, and perhaps says something about the mathematics illustrated.

15. **Appendix: Geometric Problem Solving Strategies**

   This appendix categorizes many of the standard methods to show properties of
   geometric figures.
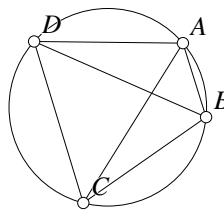
Figure 1.1: When do four points lie on a circle?
`Intro/FourCircPts.D [D]`

For example, to continue with the example in Section 1.1, how *does* one show
that four points lie on a circle? Here are some approaches if the points are,
in clockwise or counterclockwise order, A, B, C, and D (see figure 1.1). (Do
not worry if these techniques seem mysterious. They will be covered in Ap-
pendix A.)

   - Show that all four points are the same distance from some fixed point
     (which would be the center of the circle).
   - Show that $\angle DAB = \angle BCD = 90°$ (or that $\angle ABC = \angle CDA = 90°$).
   - More generally, show that $\angle DAB + \angle BCD = 180°$ (or that $\angle ABC + \angle CDA = 180°$).
   - Show that $\angle ABD = \angle ACD$ (or that $\angle BCA = \angle BDA$ or $\angle CDB = \angle CAB$ or $\angle DAC = \angle DBC$).
   - Show that all four points lie on some well-known circle, such as the nine
     point circle.
   - (Ptolemy's Theorem) Show that $AB \cdot CD + BC \cdot DA = AC \cdot BD$.

The entire appendix is organized a bit like a thesaurus. You look up the sort
of problem you are attempting to solve and you will find a list of possible ap-
proaches to that kind of problem.

1.4 Proofs in this Book

> We have a habit in writing articles published in scientific journals to make the work as finished as possible, to cover up all the tracks, to not worry about the blind alleys or describe how you had the wrong idea first, and so on. So there is not any place to publish, in a dignified manner, what you actually did in order to get to do the work.
> *Richard P. Feynman*

> I have had my results for a long time: but I do not yet know how I am to arrive at them.
> *Karl Friedrich Gauss*

Most geometry texts present slick and polished proofs of each theorem—usually the best one that the author knows. In virtually every case, the first discovered proof of a new theorem is *not* "slick and polished". The slick and polished proofs arise as mathematician after mathematician improves upon what has been done before.

There are different kinds of proofs. Some are seemingly miraculous, and it is almost impossible to imagine how they were discovered. Some are beautiful in a different way in that they make it completely obvious why a certain fact or theorem is true. Some are mechanical, and relatively easy to construct, although they may not be brilliant and may not shed too much light on the problem.

One of this book's goals is to show how to use computers to help solve problems. In geometry this often means you are trying to demonstrate some property that you have never proved before. Thus the proofs here will tend to be of a rougher form, more mechanical and less magical than what you will find in geometry textbooks. The advantage to this is that it is often much easier to see how to arrive at that proof.

Usually we will not concern ourselves with the process of polishing the proof to make something beautiful, but from time to time there is such a wonderful and magical proof available that it would be aesthetically criminal to omit it.

1.5 Illustrations in this Book

> Uninvited advice is usually ignored; yet I wish to offer some in the hope that it will be helpful. No book on mathematics can have enough illustrations or formulas. The thorough reader must always work with pencil and paper at hand.
> *Nicholas D. Kazarinoff*

Although Kazarinoff's claim from [Kazarinoff, 1961] is true, illustrations take a lot of space so many books have fewer than they should. The author remembers how depressed he was once when he opened a book whose title claimed that it presented the geometric viewpoint of complex variables, and there were almost no illustrations.

But with a computer that has decent graphics and a suitable geometry program, it is possible to present not just one illustration of a geometric fact, but thousands or even millions of variations of that illustration. The final sentence of Kazarinoff's quote should be amended to "work with pencil, paper, and *computer* at hand." (By the way, his book ([Kazarinoff, 1961]) is pretty good too.)

This book contains quite a few illustrations, and every one of them is also available as a **Geometer** file that can be viewed and modified dynamically. In addition to being simply a displayer of these geometric diagrams, however, **Geometer** allows both the modification of existing diagrams and the creation of new ones.

Following each figure's caption is a line that tells you where to find the **Geometer** diagram used to produce the figure, including the directory and file names. This name is relative to where you installed **Geometer**. All **Geometer** file names have a suffix of ".T" or ".D". There is no difference between the two file types from **Geometer**'s point of view, but if the figure cannot be reasonably manipulated and is only used as a drawing, then the .D suffix is used. Think of the ".D" as standing for "drawing" and of ".T" as standing for "theorem".

Following the file name in brackets is a single letter that indicates what sort of a diagram it is. There are four possibilities:

D "Drawing". These are the least interesting and usually have the .D suffix. A simple drawing was needed for the text and although some parts may be movable, if you move them, the figure may no longer illustrate what it was supposed to. But even these figures can be interesting. If you are wondering how on earth some of the diagrams were drawn with **Geometer**, you can simply load the diagram and use the *Edit Geometry* command to see. (Or examine the .D file in your favorite text editor.)

M "Manipulation". The drawing can be manipulated by moving various points. Usually it is obvious from the text what parts should be moved, but if not, the diagram itself will usually contain some clues in the form of text labels. In a tiny percentage of these manipulation diagrams there are some extra labels that are not tied to the geometry correctly but are placed where they are to make a diagram suitable for publication. The drawing changes correctly, but sometimes the labels are left behind.

P "Proof". When you see this sort of diagram, it is almost always a good idea to load it into **Geometer**. It is either a proof or construction where you are led, step by step, to the solution. To advance to the next stage, use the *Next* button in the control area of the screen, or simply type the letter "n" on your keyboard. To go back, use the *Prev* button or type "p" on your keyboard. To restart the proof or construction, use the *Start* button. In almost every proof, you can also manipulate the figure, both before or during the proof.

S "Script". A script is a **Geometer** file that includes a scripted transformation of the drawing. These diagrams are also worth loading into **Geometer**. The built-in script can be accessed by pressing the *Run Script* button. To interrupt the

script, simply click the mouse button with the cursor anywhere in the **Geometer** window. A few of the scripts do not do too much—they were simply used to guarantee evenly-spaced points for diagrams for publication, but the majority are very useful to watch. In most cases, you can also manipulate the figure before running the script. For example, a script that draws a conic section given information about certain lengths and foci can be run, then the lengths and/or foci changed, and when the script is run again you can see how the modifications affected the result.

If a picture is worth a thousand words, then are a thousand pictures worth a million words? Even if not, a thousand pictures is surely better than one. In addition, the fact that on a computer screen you can watch each of the thousand images blend into the next tends to provide an even better idea of what is going on.

Another less obvious advantage of using **Geometer** is that since every illustration in this book is also a **Geometer** file on the CD, you can view and manipulate that image on your computer screen as you read about it and it is easy to continue viewing the diagram when you turn the page.

In a few cases, there are two or more versions of the **Geometer** file on the CD, since a slightly different version was used to generate the figure in the text. A common situation is when the basic figure appears on the left and another version with some additional constructions appears on the right. The file whose name appears with the illustration is usually neither—it is usually the basic figure plus the constructed lines.

Finally, on the first page of each chapter there appears an illustration that is more artistic than the usual figure in a Euclidean geometry text. Appendix B provides a short description of each, and the name of the **Geometer** file and (if appropriate) the name of the "C" file that generated that **Geometer** file.

1.6 How to Use this Book

One of the reasons Americans on average do not do well in mathematics is that nowadays we are conditioned by television and our society in general to expect instant gratification. Unfortunately, whether you like it or not, the best way to learn anything about mathematics is to think hard about it and to struggle with problems. If you simply read a problem and then immediately read its solution, you will not get too much out of it. *The more you struggle with a problem, even if you do not solve it, the more you will learn.* If you do solve it by yourself, great! If you needed a hint or two, that is great too. But even if you get totally stumped, at least you will read the solution with much more interest.

Much of this book is organized as as follows: A problem is presented followed by a section that discusses how **Geometer** might be used to search for approaches to it. Finally, a complete solution to the problem is presented. You will learn the most, of course, if you read the problem and try to use **Geometer** yourself to try to solve it. See what you can learn. Only then should you look ahead at the solution.

Chapter 3 ends with a long set of construction problems. Each has a complete solution in a **Geometer** file on the CD, but of course there is no discussion of how that construction was imagined. Similarly, Chapter 12 consists entirely of problems, each of which has a corresponding proof in a **Geometer** file.

As you work on each problem, try not to get stuck in a rut. Draw pictures (even better, draw pictures with **Geometer**). Load the associated **Geometer** file, if there is one. Look at the solution in the **Geometer** file only after you have worked on the problem yourself, whether you solved it or not. Even if you did solve the problem, check the **Geometer** solution—it may be different from yours, and may shed additional light on the problem.

1.7 Notation

> A foolish consistency is the hobgoblin of little minds, adored by little statesmen and philosophers and divines. With consistency a great soul has simply nothing to do.
> *Ralph Waldo Emerson*

> ♦ For fools rush in where angels fear to tread.
> *Alexander Pope*

> ♦♦ For mathematicians rush in where fools fear to tread.
> *Tom Davis*

Some topics in this text are more difficult than others. Even within a topic, some paragraphs are difficult and can be skipped on the first reading. Others are *very* difficult, and only experts (or angels or fools) should wander into them on the first pass.

Difficult and very difficult paragraphs are indicated with black diamonds that look like this: ♦ and this: ♦♦. If you are a skier, you will recall that the most difficult slopes are often marked "double black diamond". Sometimes entire sections are marked with a single or double black diamond in which case everything in the section has the indicated difficulty. You may even encounter a "triple black diamond": ♦♦♦. If an entire section is marked with a warning and if some subsection or paragraph within it is also marked with a warning, it means that the subsection is relatively more difficult than the section within which it lies.

This book's mathematical notation is fairly standard, but following Emerson, it is not 100% consistent.

We use the symbol "≅" to indicate congruence between two equal geometric figures. The symbol "=" is for equality of lengths, angle measure, area, et cetera. The symbol "∼" indicates two similar figures.

Generally, points or vertices are indicated with upper-case letters, as: $A, B, \ldots Z$. When they have a name, lines or line segments are indicated with lower-case letters:

$a, b, \ldots, z$. A segment or line can be indicated by two points on the line, as: "The line AB".

$\triangle ABC$ is the triangle whose vertices are the points A, B, and C. For more complex polygons, we will normally just list the vertices and include a word to say what kind of polygon we mean, such as: "quadrilateral $ABCD$" or "hexagon $ABCDEF$".

The same notation is used for a segment as for the length of the segment, and that should not cause any confusion. If it appears in some equation, such as $AB : CD = \sqrt{2}$ or $AB + BC \geq AC$, we are obviously talking about the lengths of the segments.

Ratios of numbers, lengths or areas are often written as follows: $a : b = c : d$. This is almost equivalent to $a/b = c/d$, but remains true if $b = d = 0$, as long as a and c are not zero.

If angles have a name, it will usually be a Greek letter, like "$\alpha, \beta, \gamma, \ldots$". If there is no name, the notation: "$\angle ABC$" means the angle traced out by drawing a line from point A to point B and then on to point C. Sometimes the name of the vertex may be used as the angle name if there is no possibility for confusion. For example, in $\triangle ABC$, if there are no other lines passing through vertex A, then "$\angle A$" is the same as "$\angle BAC$".

If two lines or segments AB and CD are parallel we write $AB \parallel CD$. If those same two segments were perpendicular, we would write $AB \perp CD$.

If two points A and B lie on a circle, the symbol $\overarc{AB}$ refers to the arc from A to B. If there is any chance of confusion, the arc is the part of the circle beginning at A and going to B in a counterclockwise direction, so the arcs $\overarc{AB}$ and $\overarc{BA}$ together make up the entire circle.

We will use the symbol $\mathcal{A}$ to indicate the area of a polygon. $\mathcal{A}(\triangle ABC)$ stands for the area of $\triangle ABC$.

A few sets are important. $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$ represents the set of natural numbers. Some people include zero in the set of natural numbers and some do not. In this text, we will include it. $\mathbb{Z} = \{\ldots, -4, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ is the set of integers. $\mathbb{Q}$ is the set of rational numbers, $\mathbb{R}$ is the set of real numbers, and $\mathbb{C}$ is the set of complex numbers.

1.8   Where to Go from Here

The main reason to read this text is to learn some computer techniques that are useful in Euclidean geometry, but in the process of learning those techniques you will be exposed to quite a few interesting and beautiful theorems that do not appear in high school textbooks. If you want to learn more geometry or have more problems to try, there are plenty of other sources available.

One computer technique not mentioned up to now is simply the availability of a great deal of information on the internet. Here is a short list of some sites that deal with various aspects of geometry. Unlike standard bibliographic references, it is impossible to guarantee that the addresses below are still valid when you read this book.

- `http://www.geometer.org/geometer/`
  The Geometer home page. Here you can obtain the latest version of the **Geometer** program. There is additional material here of various sorts, including additional **Geometer** diagrams.

- `http://freeabel.geom.umn.edu/docs/forum/`
  The Geometry Forum. This is an electronic community focused on geometry and math education based at Swarthmore college.

- `http://www.ics.uci.edu/~eppstein/junkyard/`
  The Geometry Junkyard. A collection of usenet clippings, web pointers, lecture notes, research excerpts, papers, abstracts, programs, problems, and other stuff related to discrete and computational geometry.

- `http://www.geom.umn.edu/`
  The Geometry Center. Although the Geometry Center at the University of Minnesota is now closed, its website is still available and filled with interesting items.

- `http://dmoz.org/Science/Math/Geometry/`
  Open Directory Project Science: Math: Geometry. An index to many more interesting web pages related to geometry.

- `http://www.cut-the-knot.org/ctk/index.shtml`
  Cut The Knot! An internet column by Alex Bogomolny with many wonderful articles mostly about geometry, and many with associated animations.

- `http://faculty.evansville.edu/ck6/tcenters`
  Encyclopedia of Triangle Centers. Clark Kimberling's collection of over 1000 triangle centers. See Chapter 7.

Finally, here is a list of some books that may be interesting. See the bibliography for complete citations.

- *Geometry Revisited* ([Coxeter and Greitzer, 1967]) is one of the best books on general advanced Euclidean geometry. The information is quite dense, and it is amazing how much is packed into one small book.

- *College Geometry* ([Altshiller-Court, 1952]) is an old classic textbook that is also loaded with information. It contains an extensive set of problems and can often be found in its paperback version in used bookstores.

- *Advanced Euclidean Geometry* ([Johnson, 1929]) is another classic. It is a bit more difficult to find than the volume above, but it is worth the search. It was also more recently (1960) printed in a Dover Publications edition.

- *A Survey of Geometry* ([Eves, 1965]) is yet a third classic worth checking.

- *Geometry Turned On!* ([King and Schattschneider, 1997]) contains a collection of papers on the use of computer geometry programs both in and out of the classroom.

- *Geometric Inequalities* ([Kazarinoff, 1961]) is a good place to learn about inequalities. The topic of geometric inequalities is almost completely ignored in the book you are holding.

- *Geometric Transformations I, II,* and *III* These three books by Irving Yaglom ([Yaglom, 1962a], [Yaglom, 1962b], [Yaglom, 1962c]) approach geometry from the point of view of transformations. This is another topic barely touched upon in the current text.

- Dan Pedoe has written a couple of interesting books: *Circles: A Mathematical View* ([Pedoe, 1995]), as advertised, tells a lot about circles. Another of his books, *Geometry, A Comprehensive Course* ([Pedoe, 1988]) describes in great detail vector methods to solve geometric problems. This second text is a Dover republication of his *A Course of Geometry for Colleges and Universities* published in 1970 by Cambridge University Press.

- Berger's texts: ([Berger, 1987a], [Berger, 1987b] and [Berger et al., 1984]), including the problem book have their good and bad points. On one hand, they are wonderful to look through—you will get hundreds of ideas just by looking at the illustrations. On the other hand, this author finds the proofs and explanations *very* difficult to follow—Berger seems to think about geometry in a different way. If you have trouble following this book, perhaps you think like Berger, and his books will make more sense to you.

- Two books called *Proofs Without Words* ([Nelson, 1993]) and *Proofs Without Words II* ([Nelson, 2000]) provide interesting geometric methods to visualize proofs from many areas of mathematics, geometry included.

- *Challenging Problems in Geometry* ([Posamentier and Salkind, 1988]) contains a set of problems that are nice and not too difficult. A more challenging set can be found in the Russian book *Problems in Plane Geometry* ([Sharygin, 1988]). Both contain a section of problems, another section of hints, and then a section of solutions.

- Two fun books include *Excursions in Geometry* ([Ogilvy, 1969]) and *The Penguin Dictionary of Curious and Interesting Geometry* ([Wells, 1991]). Both contain lots of topics, many of which are a bit off the beaten path.

- For bigger view of geometry—not just Euclidean, but the whole gamut—take a look at Coxeter's *Introduction to Geometry, Second Edition* ([Coxeter, 1989]) and Hilbert's *Geometry and the Imagination* ([Hilbert and Cohn-Vossen, 1983]).

- A more rigorous approach to geometry can be found in *Euclidean and Non-Euclidean Geometries, Third Edition* ([Greenberg, 1993]), and in *Companion to Euclid* ([Hartshorne, 1997]), or *Elementary Geometry from an Advanced Standpoint* ([Moise, 1963]).

- The *CRC Concise Encyclopedia of Mathematics* ([Weisstein, 1999]) contains an amazing collection of mathematical information (including much geometry).

- Although there is probably no hope of finding these particular books, ancient high school and college geometry texts (leather-bound, even) can often be found in used bookstores. They often present a very different view of the subject from what is usually seen in modern texts. Two examples are *New Plane and Solid Geometry* ([Beman and Smith, 1900]) and *Geometrical Problems Deducible from the First Six Books of Euclid* ([Bland, 1819]), but there are plenty of others. Note that there is no error in this last citation; Bland's book was really published in 1819, not 1919.
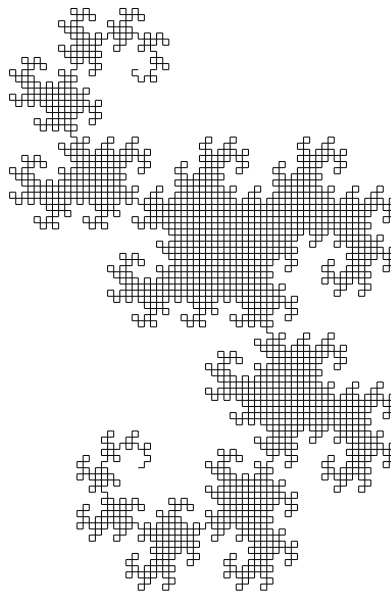
Chapter 2

Computer-Assisted Geometry

There are many computer programs that aid in the visualization of Euclidean geometry. Some available commercial ones at the time of this writing include Geometer's Sketchpad®, Cabri Geometry®, and Cinderella®.

The CD accompanying this book contains yet another called simply "**Geometer**" that provides many of the features of the programs above, but is in the public domain. **Geometer** runs on PCs, Macintoshes running OS X and Linux machines. All the examples in this book were created with **Geometer**, but most of them could have been drawn equally well with any of the other commercial programs. Similarly, every illustration is a **Geometer** diagram, so if you read this with a computer at your side so you can manipulate the **Geometer** diagram at the same time that you read about it. (The drawings of geometric figures produced by **Geometer** are called "diagrams".)

To understand this chapter you need to have some idea of how **Geometer** or some other computer geometry program works. If you use **Geometer**, the CD contains complete documentation including a tutorial. If you have never used a computer geometry

15

program before, it is probably worthwhile to work your way through the **Geometer** tutorial before reading too much farther.

Finally, there is a chapter in the **Geometer** reference manual on the CD that contains another list of suggestions for effectively using the program to study geometry and to learn to prove theorems.

In each section of this chapter a different strategy or technique that uses a computer geometry program is discussed, together with one or two examples. Some uses are obvious, but in certain cases the example is followed by a more detailed technical description. You will not miss much if you skip the technical details on first reading, but you may find it valuable to return to them when you try to apply that technique to your own problem. Remember too that you can always load any **Geometer** file into a text editor and discover exactly how it was constructed.

2.1   Accurate Drawings

A computer geometry program can produce quite accurate pictures. If you have tried to make drawings with a straightedge and compass, you know how hard this is—the three lines that are supposed to intersect at a point in fact form a little triangle that has eighth-inch sides, or perhaps the line that is supposed to be tangent to a circle clearly misses or cuts off a sizeable chunk.

The more complicated the drawing, the more the errors compound themselves. Most people agree that it is almost impossible to draw, with a physical straightedge and compass, a reasonable approximation of the construction of the regular heptadecagon (17-sided figure). On a computer screen, the construction is easily accurate to the nearest pixel.

This section contains three examples. The first illustrates an interesting relationship between a pair of constructions. It is simple enough that you can do it by hand with a pencil and paper and then compare your results with the illustration in the book (or on your computer screen). The second example shows a situation where a hastily-drawn figure may lead to an incorrect result which could not happen if a computer geometry program were used. Finally, the third example demonstrates the construction of a regular heptadecagon which is so difficult that it is virtually impossible to do without a computer drawing program.

2.1.1   Harmonic Point Sets

For now, do not worry about what is meant by the term "harmonic point sets"; that will be covered in Chapter 10.

You may find it interesting, however, to try to duplicate the drawing in figure 2.1 using a straightedge and compass. The difficulties are best illustrated if you do not look at the figure, but simply read and follow the directions in the following paragraph. But you will probably find the construction difficult even with the figure at hand.
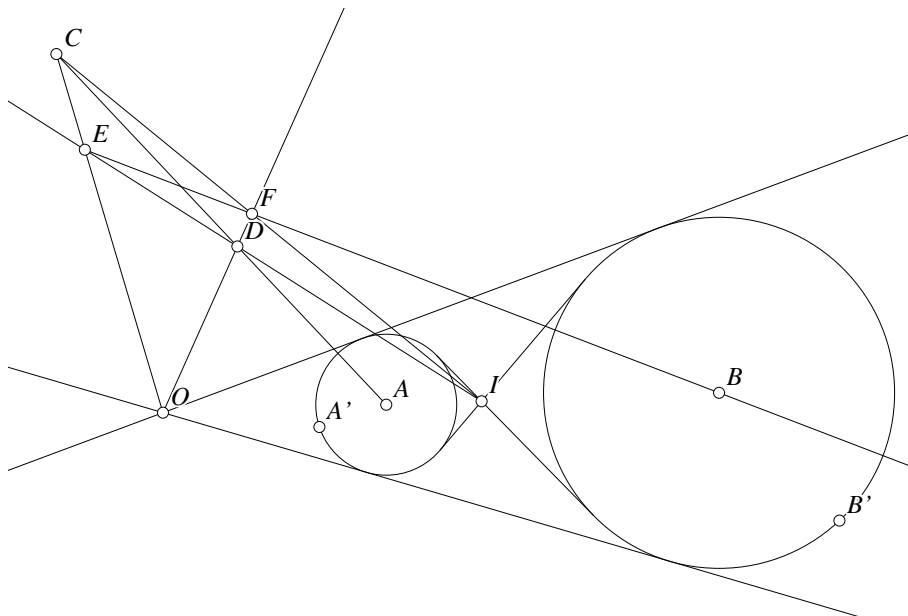
Figure 2.1: A Set of Harmonic Points
`Computer/Harmonic.T` [M]

Construct two unequal circles centered at A and B that neither intersect nor lie one inside the other. Draw the two lines that are the common internal tangents and label their intersection "I", and similarly, draw the pair of common external tangents and label their intersection "O". Now choose any point C not on the line IO and draw the segments CO, CA, and CI. Select any point D on the segment CA and draw the rays ID and OD that intersect the lines CO and CI at points E and F, respectively. Seemingly miraculously, the line EF passes through the point B, and this will be true no matter where you choose to place the points C and D, and independent of the sizes and positions of the original circles.

When you are done, be sure to load the **Geometer** diagram for figure 2.1 and manipulate the sizes and locations of the circles as well as the positions of the points C and D.

2.1.2 "Obviously True", but False

Here is a classic example found in some high school geometry texts that demonstrates the dangers of having a proof rely on a drawing. This sort of thing is a big danger for hand-drawn figures, but it becomes far less of a problem when a computer geometry program is used.

Although figure 2.2 was drawn with **Geometer**, it was not drawn using the built-in constraints, and is, in fact, wrong, as we shall see. It is, however, very much like a drawing that might be made by hand.
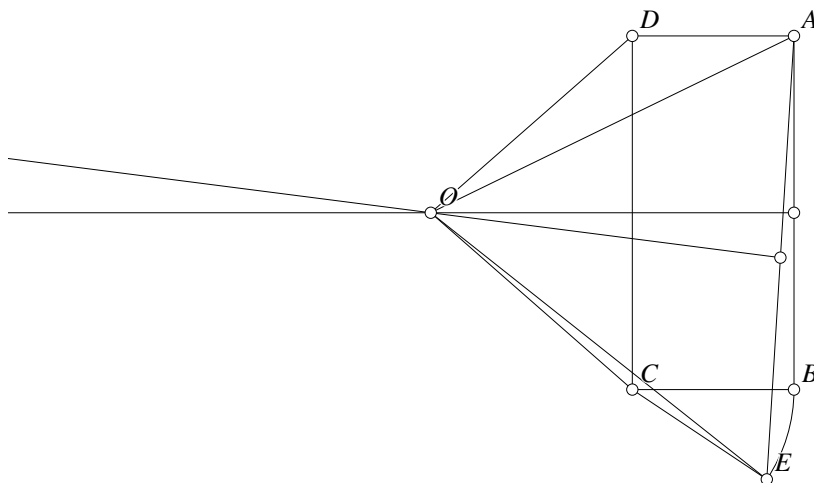
Figure 2.2: Bogus Proof
Computer/Incorrect.D [D]

Start with an arbitrary rectangle $ABCD$ as in figure 2.2, and using a compass, draw an arc of a circle centered at C and passing through B, that goes a little outside the segment BC—to point E.

Since $ABCD$ is a rectangle, $\angle ADC = \angle BCD = 90°$. Since $\angle BCE$ is a little bigger than zero, $\angle DCE$ is a little bigger than $90°$. We will "prove" that $\angle ADC = \angle DCE$—something that is obviously not true.

Here is the bogus argument. Construct the perpendicular bisectors of the segments AB and AE. Since those perpendicular bisectors are not parallel, they will meet at some point O.

Since $ABCD$ is a rectangle, the perpendicular bisector of AB is also the perpendicular bisector of DC, so all the points on it are equidistant from D and C. Therefore $DO = CO$. By similar reasoning, O is equidistant from A and E, so $AO = EO$.

Points B and E are on the same circle centered at C, so $BC = CE$, and since $ABCD$ is a rectangle, $AD = BC = CE$.

To summarize, $DO = CO$, $AD = CE$, and $AO = EO$. Therefore the two triangles $\triangle ADO$ and $\triangle ECO$ are congruent since they share three equal pairs of sides (using SSS congruence). Therefore $\angle ECO = \angle ADO$, and we can subtract the equal angles $\angle CDO$ and $\angle DCO$ to obtain the result we want—that $\angle ADC = \angle ECD$.

The result is clearly not true, but every step seems correct. What is wrong? The answer is that the diagram is not drawn accurately. (In fact, the author cheated and had to misuse **Geometer** to get the desired misleading effect.)

Figure 2.3 shows an accurately drawn diagram of the situation, and it is instantly obvious what went wrong—the line segment EO lies on the outside of the rectangle. In fact, if you check in the correct figure, $\angle ADO = \angle ECO$, which is what we proved.
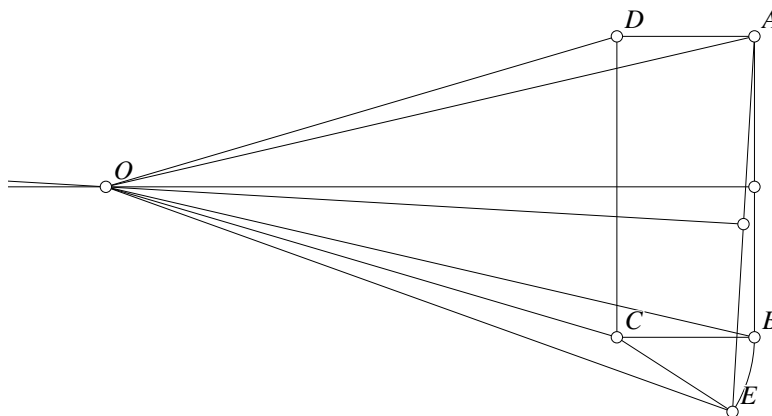
Figure 2.3: Correct Figure
Computer/Correct.T [M]

The point is that an accurate figure can correct some very mysterious problems, and computers allow you to draw extremely accurate figures without too much effort.

2.1.3 Construction of a Regular Heptadecagon

Here is a construction due to Richmond of the regular heptadecagon (17-sided figure). **Geometer** produces figure 2.4. See [Conway and Guy, 1996] for more details of this construction.

To truly drive home the advantages of a computer geometry program over a drawing with pencil, paper, straightedge and compass, you might try (even while looking at figure 2.4) to do the construction yourself as described below.

Even with the figure and the text that describes the construction, you will see that the construction is a bit difficult to follow. If you load the **Geometer** diagram for this figure (found in the directory Computer/Heptadecagon.T on the CD accompanying this book) you can go through the construction one step at a time by clicking repeatedly on the *Next* button in **Geometer** (or pressing the **n** key). This will make the construction much easier to follow.

Begin with a circle in which the regular heptadecagon is to be inscribed. Find its center, and construct horizontal and vertical diameters passing through N, S, E, and W at the "north", "south", "east", and "west" edges, respectively.

By finding the midpoint of NO and then the midpoint between that new point and O again, construct a point A that is $1/4$ of the way from O to N.

Construct the circle centered at A passing through E and use that to bisect the angle $\angle EAS$ twice so that $\angle SAB = \angle EAS/4$. This new angle will cut the horizontal diameter EW at B.

Now construct at A a $45°$ angle relative to AB intersecting EW at C. In other words, construct C such that $\angle BAC = 45°$.
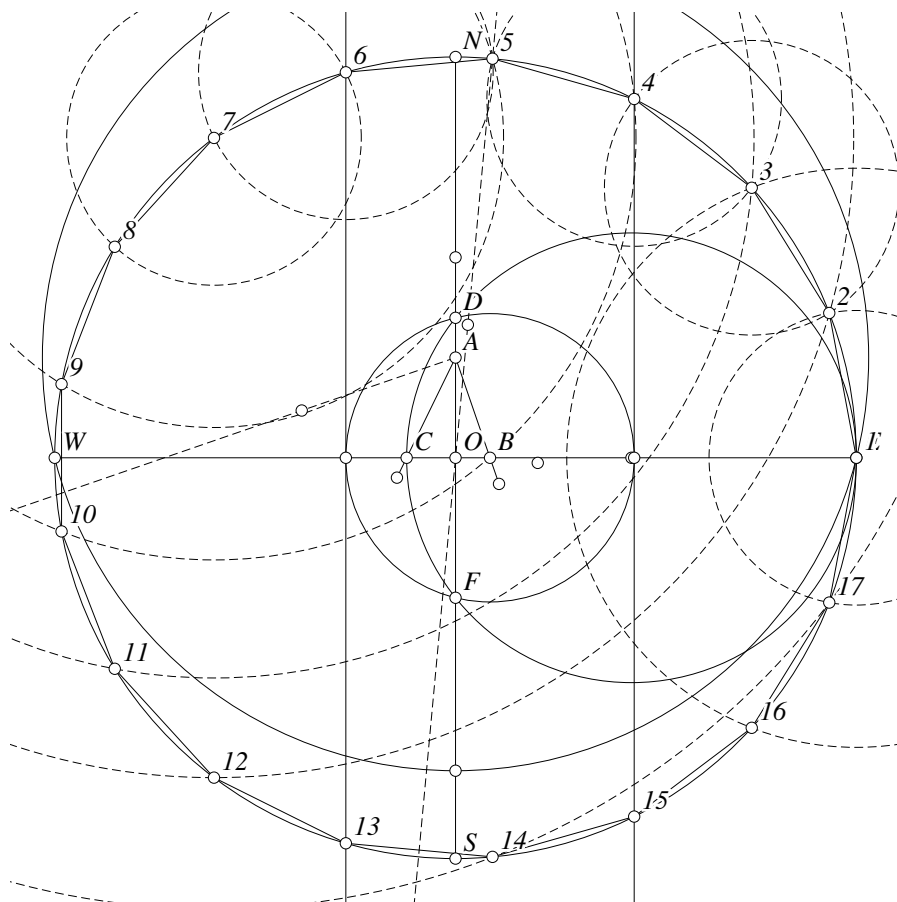
Figure 2.4: Construction of a regular heptadecagon
Computer/Heptadecagon.T [P]

Find the intersections D nd F of the circle having EC as diameter with the line NS.

Construct a circle with center B passing through D (or F), and construct tangents to it that are parallel to NS (this takes a few steps—first find the intersections of the circle with EW, and then construct the tangents there).

If we call the point E "1", and label the intersections of the vertical tangent lines from the last paragraph with the original circle "4", "6", "13", and "14" as shown in the figure, we have 5 points on the heptadecagon, and it is easy to construct the rest. For example, we could bisect the arc $\overset{\frown}{46}$ to find point 5, and once we have the arc $\overset{\frown}{45}$ we can copy it all the way around the circle.

2.2 Drawing Manipulation

Once a figure is drawn in a computer geometry program, it can be manipulated. If you draw one figure, it may be that some amazing relationship holds just by chance, but when you manipulate the diagram, it is extremely unlikely that the chance relationship will continue to hold. This is particularly true for students in their first geometry course. When told to draw a triangle, they will almost invariably draw one that is very nearly equilateral, and for equilateral triangles, all sorts of amazing relationships hold.

In this section we will examine a result that is not as well-known: Miquel's theorem.

2.2.1 Miquel's Theorem
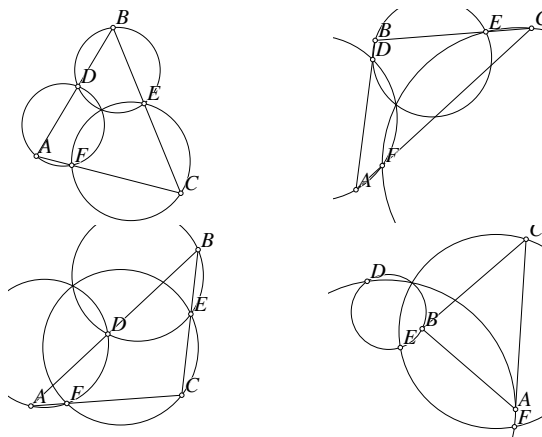
Figure 2.5: Miquel's Theorem
`Computer/Miquel.T` [M]

Given any triangle $\triangle ABC$, choose points D, E, and F on the line segments AB, BC, and CA, respectively. Now, draw three circles through the sets of points $\{A, D, F\}$, $\{B, E, D\}$, and $\{C, F, E\}$. Notice that all three circles seem to pass through the same point. (See the triangle in the upper left of figure 2.5.)

Is this just chance? Try moving the points A, B, C, D, E and F to obtain the other three drawings in figure 2.5. In every case, the three circles seem to pass through a common point, although that point may not lie in the interior of the triangle $\triangle ABC$ (see the example in the lower left), and it even seems to work if the points D, E, and F are anywhere on the lines: D does not have to be between A and B, for example—it just needs to be somewhere on the line AB. (See the example in the lower right.)

This result, called Miquel's theorem, is amazing in that it seems to work no matter what the shape of the triangle, and where the only constraint on the other three points is that they be constrained to the appropriate lines. Well, not quite anywhere: to determine a proper circle, the points A, D, and F must be distinct, et cetera. But the theorem can

even be extended to make sense in that case. If A and F are the same, for example, then the circle passing through A, D, and F can be replaced by the circle tangent to the line AC at A and passing through D.

Load the **Geometer** diagram from `Computer/Miquel.T` and manipulate it to check other configurations, to experiment with the degenerate cases, and generally to amaze yourself with the beauty of Miquel's theorem.

2.3 Finding and Testing Conjectures

2.4 conjecture testing

Accurate drawings allow you not only to test conjectures, but to discover new ones. The various different configurations of Miquel's theorem in Section 2.2.1 above show us that not only does the theorem seem to hold since the circles always seem to intersect in a point, but that the theorem is perhaps more general than we first imagined. It seems to hold, for example, even if the points on the sides of the triangle are "outside" what we normally consider to be a triangle.

2.4.1 Morley's Theorem

Here is another wonderful example of the same sort of thing, based on Morley's theorem.
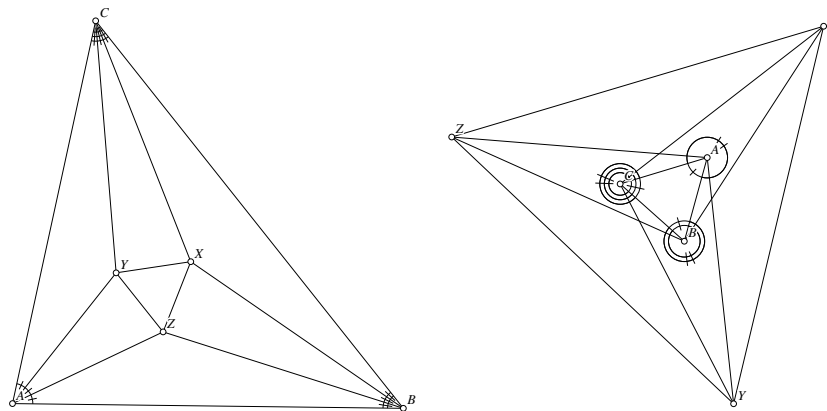
Figure 2.6: Morley's Theorem
`Computer/Morley.T` [M]

Morley's theorem states that if all three of the angles of an arbitrary triangle are trisected as in the diagram on the left in figure 2.6, then the intersection points of adjacent

pairs of trisectors form an equilateral triangle. In the figure, $\triangle XYZ$ is equilateral, and $\triangle ABC$ is arbitrary.

This is extremely interesting, and when you adjust the **Geometer** diagram, it seems to continue to work. But even more interesting is what happens if you turn the triangle $\triangle ABC$ "inside out", as in the diagram on the right of the same figure. (In other words, make the vertices A, B and C go clockwise instead of counterclockwise.) The labeling is the same, and the triangle $\triangle XYZ$ continues to be equilateral, but at first glance, the angles of $\triangle ABC$ no longer seem to be trisected. On closer inspection, however, they *are* trisected, in the sense that it is now the exterior angles that are trisected. (Remember that the trisectors are lines and the figure only shows a segment of the trisecting lines. Extend them in the opposite direction and it will be more obvious that they are the trisectors.) Convert the trisector lines from segments to infinite lines and it is clearer, perhaps, that they remain angle trisectors.

Thus Morley's theorem holds if all the trisectors are of interior angles or all are trisectors of exterior angles. What if two are interior and one exterior? What if two are exterior and one interior? Try it—you will probably be surprised at what you learn.

2.4.2 New Triangle Centers

Quite often, if you draw a figure with some sort of symmetry in the construction, you will find that interesting relationships hold. Thus if you are trying to make up problems for a math class or examination or contest, symmetric "doodling" will often generate interesting relationships.

As a quick example, let us find a couple of triangle centers. The most familiar centers are the centroid (where the three medians of a triangle intersect), the incenter (the intersection of the three angle bisectors), the circumcenter (the intersection of the three perpendicular bisectors of the sides) and the orthocenter (the intersection of the three altitudes). All are interesting

Figure 2.7: An Unknown Triangle Center
`Computer/Tcenter.T` [D]

because in general, three lines do not meet at a point.

It is surprisingly easy to find other triangle centers, and in fact, all of Chapter 7 is devoted to doing so. But figure 2.7 illustrates a simple one.

In $\triangle ABC$, each vertex is reflected across the opposite side resulting in vertices A', B' and C'. If the centers of the circumcircles of $\triangle A'BC$, $\triangle AB'C$ and $\triangle ABC'$ are connected to the opposite vertex as in figure 2.7, all three of those lines intersect at a point which will be a new triangle center. (Notice also that the three circumcircles also seem to intersect at a point. Proving either of these concurrences might be suitable problems for a math exam or contest. Of course you had better make sure you know how to solve them before you assign them as problems. They might be very difficult to prove.)

2.5 Ease of Correction

It is easy to erase your mistakes on a computer screen. This is probably most useful when you accidentally draw the wrong thing, but there are other uses as well.

For example, if you suspect that four points in a diagram lie on a circle, just draw a circle through three of them and see if it goes through the fourth. If it does, manipulate the figure a bit to see if the circle went through that fourth point by chance. In either case, you can just erase the circle afterwards if you want to. If you are working with a pencil and paper, you will probably have to draw a couple of extra construction lines just to draw the circle, so erasing it is more difficult. Besides, if it does go through the fourth point and you want to see if that occurred by chance, you may have to redraw the entire diagram with different measurements.
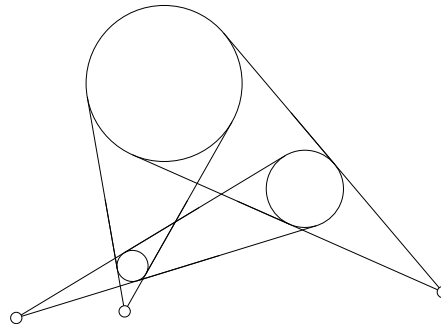
Figure 2.8: Monge's Theorem
`Computer/Mongedemo.T [D]`

Here is a very simple example where the ability to erase might be useful. Draw three unequal circles where none of them intersect and none of them lies completely within another. For each of the three pairs of those circles, draw the common external tangents and identify where those tangents meet. It appears that those three points lie on a line. Is it true? Draw a line through two of them and see if it passes through the third. If so, manipulate the figure and see if it continues to hold. See figure 2.8.

With a physical paper and pencil, lots of auxiliary lines must be drawn to construct the common tangents, and if you wish to test the theorem in different configurations, then far more erasure would be required.

The intersections do indeed lie on a line. The result is known as Monge's theorem and it is proved in Section 9.3.

2.6 Stepping through Proofs

Both teachers of geometry and students may find this feature valuable. The proof of many interesting theorems requires a number of steps and construction lines, and if a student (especially a beginner) is faced with a complex diagram covered with additional construction lines, it is difficult to look through the diagram and see why $\triangle ABC \cong \triangle DEF$. The student must find the points $A, B, \ldots, F$ in a possibly very complex diagram and then notice that the two triangles determined by them are congruent.

If a **Geometer** diagram is constructed as a "Proof", then when the student steps through the proof and reaches the statement that $\triangle ABC \cong \triangle DEF$, both $\triangle ABC$ and $\triangle DEF$ will be highlighted. In addition, the student may then manipulate the

diagram at this point and see that the two triangles remain congruent under all the manipulations. When the student continues to the next step of the proof, the subject of that step will be highlighted, and the highlighting for the previous step will disappear.

Teachers of geometry can prepare, with a little bit of effort, proofs of theorems that can be displayed step by step on a computer screen. The actual construction of such proofs is beyond the scope of this book. To learn the techniques for constructing your own proofs using **Geometer**, see the reference manual, especially the chapter entitled "Teacher's Tutorial", where about 35 pages of text are devoted to this topic.

Fortunately, the CD contains many **Geometer** diagrams in the form of "proofs", for theorems ranging from those covered in a typical high school text to the more interesting advanced theorems covered here. Most of them can be used in a classroom without modification, so teachers of geometry can use them without learning to construct them. The many proofs on the CD also provide good examples for a teacher who wishes to construct her own.

However, it is worth reading through the example below, since it does illustrate the sorts of things that can be done to make it easier to follow a proof on a computer than on a printed page of a book.

We will use the nine-point circle theorem to demonstrate **Geometer**'s ability to present proofs. If you can use a computer, read the rest of this section as you view the diagram on your screen. Load the file `Computer/Ninepoint.T` and you should see an image on your screen roughly like figure 2.9, but with some additional accompanying text.

The statement of the theorem is this:

Theorem 2.1 (The Nine-Point Circle) *Given any triangle, the three midpoints of the sides, the three feet of the altitudes, and the three points midway between the orthocenter and the vertices all lie on the same circle, called the nine-point circle.*

In figure 2.9, let $\triangle ABC$ be the triangle, the midpoints of the sides are A', B', and C', the feet of the altitudes are E, F, and G, the orthocenter (the point where the three altitudes meet) is H, and the points Q, R, and S are midway between H and the vertices A, B, and C. All nine points: A', B', C', E, F, G, Q, R, and S lie on a circle.

With a printed version of the diagram, you are stuck. You can see that it works there, but that is it. On the computer screen, you can move points A, B, and C, and watch how all nine points move in response and how, although it changes size and position, there is always a single circle that passes through the points. You see that it remains true whether the feet of the altitudes lie inside or outside the triangle. You see how for some triangles, the circle is enormously larger. You see that the ordering of the points may change on the circle, but still the theorem holds.

There is more, of course. The illustration in the text is in black and white. On the computer screen, the altitudes and their feet are in red, the midpoints of the sides are in yellow, and the other points Q, R, and S, are in green. Finally, the text on the screen restates the theorem, suggests things you might do to test it, and finally, it says, "Press 'Next' to continue . . . ".

Each time you click on the *Next* button (or press the **n** key on your keyboard), the proof advances by one step. The text under the illustration explains and justifies the step, and the object or objects described by the text can be highlighted in some way. For example, the first step of this nine-point circle proof is to show that $\triangle ABC \sim \triangle C'BA'$. The text explains why this is the case, but in addition, the outlines of those two triangles blink in different colors. This is not so critical in a diagram as simple as this, but it is easy to imagine a complex figure where there are so many points and lines that it is difficult even to find the triangles mentioned. When you advance to the next step of the proof, the colors of the triangles return to normal, and the features important in the next step are somehow highlighted, et cetera.

Figure 2.9: The Nine-Point Circle
`Computer/Ninepoint.T` [P]

Construction lines can appear and disappear at different steps of the proof so as to avoid cluttering the diagram. They appear only when needed.

Another advantage of a **Geometer** presentation is that at any stage in the proof, you can manipulate the points of the diagram. In this proof a nice example occurs in the third step where the segments $A'C'$ and SQ are shown to be parallel. As you look at that step, you can move the point B up and down and notice not only that the lines remain parallel, but that they can swap positions. So if your proof, for some reason, depended upon the fact that $A'C'$ lies above SQ, it is clear that the proof would not be valid.

Of course you can go forward and back through the proof using the *Next* and *Prev* buttons (or the **n** and **p** keys on the keyboard), and get a very good view of how the proof works, including in situations with different configurations of the original triangle $\triangle ABC$.

Any of the diagrams in the book that are identified with a "[P]" after the file name in the caption are proofs (or constructions) whose steps can be followed in the same way with the *Next* and *Prev* buttons.

2.7   Making Measurements

Many geometry problems consist of a diagram with a few of the lengths and/or angles specified and require the calculation of others. Computer geometry programs usually allow you to assign various lengths to lines, et cetera, and then can display the lengths of other lines. As a trivial example, if you construct a right triangle with legs of lengths 3 and 4 and ask to have the length of the hypotenuse displayed, you will be shown a 5 (or, thanks to computational error, perhaps 5.0000001, but that is at least a pretty good hint that the true answer is 5).

This sort of problem is more typical of high school problem sets or mathematics contests, and although **Geometer** can do this, it was primarily designed to work with more general situations, so measurement manipulations may be a bit clumsy. Most of the commercial programs are aimed specifically at high school students and hence may be easier to use in this manner.

2.7.1 Finding an Unknown Length

The most obvious usage is for problems like the one shown in figure 2.10. Given the measurements and the right angles in the drawing, find the unknown length x.

In **Geometer** it is easy to draw a figure to scale. Then if we divide the length of the segment labeled x by the length of the segment labeled 2 and multiply that result by 2, we obtain 10, the length of the segment x. This one is pretty easy to do without a program like **Geometer**, but there are plenty of problems that look this easy but are fiendishly difficult.

Figure 2.10: An Unknown Length
`Computer/Measure.T` [D]

Even when **Geometer** says that $x = 10.000$ we do not have a solution; it might be that $x = 10.0001$. But suspecting that $x = 10$ can lead to a quick solution: The long sides are $AB = 10 + 2 = 12$ and $AC = 5 + 3 = 8$. Since $12 : 3 = 8 : 2$ maybe the triangles are similar: $\triangle AEB \sim \triangle ADC$. They obviously are since both are right triangles sharing the angle at A, so knowing the answer (or, more accurately, strongly suspecting that you know the answer) makes it trivial to prove that the answer is correct.

2.7.2 Finding an Unknown Relationship

In figure 2.11, two circles with centers at A and B are tangent to a line. If a third circle, centered at C, is tangent to both and to the line as shown in the figure, what is the relationship between the radii of the three circles?

The **Geometer** diagram in figure 2.11 is to help gather data. In the diagram, you can set the values of the radii of the circles centered at A and B, and the three
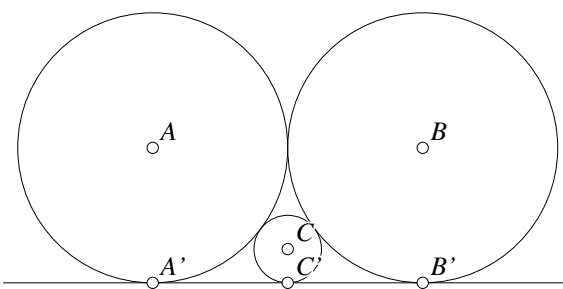
Figure 2.11: Mutually Tangent Circles
`Computer/TanCircs.T` [D]

radii are printed on the screen. (To change the radii, use *Edit Geometry* and change them with the text editor.) For the diagram in figure 2.11, the values $AA' = 0.500$, $BB' = 0.500$ and $CC' = 0.125$ are printed on the screen together with the diagram. Actually, the **Geometer** diagram prints the values of both CC' and $1000CC'$ so you can see more decimal places—**Geometer** normally shows only 3 places of precision when it displays floating point numbers.

It is possible to make a diagram like the one above where you can set the radii with the mouse, but due to the fact that a mouse can only select to the nearest pixel, it is impossible in most cases to obtain nice round values for the inputs.

By repeatedly modifying the initial values for the radii, it is possible to obtain a table of values like this:

AA'	BB'	CC'
1.000000	1.000000	0.250000
1.000000	0.250000	0.111111
1.000000	0.111111	0.062500
0.250000	0.111111	0.040000
0.111111	0.111111	0.027778

You can make this table as large as you want, but we have cheated a bit, selecting values that may make the relationship a bit more obvious.

From the table above, it may be a lot easier to infer that the relationship is this: if $AA' = 1/m^2$ and $BB' = 1/n^2$, then $CC' = 1/(m+n)^2$. Compare this with the Farey circles in Section 2.11.2. (Remember that the values in that section are diameters, not radii.)

♦ Actually, there generally are two circles tangent to the given circles and the line, but will lie outside both circles. The only time this will not occur is if the original circles are equal in size. If the radii of the original circles are $1/m^2$ and $1/n^2$ the outer tangent circle will have radius $1/(m-n)^2$. Note that if the two circles have equal radii, then the outer tangent "circle" essentially becomes a line which is much like a circle with infinite radius. If $m = n$, then $1/(m-n)^2$ is very much like infinity.

2.7.3 ♦Technical Details

Given the two radii as inputs, it is easy, using the pythagorean theorem, to determine the separations of the centers of the circles along the line. If the radii are r and R, the separation of the centers is $2\sqrt{rR}$ which is easily constructed or calculated in a computer geometry program.

The construction of the circle that is tangent to the two circles and to the line is a bit more difficult. The easiest way to do it is to invert the two circles and the line through a circle that is centered at the point of tangency of the two circles (see Chapter 8). Since both circles pass through the center of the point of inversion, they will be inverted to two parallel lines and the line will be inverted to a circle between them that is tangent to both. From here, it is easy to find another circle of the same radius that is tangent to the inverse of the line and to the two parallel lines which, when re-inverted, becomes the circle tangent to all three.

2.7.4 Checking a Calculation
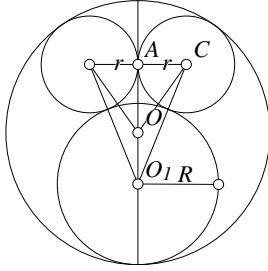
Figure 2.12: Tangent Circles
`Computer/MMouse.T` [M]

In Section 2.7.2 we used **Geometer** to find a mathematical relationship between measurements. Sometimes it can be used in the opposite way, to check an algebraic calculation.

In figure 2.12 we see a large circle with three mutually-tangent interior circles that look a bit like Mickey Mouse. If the radius of the surrounding circle is 1, the radius of the lower inscribed circle is R and the two equal circles above have radius r, find the mathematical relationship between R and r.

If x is the length of AO, we know that $AC = r$, $CO = 1 - r$, $CO_1 = R + r$, and $AO_1 = x + 1 - R$. The pythagorean theorem can be applied to triangles $\triangle AOC$ and $\triangle AO_1C$ to obtain two equations that relate x, r and R. A fair amount of tricky algebra finally yields the following relationship between R and r:

$$r = \frac{4R(1 - R)}{(1 + R)^2}. \tag{2.1}$$

A couple of obvious values can be checked: $R = 0$ and $R = 1$ and these seem to yield the correct value of r, but it is difficult to be certain of the calculation.

One way to check is to draw the figure and then use **Geometer**'s ability to calculate and present lengths and to display both the value of r and of $4R(1 - R)/(1 + R)^2$. As the figure is modified, you can see immediately if the two values differ. In this case, they remain the same, so you can be much more confident of the calculation.

2.7.5 Technical Details

It is a bit tricky to construct a **Geometer** diagram where the radius R can be modified and the two equal mutually-tangent circles are drawn. The easiest method involves inversion in a circle that is discussed in Chapter 8. In addition, you will need to know how to do calculations in **Geometer**, and that is discussed in Section 11.5. To see the details, examine the text of the **Geometer** file `Computer/MMouse1.T`. (The file `Computer/MMouse.T` is simply for making the illustration in the text; it does not contain the **Geometer** code that does the calculation.)

Another alternative is to use the result of your algebraic calculation in equation 2.1 to draw the two circles. The centers of those circles must lie on a circle of radius $1 - r$ from the center of the surrounding circle of radius 1 and must also lie on lines parallel to AO and at distances of r from it. If the calculation is correct, two circles of radius r drawn about these centers should appear to remain tangent to each other and to the other two circles as the radius R is modified. You can modify R and visually check

the tangencies. You can examine the text of a **Geometer** file that uses this technique in
`Computer/MMouse2.T`.

One other minor technical point that is illustrated in the file `Computer/MMouse1.T`
is that in a complicated diagram it is sometimes hard to identify the points needed to
do a calculation. Which are the points, after all, that determine the radius of the various
circles? Using the graphical user interface, simply color those points differently and
then it is easy to see in the text editor which points are of interest. For example, in this
diagram, the points at the center and edge of the circle of radius R are in red, so in the
text editor we merely need to look for points with the text `.red` in their description.

2.8 Searching for Patterns

The **Geometer** program has a mode where
the program searches for patterns that may
hold in a drawing. After a diagram is drawn,
Geometer can be instructed to look for pat-
terns, at which point it searches for unusual
situations in the diagram such as three points
that lie on a line, or four points on a circle,
or angles that are equal or complementary or
supplementary, et cetera. Then as you manip-
ulate the diagram, **Geometer** rechecks these
relationships over and over to see if they con-
tinue to hold. When you tell **Geometer** to

Figure 2.13: Searching for a Proof of
Miquel's Theorem
`Computer/Pattern.T [M]`

stop searching for patterns, it will tell you about any relationships that always held.
Obviously these are not necessarily always true, and they are "true" only to the degree
of accuracy to which **Geometer** can do arithmetic calculations.

But knowing, for example, that four seemingly unrelated points always lie on a
circle is often tremendously important in finding a proof of whatever theorem you are
trying to prove.

As a simple example, let us consider again Miquel's theorem that we examined in
Section 2.2.1. This time, instead of simply convincing ourselves that the theorem is
true, we will try to find a proof.

Begin with the same diagram that we used in Section 2.2.1 and we will add a point
G that is the intersection of circles BDE and AFD. Connect G to all six other labeled
points, and then ask **Geometer** to search for patterns using the *Test Diagram* and *End
Test* commands. The diagram should look something like what is shown in figure 2.13
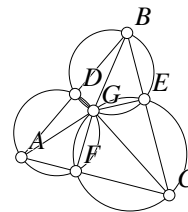
Move points A, B, and C a bit, and then end the test. A small auxiliary window
appears on your screen that contains the following information:

```
Points on a circle:
   C  E  F  G
Equal angles:
```

```
    (A D G) (B E G)
Supplementary angles:
    (A D G) (B E G)
 are supplementary to
    (A F G)
Supplementary angles:
    (B C A)   is supplementary to  (E G F)
Supplementary angles:
    (B A C)   is supplementary to  (D G F)
Supplementary angles:
    (A B C)   is supplementary to  (D G E)
Concurrent circles:
    c1 c2 c3
```

First of all, it tells you that the point G is also on the third circle ($CEFG$) which you are trying to prove, and it also tells you that the three circles are in fact concurrent. (**Geometer** uses the internal names `c1`, `c2`, and `c3` of the circles since they have no user-defined names[1]. The points are all listed using the names you gave them.)

But there is also information about equal and supplementary angles. Just look at one: $\angle BCA$ is supplementary to $\angle EGF$. This is the key to the proof! You know that two sets of four points: $BEGD$ and $AFGD$ are concyclic, so $\angle FAD$ is supplementary to $\angle DGF$ and $\angle DBE$ is supplementary to $\angle EGD$. The three angles of $\triangle ABC$ add to $180°$ so with a little arithmetic, it is easy to show that $\angle EGF = 360° - \angle DGF - \angle EGD$ is supplementary to $\angle ACB$, which is enough to complete the proof.

The lines connecting G to A, B, and C were unused, but since we had no idea how to start the proof, it did not hurt to test them. For all we knew, the construction of those lines might have been the route to the solution.

Geometer attempts to show only "interesting" relationships. In the example above, it only showed one set of four points lying on a circle when in fact there are three. But the first two sets were required by the constraints on the diagram to lie on a circle. Similarly, if a line is defined by two points and you constrain a third point to lie on that line, **Geometer** will not tell you that those three points are on a line—they had to be on a line.

Searching for patterns is a very powerful proof finding technique. Many of the relationships displayed by **Geometer** will be the entries on the left side of a high-school "statements—reasons" two-column proof. You just have to put them in the correct order.

Although it usually works, the technique is not foolproof. The relationships are only calculated to an accuracy limited by the floating point precision of the machine, so as far as the internals of the program are concerned, the relationships hardly ever hold *exactly*. But if they hold to a certain tolerance and continue to hold as the diagram is modified, **Geometer** will report them at the end.

---

[1]By default, only points have names, but you can, if you wish, assign names to other **Geometer** primitives.

2.9   Searching for Extremes

This is just another measurement, in a sense, but suppose you have a question like this: "What is the minimum perimeter of a triangle whose vertices lie on the edges of a given triangle?" You can draw the triangle constrained to have its vertices on the edges of the given triangle and display its perimeter. Then you can manipulate the inner triangle as you watch the value of the perimeter and easily find a very close approximation of the triangle that does the trick. If the original triangle has all acute angles, the answer is the triangle whose vertices are the bases of the altitudes of the original, which you may be able to notice by inspection. What happens if the original triangle contains an obtuse angle? (See Fagnano's problem in Chapter 12.)

A different, but similar, example is described below.

2.9.1   Maximum Area

Suppose that a pentagon is inscribed in a circle. What is the arrangement of the vertices that gives the maximum area? It is easy to create a **Geometer** file that lets you manipulate the vertices and continuously prints the area. With a random arrangement of points on the left in figure 2.14, a certain area is enclosed.

You can then slide any of the points on the circle, constantly trying to find larger areas, and after some playing around, you will come up with a picture like that on the right in figure 2.14 where it seems impossible to increase the area. Notice that the pentagon in this second figure appears to be very nearly a regular pentagon with all sides and angles equal.

In fact, the regular pentagon is the largest possible; this exercise does not prove that to be the case, but it certainly provides plenty of visual evidence which would encourage you to seek a rigorous mathematical proof. An interesting related exercise is to determine the size of the largest equilateral triangle that can be inscribed in a square.

Another interesting example can be examined in the file `Computer/MaxQuad.T` that illustrates the fact that given the lengths of the sides of a quadrilateral, the quadrilateral of maximum area that can be produced from those sides is the one that is concyclic.

2.9.2   ♦Technical Details

Geometer has a command available through the text editor to evaluate the area of a polygon. The code looks like this:

```
poly1 = .polygon(5, v6, v7, v3, v4, v5);
area = .f.area(poly1, "Area");
```

The first line above was generated by **Geometer** from the graphical user interface by clicking on *P.P=>Poly* and then clicking on the points around the pentagon in
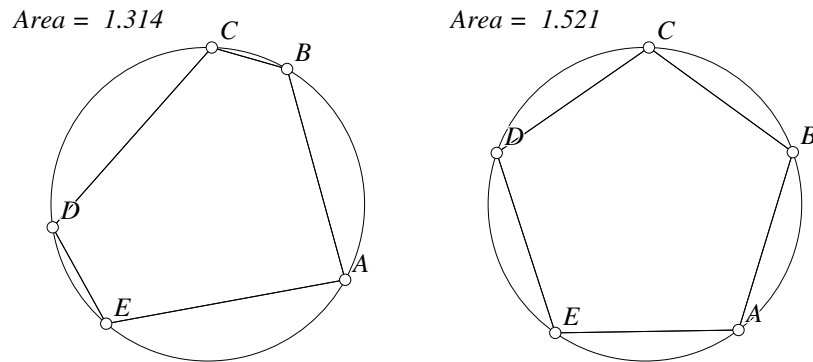
Figure 2.14: Area of a Pentagon
Computer/PolyArea.T [M]

order and then clicking on the first point to close and complete the polygon. The internal name is `poly1`, so the `.f.area` command takes `poly1` as a parameter, and a floating point number called `area` is generated. Since the object called `area` contains the text `"Area"` in its definition, the value of the float is continuously displayed on the **Geometer** screen, something like this: "`Area = 1.132`".

2.10 Locus of Points

If you are asked to find the set of points equidistant from a fixed point and a line, it is trivial with **Geometer** to construct one such point. If that point is colored in a "smearing" color, and is dragged around, you will discover that it paints hundreds of points, all of which seem to lie on a parabola. Although this is not a proof that the locus is a parabola, it provides almost overwhelming evidence. This can, of course, be done by hand, but the process of doing so is much more painful and error-prone. In addition, once the diagram that illustrates this relationship is drawn, you can move the fixed point relative to the fixed line and see how the resulting parabola changes size and shape.

2.10.1 Parabola as a Locus of Points

Here is how to use **Geometer** to draw the locus described above. First, draw a line L, a point P not on L, and constrain a point X to lie on L. All points equidistant from X and P lie on the perpendicular bisector of those points, and the point of interest is on the line through X and perpendicular to L. The intersection of those lines is a point equidistant from X on L and the point P. As X is moved along L, all such points are drawn in a smearing color which generates the diagram in figure 2.15. Note that you can move P relative to L (before sliding X back and forth) to generate different shapes of parabolas.
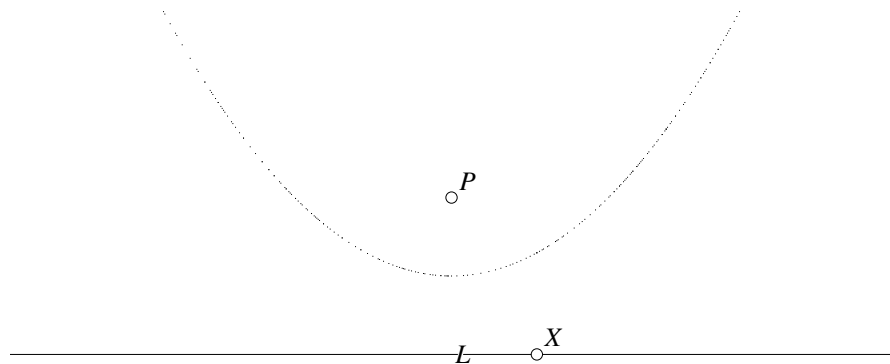
Figure 2.15: Parabola as a Locus of Points
`Computer/Parabola.T` [M]

2.10.2   Cardioid/Limaçon

It is almost equally easy to generate a far
more interesting family of curves including
the cardioid and the limaçon of Pascal. Fig-
ure 2.16 shows a limaçon since the point P
does not lie on the circle.

Given a circle K and a point P not on the
circle, the limaçon of Pascal is the set of all
points that are the projection of P on the lines
tangent to the circle K. In figure 2.16, you
can move the point M around the circle and
trace out a limaçon. If the point P happens
to lie on the circle, a cardioid is generated in-
stead. It is interesting to explore the shapes
generated when the point P is different dis-

Figure 2.16: The Limaçon of Pascal
`Computer/Limacon.T` [M]

tances from the circle, including on and inside the circle K.

2.11   ♦Computer-Generated Diagrams

In **Geometer**, every diagram is saved as a file in text format that can be inspected and
edited. **Geometer** itself provides a primitive editor, but there is no reason you cannot
use your favorite powerful text editor to do the same thing. In addition, you can write
a computer program that generates a very complicated diagram simply by generating
the text form of a **Geometer** file.

Geometer is usually used with the standard graphical user interface, but at times,
it it easier to use the built-in text editor to make some changes. One of the beauties of

the program is that you can flip back and forth: use the graphical interface when it is easy, switch to the text editor for a while, then back to the graphical interface, et cetera.

This is an advanced technique, but it can be incredibly powerful.

Warning: Unless you understand the internal form of the **Geometer** diagram files, some of the examples in this section may be a bit mysterious. The examples are not difficult, however, and anyone familiar with computer programming should have no difficulty understanding what is going on.

Some diagrams are difficult to generate because they are large, or because they require repeated complex operations, or they require lots of specific exact coordinates.

2.11.1 Drawing a Grid

As a very simple example, suppose you want to draw a grid of equally-spaced lines with 51 lines in each direction to make a grid of 50×50 squares as in figure 2.17. It is quite difficult to generate something like this by hand, but with a tiny computer program written in your favorite language (the author's favorite language is C or C++), the generation becomes trivial.

Figure 2.17: A 51 by 51 Grid
`Computer/Grid.T` [D]

What you need is simply a file that looks like 51 copies of the first 6 lines of the **Geometer** code below, but where each successive set of lines has different sets of endpoints for the lines. The second 6 lines illustrate what should appear as the second set of lines. The pattern continues for 51 sets of those six lines.

```
v01 = .pinned(-1, -1.000000, .in);
v11 = .pinned(1, -1.000000, .in);
lh1 = .l.vv(v01, v11);
v21 = .pinned(-1.000000, -1, .in);
v31 = .pinned(-1.000000, 1, .in);
lv1 = .l.vv(v21, v31);
v02 = .pinned(-1, -0.960000, .in);
v12 = .pinned(1, -0.960000, .in);
lh2 = .l.vv(v02, v12);
v22 = .pinned(-0.960000, -1, .in);
v32 = .pinned(-0.960000, 1, .in);
lv2 = .l.vv(v22, v32);
```

Here is a tiny program written in the C language that will generate a **Geometer** file that will do the trick:

```c
#include <stdio.h>
#define GRIDSIZE 51

main() {
    int i;
    double x;
    int n = 1;

    for(i = 0; i < GRIDSIZE; i++) {
        x = -1.0 + i*2.0/(GRIDSIZE-1.0);
        printf("v0%d = .pinned(-1, %f, .in);\n", n, x);
        printf("v1%d = .pinned(1, %f, .in);\n", n, x);
        printf("lh%d = .l.vv(v0%d, v1%d);\n", n, n, n);
        printf("v2%d = .pinned(%f, -1, .in);\n", n, x);
        printf("v3%d = .pinned(%f, 1, .in);\n", n, x);
        printf("lv%d = .l.vv(v2%d, v3%d);\n", n, n, n);
        n++;
    }
}
```

The 306-line file generated by the program above produces figure 2.17.

2.11.2 Farey Circles

The example above is quite simple; here is one that is more interesting—it displays some of the so-called Farey circles.

The Farey circles consist of all circles of diameter $1/q^2$ that are tangent to the real number line at the point p/q, where p/q is a rational number reduced to lowest terms. See figure 2.18. The amazing thing about these circles is that none of them pass through another, but many pairs are tangent.
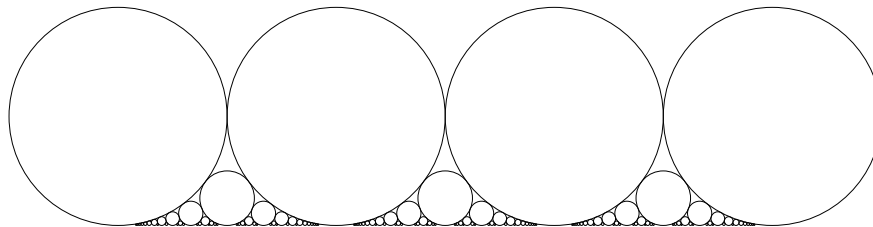
Figure 2.18: Some Farey Circles
Computer/farey.T [D]

To generate a list of circles for **Geometer** to draw, we would like to make a sequence of entries in a **Geometer** file that look like this:

```
v6 = .pinned(2.500000, 0.125000, .in);
l6 = .l.vlperp(v6, l, .in);
vv6 = .v.ll(l, l6, .in);
c6 = .c.vv(v6, vv6);
```

In the example above, the circle above the point 2.5 on the real line corresponds to the rational number $5/2$, or $p = 5$ and $q = 2$. This circle must be centered at $(2.5, .125)$ since its diameter is $1/q^2$ and hence must have radius $1/8 = .125$. To construct this circle in the example code above, the point v6 is pinned with coordinates $(2.5, 0.125)$. From that point, a perpendicular line l6 is dropped to the line l, which is defined once in the file as the real axis. The point vv6 is the intersection of those lines, and finally, the circle that interests us is centered at v6 and passes through vv6.

Here is the C code that generates the complete **Geometer** file, including code to generate the line l:

```c
#include <stdio.h>
#define NMAX 12

int gcd(int i, int j) {
    if (i % j == 0) return j;
    return gcd(j, i%j);
}

void outcircle(int q, int p) {
    static int n = 0;
    printf("v%d = .pinned(%f, %f, .in);\n", n,
            ((float) p)/((float) q), 0.5/((float)q*q));
    printf("l%d = .l.vlperp(v%d, l, .in);\n", n, n);
    printf("vv%d = .v.ll(l, l%d, .in);\n", n, n);
    printf("c%d = .c.vv(v%d, vv%d);\n", n, n, n);
    n++;
}

main() {
    int p, q;

    printf("v00 = .pinned(0, 0, .in);\n");
    printf("v01 = .pinned(1, 0, .in);\n");
    printf("l = .l.vv(v00, v01, .in);\n");
    outcircle(1, 0);
    for (q = 1; q <= NMAX; q++)
        for (p = 1; p <= 3*q; p++)
            if (gcd(q,p) == 1) outcircle(q,p);
}
```

The constant NMAX is 12, which is the largest denominator that the program will consider. The routine gcd finds the greatest common divisor of two numbers which is used to make certain that the fractions are reduced to lowest terms. The routine outcircle prints all the data needed to draw a circle over the point p/q with radius $1/q^2$. The main routine prints **Geometer** code to draw the line l at $y = 0$ and then marches through all possible combinations of q and p and generates circles when p/q is a fraction reduced to lowest terms. This C code with NMAX equal to 12 generates a file containing 559 lines of **Geometer** code.

Here is an amazing property of the Farey circles that you may wish to investigate. If the two circles corresponding to the fractions a/b and c/d are tangent to each other and to the line $y = 0$, then the circle corresponding to $(a + c)/(b + d)$ is tangent to both of them as well as to the line $y = 0$. See [Conway and Guy, 1996].

Note: the radii of the Farey circles can be calculated using the Descartes circle theorem (see Section 2.14) where we consider the straight line to be a circle of infinite radius (or equivalently, zero curvature).

2.12 Animations

This is again a useful feature for teachers of geometry. The **Geometer** diagrams can be programmed to run through a fixed sequence of steps when the user presses the *Run Script* button. An animation of a process often makes things much easier for the students to visualize. This is also an advanced technique that requires some programming in the **Geometer** diagram file.

The spirograph example in Section 2.15.3 is really an animation where the angles in the virtual spirograph disks are incremented in a series of successive equal steps, but a far more interesting one is illustrated in the series of images in figure 2.19 that can be used to convince a student that the area of a circle of radius r is given by the formula πr^2.

The animation divides a circle into a number of wedges, moves the wedges into two lines, and then pushes the lines together to form a rectangle. The three figures illustrate successive steps in this process which is continuous on the computer screen.
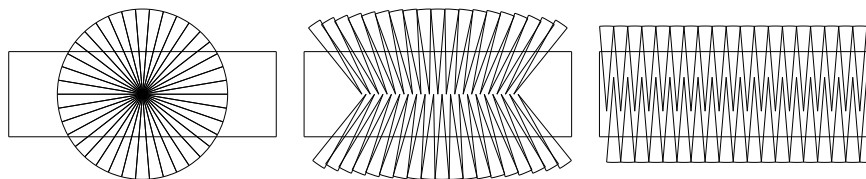
Figure 2.19: Circle in Wedges
`Computer/Circle.T` [S]

2.12.1 ♦Technical Details

This particular demo is perhaps the most complex that has ever been written for **Geometer**. For each of the wedges, a computer program written in C generated a translation/rotation pair that would gradually move it from its initial position on the circle into the final alignment. A script was written that steps a variable from 0 to 1, and while the variable is less than 0.5, the translation/rotation is performed. While it is larger than 0.5, a translation from the final "unwrapped" position is performed to push the two halves together. The coordinates for the slices were also computer generated.

2.13 Drawing for Publication

Geometer is a great drawing program, and in fact, it was used to generate every illustration in this book. When you get a diagram that you like on the screen, **Geometer** can produce from that diagram an "EPS" (Encapsulated PostScript) file that will draw the same figure at any magnification, and in a form that can be included in many text manipulation programs. Painting and image manipulation programs like Adobe's Photoshop® can convert EPS files to almost any other graphics file format. All of Chapter 11 is devoted to making drawings and presentations easier to understand.

One advantage of a computer geometry program for producing illustrations that usually becomes obvious only when you try to do it is that with a program like **Geometer**, after drawing the figure, you can adjust the size and shape to fit the space available on the page, you can adjust the shape so that important labels do not fall on top of each other, et cetera.

Figure 2.20: Dirichlet Domains
`Computer/Dirichlet.T` [M]

Here, the word "publication" means anything where you would like to draw an accurate figure, from a published book to a handout for class. All the figures in this book were produced using **Geometer**, but the example in this section was originally prepared for a talk on computational geometry, not Euclidean geometry.
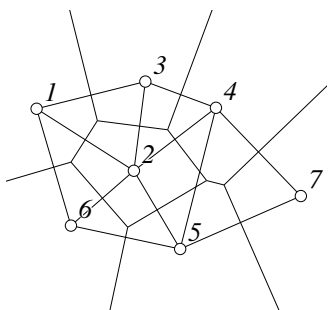
Figure 2.21: Dirichlet Domain Construction
`Computer/Dirichlet.T` [M]

A set of n distinct points in the plane divide the plane into n distinct regions called Dirichlet domains where the points in each region are closest to one of the points. Points on the boundaries of the regions are equidistant from two or more points. Figure 2.20 illustrates the Dirichlet domains for a set of 7 points.

It is not difficult to generate the domains. All the boundaries between the domains are segments or rays of the lines that are perpendicular bisectors of the segments between the points. Some of those segments connecting pairs of points are displayed in figure 2.21 and it is obvious that the boundaries of the Dirichlet domains are perpendicular to those segments.

To construct the diagram, the perpendicular bisectors were drawn, their key intersections were identified with additional points, and then the lines themselves were made invisible. Next, visible lines were drawn only along the true boundaries of the

domains. Notice that not all the perpendicular bisectors are used; the ones used depend on the locations of the points.

Since this diagram was solely for publication, and not for manipulation, it will work only if the points determining the domains are not moved too much. To see what happens with too much movement, load the **Geometer** diagram and try making both small and large movements of some of the points. The diagram is slightly flexible, but with too much variation, it breaks down. Since it was only to be shown in printed form, that was not a problem.

2.13.1 ♦Technical Details

To draw the Dirichlet domains, first the set of seven points was placed in reasonable positions. Then all the perpendicular bisectors were drawn as lines that extend forever in both directions in an unusual color. By examining the particular configuration of the points, it was easy to see by inspection which sections of the long lines mattered in the final figure.

Points were placed at the key intersections, and those were connected by line segments of the final color. Once the key part of a perpendicular bisector was redrawn as a segment in the correct color, the infinite bisector was changed to the invisible color. This process was repeated until the entire Dirichlet domain illustration was complete.

Obviously, this only works for small movements of the points; larger movements will generate totally bogus results.

2.14 Computer Algebra Systems

Although we will do almost nothing with them in this book, there are commercial computer algebra systems such as Mathematica®, Maple® and Macsyma® that are designed to manipulate algebraic expressions. Since it is often easy to write down a set of equations that relate various geometric measurements but very difficult to solve them, computer algebra systems can provide a powerful tool to manipulate and solve these equations.

As an example, consider the configuration of circles illustrated in figure 2.22. Given three pairwise-tangent circles, there are two other circles that are tangent to all three. In the figure, the one enclosed by the three is shown, and there is an additional circle (not shown) that surrounds the three circles. Circles in this configuration are sometimes called **Soddy circles**.

A relationship between the curvatures of the circles is given by the circle theorem of Descartes. The curvature of a circle is the inverse of the radius, so if the radii of the four circles are r_a, r_b, r_c and r_d, then their curvatures are $a = 1/r_a$, $b = 1/r_b$, $c = 1/r_c$ and $d = 1/r_d$ which are related by the following formula:

$$2(a^2 + b^2 + c^2 + d^2) = (a + b + c + d)^2.$$

In the figure, we can see that four triangles are formed—the outer one, $\triangle ABC$, and it is composed of three inner triangles. A similar diagram can be drawn would relate the radii with the surrounding tangent circle. The sides of the outer triangle have lengths $r_a + r_b$, $r_b + r_c$ and $r_c + r_a$ and a substitution of those lengths into Heron's formula (see Section 5.6) tells us that the area $\mathcal{A}(\triangle ABC)$ is given by:

$$\sqrt{r_a r_b r_c (r_a + r_b + r_c)}.$$

Exactly similar expressions exist for the three triangles that compose it and their areas add to give the area of the larger triangle so the radii satisfy the following equation:

Figure 2.22: The Descartes Circle Theorem
`Computer/Soddy.T [M]`

$$\sqrt{r_a r_b r_c (r_a + r_b + r_c)} = \sqrt{r_a r_b r_d (r_a + r_b + r_d)} + \sqrt{r_a r_c r_d (r_a + r_c + r_d)}$$
$$+ \sqrt{r_b r_c r_d (r_b + r_c + r_d)}.$$

It is conceptually easy to eliminate the radicals from the equation above. If we let A, B, C and D represent the expressions under the radicals, the equation looks like this: $\sqrt{A} = \sqrt{B} + \sqrt{C} + \sqrt{D}$ which can be simplified as follows:

$$\begin{aligned}
\sqrt{A} &= \sqrt{B} + \sqrt{C} + \sqrt{D} \\
\sqrt{A} - \sqrt{B} &= \sqrt{C} + \sqrt{D} \\
(\sqrt{A} - \sqrt{B})^2 &= (\sqrt{C} + \sqrt{D})^2 \\
A + B - 2\sqrt{AB} &= C + D + 2\sqrt{CD} \\
A + B - C - D &= 2\sqrt{AB} + 2\sqrt{CD} \\
(A + B - C - D)^2 &= (2\sqrt{AB} + 2\sqrt{CD})^2 \\
(A + B - C - D)^2 &= 4AB + 4CD + 8\sqrt{ABCD} \\
(A + B - C - D)^2 - 4AB - 4CD &= 8\sqrt{ABCD} \\
((A + B - C - D)^2 - 4AB - 4CD)^2 &= 64ABCD
\end{aligned}$$

This seems simple enough, but recall that A, B, C and D are each complex expressions. When the final equation above is expanded, there are approximately 400 terms of degree 16. This can be factored and substitutions made of $1/a$ for r_a, et cetera, and Descartes' theorem drops out. When printed, the complete Maple calculation requires

about 18 pages of equations. This is not something that you would like to do by hand, but it only requires a few minutes on a computer.

Note: Descartes' circle theorem is quadratic so for every set of three radii (or curvatures) there are two solutions for the fourth. If the solution is negative, the result is the surrounding circle; if positive, it is the surrounded circle.

2.15   Non-Mathematical Uses of Geometer

We can often use **Geometer** as a general drawing program where the ability to satisfy geometric constraints is simply a useful feature. It has been used to illustrate mechanical linkages, to draw illustrations from computational geometry and chemistry, or simply to produce works of art. In fact, Appendix B contains a description of the somewhat artistic illustrations that appear at the beginning of each chapter of this book.

This section includes three uses for the **Geometer** program that are a bit outside the realm of geometry, one in mechanical engineering, one in optics, and one that simply generates a pretty picture.

2.15.1   Peaucellier's Linkage

For many years, it was unknown whether a mechanical linkage could be designed that would convert circular motion (such as the output of an engine) into perfectly linear motion. Peaucellier's linkage does just that, as shown in figure 2.23. (The diagram is

Figure 2.23: Peaucellier's Linkage
`Computer/Peaucellier.T [P]`

marked with a "[P]" because not only does it allow you to manipulate the linkage, but it contains a **Geometer** proof that the linkage does, in fact, generate a straight line.)

In the figure, the point O is fixed on the circle and the point A is constrained to move around the circle. The segments $OC = OD$ and $AC = CA' = A'D = DA$ are solid bars that are connected together with pivots at O, A, C, A', and D. You can change the lengths of those bars by moving the endpoints of the segments labeled "l" and "r" on the left side of the figure.

As you move the point A around the circle, the point A' traces out a straight line, a few points of which are shown in figure 2.23.

It is interesting to construct a similar figure where the point O does not lie on the circle. What is the relationship then between points A and A'?

2.15.2   Focus for a Spherical Lens

Snell's law in physics allows you to calculate the angle through which light is bent when it passes a surface between two substances with different indices of refraction. In

fact, the index of refraction is just the ratio of the sines of the angles of incidence and refraction.
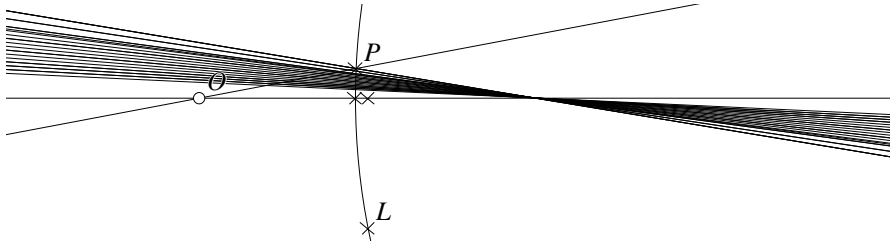
Figure 2.24: Good Focus with a Spherical Lens
`Computer/lens.T` [M]
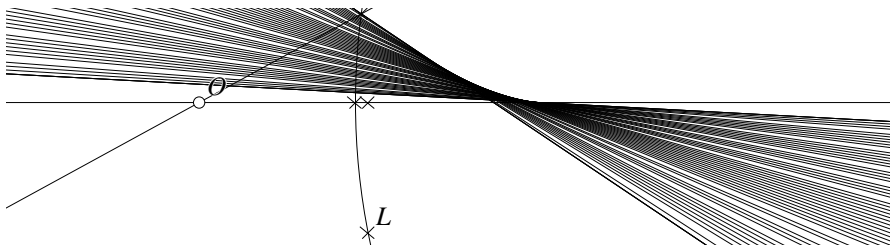
Figure 2.25: Bad Focus with a Spherical Lens
`Computer/lens.T` [M]

In this example, we imagine a sphere of glass with index of refraction 1.5 and light emitted from a point O outside the glass. When light from O strikes the glass at a point P, the **Geometer** diagram finds the angle of incidence with the glass, calculates the angle of refraction and draws a ray from P in that direction. As you move the point P on the surface of the glass, all those rays are smeared onto the diagram.

Figure 2.24 shows that if P is only a little bit away from the axis of the lens, the rays very nearly focus, but if there is more movement, as in figure 2.25, the rays do not come close to being in focus. The shape of the lens and the position of O can be changed in the diagram to experiment with different lens shapes and with different object distances.

2.15.3 Pure Art: A Spirograph

As children, most of us played with a Spirograph®—a toy consisting of geared plastic circles that could be rotated around each other, and as they did, they dragged a pencil that generated wonderful patterns. With the plastic version, you were pretty much limited to two gears, but it is easy to write a short **Geometer** program that has any number. The example shown in figure 2.26 consists of three virtual gears, and requires only a 21 line **Geometer** file.

2.15.4 ♦Technical Details

The lens demonstration began with standard **Geometer** commands to draw and control the positions and shape of the lens and the position of the object point to be focused. Then a point was constrained to be on the surface of the lens, and a line drawn from there to the object. A line from the center of the circle was drawn to the point on the surface (which would be perpendicular to the lens surface). This and the incoming line segment form an angle whose sine is calculated using **Geometer**'s internal calculator and multiplied by an appropriate constant to find the sine of the angle of refraction. A new line in the smearing color was generated passing in that direction, and all other construction aids were changed to the invisible color.

Figure 2.26: A Fancy Spirograph Drawing
`Computer/Spirograph.T` [S]

The spirograph points are simply calculated as a series of three rotation/translation pairs, as if three gears with different radii and hence numbers of teeth were hooked one to the other in a sort of super-spirograph. Examine the **Geometer** code to see exactly how this is done. Using a script that takes all the gears through one complete cycle, a single point at the origin is put through the three translation/rotation transformations and a point is drawn in the smearing color at the final resulting location.

Chapter 3

Geometric Construction

To use **Geometer** or any computer geometry program to investigate new problems you will need to be able to draw diagrams that correspond to those problems. In principle, a vast collection of drawings can be made with only two tools—a straightedge and a compass—but to draw complex diagrams quickly, the ability to use a larger set of tools is extremely valuable.

The classical Greek construction problems that allow only the straightedge and compass as tools are certainly interesting in their own right, and there is a large literature associated with them. In this chapter when we wish to refer to that sort of problem, we will call it a **classical construction** problem.

We will examine the classical construction problems later in Section 3.11, but if you are planning to use a computer geometry program, you are going to be forced to face your own construction problems right away.

If you would like to illustrate a theorem with a computer drawing that can be manipulated, you need to draw it so that it has the same degrees of freedom that the theorem

allows. Luckily, computer arithmetic, although not perfect, is accurate enough that your drawings will seldom be in error by more than a pixel or so, if they are drawn correctly.

Every computer geometry program, including **Geometer**, certainly supports the ability to do classical straightedge and compass constructions, but all of them have additional capabilities. Some of these capabilities are just shortcuts to skip some steps in a classical construction, but some allow you to do things that are far more powerful than what a straightedge and compass would allow (see Section 3.12).

Of course all the examples here are based on **Geometer**'s tools, although it is likely that similar operations are available in other programs.

3.1   The Circumcenter

As a warm up, suppose you want to examine the circumcenter of a triangle and to see how it moves as the size and shape of the triangle changes. (The circumcenter is the center of the circle passing through the three vertices of the triangle.) Your goal, therefore, is to make a **Geometer** diagram with a triangle $\triangle ABC$, all of whose vertices are movable, and with the circumcenter of that triangle visible and constructed in such a way that it adjusts its position automatically as the vertices of the triangle are moved.

The construction can be performed with classical techniques. The center of the circumcircle must be equidistant from the three vertices, so one way to construct it is to find the perpendicular bisectors of the segments AB and AC (or any other pair of the three segments) and their intersection is the circumcenter. All the points on a perpendicular bisector of a segment are equidis-

Figure 3.1: Construct the Circumcenter
`Construction/Circumcenter.T` [M]

tant from the endpoints, so the point at the intersection of the bisectors will be equidistant from all three vertices. This will be the required center of the circumcircle.

The usual classical construction of the perpendicular bisector of a segment involves drawing circles centered at each endpoint and passing through the other endpoint. The two intersections of those two circles determine a line. This line is the perpendicular bisector of the segment. The entire classical construction of the circumcenter is illustrated in figure 3.1.

If you are beginning to use **Geometer**, this construction is a good exercise. You will begin to learn where the controls are, and you will see how a construction can be carried out with a minimum number of tools. As you are drawing, especially since you may have to stop and think about where to find the tools you need, you may get confused about which parts of the construction are which. Remember that you can

change the colors of lines, points and circles, and you can use that sort of marking to help remind you where you are in the construction. For example, color the lines of the original triangle differently, and use red construction lines to get the perpendicular bisector of the red side and green ones when you are working on the green side.

The classical approach requires eleven steps after the triangle is available. Four circles are drawn, four intersections of the circles are determined, two lines are drawn through these intersection points, and finally the intersection of those lines is the required circumcenter.

This is a perfectly good construction, but as you can see, there are a lot of auxiliary lines, and altogether it required eleven steps to find the circumcenter.

Geometer has shortcuts. For example, there is a built-in command to construct the perpendicular bisector, after which only their intersection would need to be found for a total of three steps. But there is an even faster way: **Geometer** has a command (called *PPP=>C* in the command area under the *Circles* label) to construct a circle passing through three points (this would be the circumcircle) and another to find a point at the center of a circle, so the lazy **Geometer** user can construct the circumcenter of a triangle in two steps. The command to find the point at the center of a circle is in the *Primitives* pulldown menu under the *Point* submenu and is called *C=>P Ctr*.

3.2 Available Tools

Even if you do not know how to use all of them, it is good to have at least a vague idea of what tools are available in your computer geometry program. Then if you are having trouble with a construction, you may decide that it would be a good idea to learn exactly how a couple of likely commands might be used.

In **Geometer**, the more common construction tools are listed in the menu to the right of the drawing area, organized by the type of primitive they create—at the top, the point and line creation commands, and so on. This list is not complete, however. In the *Primitives* pull-down menu are a lot of others, all of which are available via the graphical user interface.

If you are willing to write **Geometer** textual descriptions directly, there are still more commands available, but most people will not need to use these. The textual descriptions include quite powerful calculation tools that will allow you to draw almost anything.

3.3 The Structure of a Geometer Diagram

Think of a **Geometer** diagram as a list of commands, each of which is executed exactly once beginning with the first and continuing to the end of the list. There is no looping

and branching can occur only in a very limited sense. The *vast* majority of diagrams consist of a linear list of commands that are issued one after the other[1].

When you view a diagram on your computer screen, you can only move points, and only certain "free" points at that. There are completely free points that you can move anywhere, and there are partially constrained points that you can move, but they must remain on a line or a circle or a conic section. Other than these free points, everything else is constrained by other geometric objects.

A line that is constrained to pass through two points will change if you move either of the two points upon which it depends. Those points can change either because they were free and you moved them directly with a mouse, or because they were constrained by other objects that moved. For example, if a line is constrained to pass through a free point and you move that point, the line will move, and so will the point defined to be the intersection of that line with a circle. So if another line passes through that line-circle intersection, it too will move, and so on.

Sometimes a constraint fails to hold. For example, if a point is defined to be at the intersection of a line and a circle and the line moves so that it no longer intersects the circle, the intersection point becomes undefined, and it, as well as any other objects that depend on it, are simply not drawn.

Remember that there are hundreds of examples of **Geometer** diagrams on the CD and in addition to examining any of them with the **Geometer** program, you can also look at them with a text editor—either the one contained within **Geometer** itself or with your own favorite. The **Geometer** diagrams are pure text files.

3.4 ♦ Interpreting Geometer Files

For complete details, see the **Geometer** reference manual on the CD, but we will examine here a small **Geometer** program and see how to interpret it. We will use as an example the classical construction of the circumcenter discussed earlier and displayed in figure 3.1, above. Here are the text file contents:

```
.geometry "version 0.60";
v1 = .free(-0.210169, -0.634407, "A");
v2 = .free(-0.39322, 0.484237, "B");
v3 = .free(0.183051, -0.18017, "C");
l1 = .l.vv(v1, v2);
l2 = .l.vv(v2, v3);
l3 = .l.vv(v3, v1);
c1 = .c.vv(v1, v2);
c2 = .c.vv(v2, v1);
v4 = .v.cc(c1, c2, 2, "D");
v5 = .v.cc(c1, c2, 1, "E");
c3 = .c.vv(v3, v2);
c4 = .c.vv(v2, v3);
```

---

[1]In practice, **Geometer** does not need to execute all the commands every time. Many operations leave most of the diagram unaltered, and **Geometer** can usually detect this and avoid a lot of extra calculation. But the net effect is as if it had executed every instruction from beginning to end.

```
v6 = .v.cc(c3, c4, 2, "F");
v7 = .v.cc(c3, c4, 1, "G");
l4 = .l.vv(v4, v5);
l5 = .l.vv(v7, v6);
v8 = .v.ll(l4, l5, "O");
```

The first line does nothing for the drawing but is used by **Geometer** for version control. Files missing this line will work without it.

The next three lines define three completely free points whose internal names are `v1`, `v2` and `v3`. The `.free` signifies the type of object, a free point in this case. The current x- and y-coordinates are given by the next two numbers, and the names in quotation marks are the names that appear on the screen next to those points. If there is no name in quotes, the point will appear without a name, although of course it retains the internal **Geometer** name.

All the words in the file that begin with a period, like `.free`, `.l.vv`, et cetera, are reserved **Geometer** commands; the others are variables whose names you or **Geometer** can create.

The next three lines of code draw three lines: the first connecting `v1` with `v2` and so on. The form of the command, `.l.vv` tells you a lot about what it means. The first "`l`" means "line", which is the type of object being created. The "`l`" is followed by another period, and then a list of the types of objects it depends upon, in this case, the line depends upon the positions of two points (called "vertices" in computer graphics lingo—hence the "`vv`"). Similarly, "`.c.vv`" indicates a command that makes a circle from two points (vertices)—the center and a point on the edge, and "`.v.cc`" makes a point at the intersection of the circles. Of course there are up to two intersections of a pair of circles, so the parameter "1" or "2" tells **Geometer** which one to use.

You may find it very interesting to draw diagrams using the graphical user interface and then look at the generated text form of the code. You can do this with the "Edit Geometry" command in the "Edit" pulldown menu. Try changing colors of objects, changing the line types, angle types, et cetera, and see how the code is affected. You can also edit code while you are in the text editor, although it is quite easy, at first, to mess things up pretty badly, and **Geometer** does not have a very robust error checker.

Note: If you are trying to edit the code and get completely messed up, the easiest thing to do to get out of the mess is to select all the code in the file and delete it. Then save the file, and **Geometer** is perfectly happy to read a completely empty text file and continue.

3.5 Simple Classical Examples

The best way to learn to do constructions with **Geometer** or any similar program is to examine a series of examples, beginning with easy ones and progressing to ones that are more complex.

In this first section we will show how the most basic commands can be used to build up common constructions: the midpoint of a segment, a perpendicular to a line from a point, a tangent to a circle from an exterior point, and so on.

Although almost all the examples in this section are available directly as **Geometer** commands, all can be done as a series of smaller steps with only the most primitive straightedge and compass operations.

Even if you plan to use **Geometer** for your work and the operations below are available as primitive **Geometer** commands, it is useful to read through these examples as they demonstrate how more complex operations can be built up from primitive ones. In most computer geometry programs (including **Geometer**) you can define "macros" which are combinations of simple commands that can be applied as a unit. Thus if you compose a macro to find the common exterior tangent lines to a pair of circles, you can use that macro over and over again, whenever you need to do that particular construction as part of a more complex problem.

Note that in order not to clutter up the illustrations too much, we often draw only an arc of the required circle in the sections that follow, because we only care where the circle intersects some other line or circle. In every case, we could have drawn the whole circle, but it is easier to follow the steps with less junk in the drawing.

In the first few examples, the **Geometer** commands that perform the given steps are indicated in the text. For example, the command to construct a line connecting two points is labeled as *PP=>L* on the **Geometer** control panel.

3.5.1 Bisect a Line Segment

Given a line segment AB, construct M, the midpoint of that segment.

If you use **Geometer** you can load the file `Construction/Segbisect.T` to follow the construction more easily on your computer screen. The caption under the figure tells you where to find the file: it is in the subdirectory `Construction` of the directory `Book` which is where all the **Geometer** diagrams used in this chapter are found. You can learn about the file structure on the CD by using the *Documentation* entry in the *Help* pulldown menu.
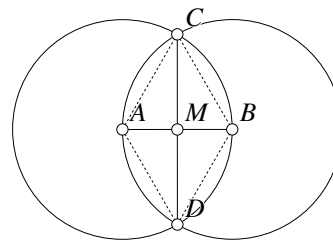
Figure 3.2: Bisect a Segment
`Construction/Segbisect.T` [P]

To step through the construction in **Geometer**, press the *Next* button (or type the **n** key on the keyboard) to see the next step. The *Prev* button (or the **p** key) goes back one step. (Remember that if the text "[P]" appears after the file name under a figure, it means the figure is a proof or construction, and that the *Next*, *Prev*, and *Start* commands will have an effect.)

The method is shown in figure 3.2. First construct circles centered at A and B, both having radius AB using *CtrEdg=>C*. We can do this because we have a point at the center, and a point on the radius of each of those circles. Those two circles will meet at points C and D, which we can identify because we can find the points at the

intersections of two circles (command: *CC=>P*). Connect C and D by a straight line (command: *PP=>L*), and that line will intersect the original segment AB at a point M in the center (command: *LL=>P*).

In a sense, we are not done; we must prove that the construction in fact bisects the line segment, so for completeness here is a quick proof (actually, two quick proofs):

Since both circles have the same radius, $AC = BC = AD = BD$. Therefore quadrilateral $ACBD$ is a rhombus, and therefore a parallelogram, since a rhombus is a special type of parallelogram. The diagonals of a parallelogram (AB and CD in this case) bisect each other so M is the midpoint of AB.

Alternatively, we know that the diagonals of a rhombus are perpendicular so all the angles at M are right angles. We also know that $AC = BC$ (the circles have equal radii), and since CM is equal to itself by hypotenuse-leg we know that $\triangle AMC \cong \triangle BMC$. Therefore $AM = MB$ so M is the midpoint of the segment AB.

3.5.2 Copy a Segment

Copy a segment AB to an equal-length segment on the line L. Set the compass to the length of AB. Place a mark P on the line L where the new segment is to appear (command: *P on L*), and sweep a circle using P as center and with the compass setting as above. In **Geometer**, the command to create a circle with a center and a radius defined by the distance between two other points is

Figure 3.3: Copy a Segment
`Construction/Copyseg.T` [P]

Ctr PP=>C. If Q is one of the points where this new circle crosses L (command: *LC=>P*) then the segment PQ has the same length as AB. See figure 3.3

3.5.3 Copy an Angle

Copy an angle $\angle ABC$ so that the vertex of the new angle lies on a line L. See figure 3.4. First pick a point P on L (*P on L*). Draw a circle centered at B on the original angle having length AB (command: *CtrEdg*) and find the point N where it intersects the other edge of the angle (command: *LC=>P*). Sweep another circle with radius AB centered at P on L (command: *Ctr PP=>C*), and let that circle intersect L at O. Now construct a circle centered at O of length AN, and the point R where this circle intersects the larger circle centered at P (command: *LC=>P*) forms an exact copy of the arc AN. Connect R to P (command: *PP=>L*) and we are done.

To show that the new angle at P is the same as the old angle at B, note that the two triangles $\triangle ABN$ and $\triangle OPR$ have all three sides equal, and are hence congruent by SSS. Thus the corresponding angles ($\angle ABN$ and $\angle OPR$) are equal.

3.5.4 Bisect an Angle
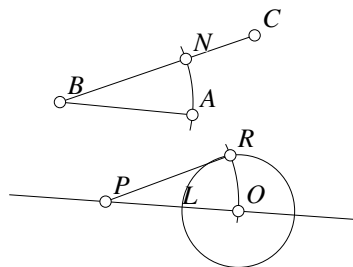
Figure 3.4: Copy an Angle
Construction/Copyangle.T [P]

Bisect an angle. See figure 3.5. To bisect ∠ABC, pick a point P on AB, draw a circle with center B through P, and find Q, where it intersects BC. Now, from points P and Q, draw circles with lengths both equal to BP which will intersect at W. The line BW bisects ∠ABC.

To show that this line does in fact bisect the angle, note that the two triangles △BQW and △BPW are congruent, since BP = BQ, PW = QW, and BW is equal to itself. Thus all three sides are equal to the corresponding sides, so △BQW ≅ △BPW by SSS. The corresponding angles are equal: ∠PBW = ∠QBW, so the angle ∠PBQ is bisected by BW.

3.5.5 Perpendicular at a Point

Construct a line perpendicular to a given line through a point on that line. See figure 3.6. To construct a line perpendicular to L passing through P, first select another point Q on L. Draw a circle centered at P passing through Q and let R be the other intersection of this circle with L. Next, sweep out two larger arcs, say of length RQ, centered at points R and Q. These arcs intersect at M, and the line PM is the required line perpendicular to L and passing through P.

Figure 3.5: Bisect an Angle
Construction/AngleBisect.T [P]

In the figure, △MPR ≅ △MPQ by SSS. This is because MP is equal to itself, PQ = PR since they are on the same circle centered at P, and RM = QM since both were drawn with an equal compass setting (in this case, equal to RQ). Since △MPR ≅ △MPQ, ∠MPR = ∠MPQ. Since together those angles make 180°, they must each be 90°, or right angles. Thus the line PM is perpendicular to L.
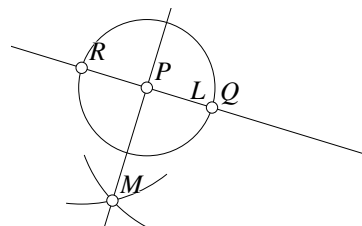
Figure 3.6: Perpendicular to a Line
Construction/PerpPtOnLine.T [P]

3.5.6 Drop Perpendicular to Line

Construct a line perpendicular to a given line through a point that is not on the line. There is no figure for this one, because in

Construction 3.5.5, figure 3.6 works perfectly well if the point P is not on L. The proof depends upon the fact that P is on the line, but can be modified to be correct. If you have doubts, do the construction yourself in **Geometer** following the written directions for the previous construction and see how it works. Be sure to move P around, both near and far from the line.

3.5.7 Parallel through a Point

Given a line L and a point P not on L, construct a line through P that is parallel to L.

A figure is not really required—do the previous construction in Section 3.5.6 twice. The first time, construct a new line L' perpendicular to L through P, and then construct another line through P, but perpendicular to L'. This second line will be perpendicular to a perpendicular to L, and will hence be parallel to L, and it was constructed to pass through P.

3.5.8 Sum and Difference of Angles
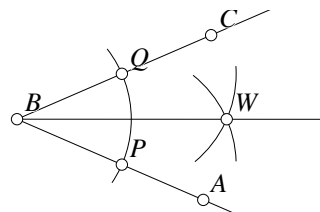
To get the sum or difference of angles is easy, assuming you know how to copy an angle (Construction 3.5.3). Lay out one of the angles, and then put the other one next to it, either inside or outside, to make the sum or difference.

3.5.9 Triangle Given Three Sides

Construct a triangle given the lengths of the three sides. This obviously cannot be done if one segment is longer than or equal to the sum of the other two, but assuming that is not the case, the construction is not too hard. See figure 3.7.

Copy one of the lengths (say AB) to a line having new endpoints A' and B'. Then draw a circle of radius AC about A' and another circle of radius BC about B'. Call the intersection C'. Triangle $\triangle A'B'C'$ is the required construction. Because each of the sides of the new triangle is an exact copy of the given lengths, it is clear that $\triangle A'B'C'$ satisfies the conditions.
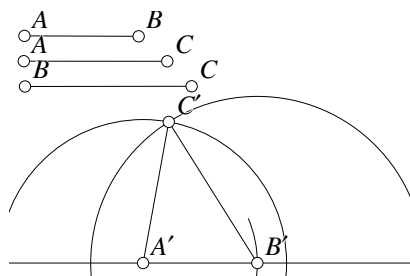
Figure 3.7: Triangle from the Lengths of its Sides

    `Construction/Triangle.T [P]`

If you are using a computer geometry program, be sure to change the three given lengths and see how the construction is modified. Also change them so that they do not specify a valid triangle (for example, make AB and BC small, and AC large) and see what happens.

3.5.10   Perpendicular Bisector of Segment

Construct the perpendicular bisector of a segment. This is exactly the same as as the construction in Section 3.5.1, except the result we want is the line, not the point where the lines intersect.

3.5.11   Circle through Three Points

Given three points, A, B, and C, construct the circle that passes through all of them. (See figure 3.8.)

    Using the construction from Section 3.5.10 twice, draw the perpendicular bisectors of both segment AB and segment BC. Those bisectors will meet at a point O which can then be used as the center of the required circle.

    The reason this works is that every point on the perpendicular bisector of a segment is equidistant from the two endpoints of the segment. So the point O is the same distance from A and B, since it is on the perpendicular bisec-

Figure 3.8: Circle Passing through Three Points

`Construction/CircFrom3Pts.T [P]`

tor of AB, and similarly, it is the same distance from B and C. Thus $AO = BO$ and $BO = CO$, so $AO = BO = CO$, so O must be the center of a circle passing through A, B, and C.

3.5.12   Find a Circle's Center

Given a circle, find its center. We basically did this (and a little more) in Construction 3.5.11. Simply pick three different points on the given circle and do the construction above. Any three different points will work, but if you are doing this with a physical straightedge and compass, it is a good idea to pick them far apart on the circle.
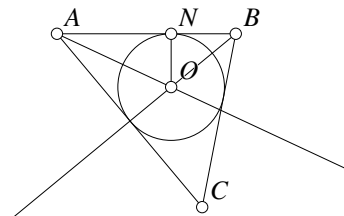
Figure 3.9: Circle Inscribed in a Triangle

`Construction/Incenter.T [P]`

3.5.13   Inscribed and Circumscribed Circles

Given a triangle $\triangle ABC$, construct both the inscribed and circumscribed circles.

    The circumcircle is the circle passing through the three points that are the triangle's vertices. See figure 3.9 for the construction of the inscribed circle (called the "incircle").

To construct the incircle, construct the angle bisectors at angles A and B. These will meet at a point O inside the triangle. From O, drop a perpendicular to line AB, intersecting line AB at point N. The required circle has center at O and passes through N.

Every point on an angle bisector is equidistant from the two sides of the angle, so the point O will be equidistant from lines AB, BC, and CA. To find that distance, we drop a perpendicular from O through any of the sides and find where it intersects that side. This will be at the shortest distance from O to the side. Since O is this same distance from all three sides, it is the center of the required circle, and the distance from O to a side is the radius of the required circle.

3.5.14 Construct Specific Angles
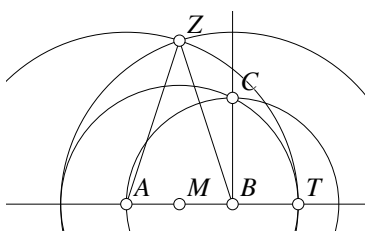
Figure 3.10: Construct a 72° Angle
Construction/Seventytwo.T [P]

Construct angles of $90°$, $45°$, $30°$, $60°$, and if you want a challenge, $72°$.

The first four constructions are quite simple. Any time you have constructed a perpendicular to a line you have constructed a $90°$ angle. If you construct a right angle and bisect it, you will have a $45°$ angle. To make a $60°$ angle, just construct an equilateral triangle (a triangle with all three sides equal, which you know how to do since you can construct a circle given the lengths of all three sides), and all three of its angles will be $60°$. If you then bisect your $60°$ angle, you will have one that is $30°$.

The construction of the $72°$ angle is shown in figure 3.10. Begin by laying off a length AB on some line. Construct a perpendicular to the line at B, and a compass centered at B and passing through A will meet the new perpendicular at C. Find M, the midpoint of AB, and draw a circle centered at M that passes through C. This circle will intersect the line AB at T. Using the length AT, make two circles centered at A and B which intersect at a point Z. The isosceles triangle $\triangle AZB$ has two angles measuring $72°$ and one measuring $36°$.
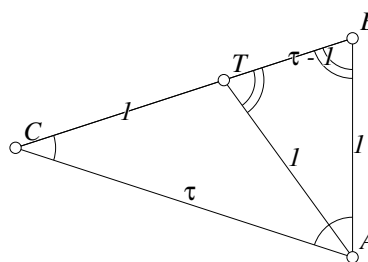
Figure 3.11: The Golden Ratio
Construction/Golden.D [D]

To understand the construction, let us first look at the properties a triangle with angles of $72°$, $72°$, and $36°$ must satisfy. See figure 3.11. In that figure, angles marked with the double arcs are $72°$ and those with a single arc, $36°$. If the line AT bisects the $72°$ angle at A, it will form two $36°$ angles, as shown.

If we let the side AB measure 1 unit, then AT and CT will also measure 1 unit because they are parts of isosceles triangles $\triangle ABT$ and $\triangle CAT$. The unknown length

is AC, so let us call that τ. Then the length of BT is $\tau - 1$. By AA, $\triangle ABT \sim \triangle CAB$, so we have: $\tau/1 = 1/(\tau - 1)$. Multiply this out to get $\tau^2 - \tau - 1 = 0$, whose solution with $\tau > 1$ is $\tau = (\sqrt{5} + 1)/2$.

What we have done is to construct τ, the golden ratio. An isosceles triangle with two sides equal to τ and the third equal to 1 will have two 72° angles in it. The value of τ is given by the formula $\tau = (1 + \sqrt{5})/2$. To see that the length AT is τ (assuming that AB has length 1), note that MB has length $1/2$, so by the pythagorean theorem, the length of MC (which is equal to MT) is $\sqrt{5}/2$. AT is the same as MT with $1/2$ added on, so $AT = \tau$.

Note: Since we know how to add and subtract angles and to bisect them, there are actually quite a few interesting angles that can be constructed from the ones above. For example, if we subtract a 30° angle from a 36° angle, we can get a 6° angle which can be bisected to give us 3°. We can then add together any number of 3° angles to get 3°, 6°, 9°, 12°, ….

These are the only angles that can be constructed that are an integral number of degrees. We know that because if it were possible to get any other angle than a multiple of 3° then we could subtract the nearest multiple of 3° to get either 2° or 1°. We could then add together either 10 or 20 of those to construct a 20° angle which we will show to be impossible in Section 3.12.

3.5.15   Construct a Regular Pentagon

Construct a regular pentagon. Assuming that you can do the construction above to make $\triangle ABC$ with two 72° and one 36° angle, you are almost done, since the exterior angles of a regular pentagon are all equal to 72°. See figure 3.12.

Figure 3.12: Construct a Regular Pentagon

`Construction/Pentagon.T [P]`

Begin with the triangle constructed in the previous section and sweep out circles centered at points A, B, and C, all having a radius equal to AB. Let M and N be the points where these circles intersect outside $\triangle ABC$, and connect them to vertices A, B, and C, as appropriate.

Since all the sides of this figure are equal in length to AB, all we need to do is show that all the angles are equal as well. In a regular pentagon, all the internal angles measure 108°, and we know that a triangle whose sides are 1, 1, and τ has angles of 36°, 36°, and 108°, so it is easy to add the angles together in the figure and check that all of the internal angles of the polygon $ABNCM$ in figure 3.12 are equal to 108°.

3.5.16   Tangent to Circle through Point

Given a circle C and point P on that circle, construct a line through P that is tangent to C.
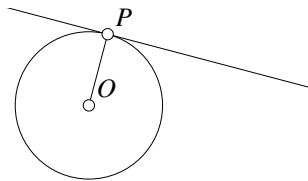
Figure 3.13: Tangent to Circle at a Point on the Circle

`Construction/Circtan.T [P]`

This one is not hard. Find the center O of circle C as we did in the construction described in Section 3.5.12. Next, construct the line OP. The line perpendicular to OP passing through point P is the required tangent line because the radius of the circle is perpendicular to the tangent line where it intersects the circle.

Obviously if you do the complete construction described above, the figure would be far more cluttered than what is illustrated in figure 3.13.

3.5.17 Tangent to Circle II

Given a circle C and a point P not on C, construct a line through P tangent to C. See figure 3.14.

First find O, the center of the circle. Then find M, the midpoint between O and P. Construct a circle centered at M and passing through O (and P too, of course). The intersections T_1 and T_2 of the two circles are the points of tangency, so the required lines are PT_1 and PT_2.

This works because the circle centered at M will have OP as its diameter. For any point T on that circle, $\angle PTO$ must be a right angle. At the particular points T_1 and T_2, the right angle will coincide with the boundary of the original circle C, so OT_1 and OT_2 are radii, and hence PT_1 and PT_2 are tangent lines.

3.5.18 The Geometric Mean

Given two segments with lengths l and m, construct a segment whose length k is the geometric mean of l and m. In other words, so that $k^2 = l \cdot m$.

One way to find the geometric mean of two segments is based on the observation that the altitude to the hypotenuse of a right triangle is the geometric mean of the lengths into which the hypotenuse is divided. In figure 3.15, CP is the geometric mean of AP and BP. All three of the triangles in the figure are similar ($\triangle ACP \sim \triangle CBP \sim \triangle ABC$), so $AP/CP = CP/PB$ which means $AP \cdot BP = CP^2$.

Figure 3.14: Tangent to a Circle from an Exterior Point

`Construction/CircPtTan.T [P]`

To construct the geometric mean, draw an arbitrary line and choose a point P on it. Next, pick points A and B such that $PA = l$ and $PB = m$. Find the midpoint of the segment AB and use it as the center of a circle passing through A and B. Thus AB is

the diameter, and any point C on the boundary will form a right angle $\angle ACB$. Draw a perpendicular to AB through P that intersects the circle at point C If k is the length of PC, then:

$$k^2 = l \cdot m.$$

3.5.19 Divide a Segment into n Equal Parts
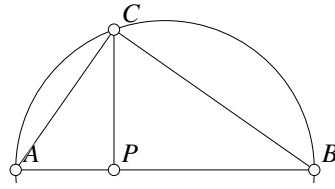
Figure 3.15: The Geometric Mean of Two Segments

Construction/Geomean.T [M]

The problem calls for a subdivision of a line segment into n equal pieces, but the following subdivision into 5 equal pieces can be modified easily to work for any n. In figure 3.16 $n = 5$ and we will divide the line segment AB into 5 equal segments.

First pick a point C not on the line, and construct line AC. Now pick a point X_1 near A, and draw a circle around it passing through A. That circle will also hit in X_2, and clearly, $AX_1 = X_1X_2$. Continue in this manner, making a circle about X_2 which passes through X_1 and its intersection with line AC gives a new point X_3. We can do this as many times as we want to get an evenly spaced sequence of points $X_1, X_2, \ldots, X_n$ along the line. In our case we can stop at X_5 since we are subdividing into 5 parts.

Draw the segment X_5B, and construct lines parallel to X_5B passing through X_4, X_3, X_2, and X_1. These lines will intersect the segment AB at evenly spaced points P_4, P_3, P_2, and P_1, which achieves the desired subdivision. The distances between the P_i are equal because the parallel lines are evenly spaced.
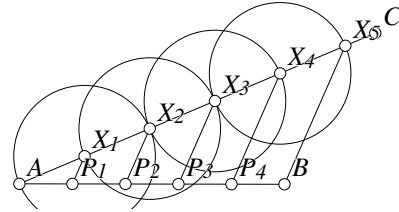
Figure 3.16: Dividing a segment into 5 equal pieces

3.5.20 Divide a Line in a Given Ratio Construction/Subdivide.T [P]

If you understood Construction 3.5.19 that divides a line into equal parts, it should be obvious how to divide it into a ratio. Use exactly the same construction as above, but when you mark off segments on AC, just mark off two of them, and make them the relative lengths of the ratios. To prove that your result is correct, look at the two similar triangles formed.

We will use this trick in the construction of the common internal tangents to two circles in Section 3.6.1.

3.6 Intermediate Examples

3.6.1 ♦ Common Tangents to Two Circles
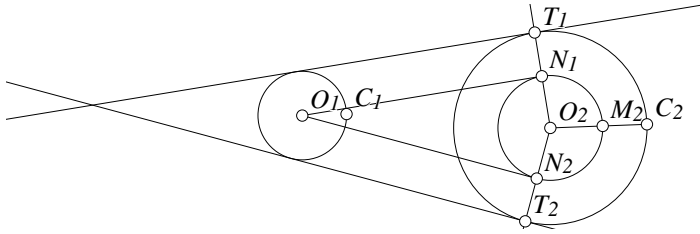
Figure 3.17: External tangents of two circles
Construction/ExtTangent.T [P]

Given two circles C_1 and C_2, find the external and internal tangent lines to both of them. These are really two different problems, so let us begin by finding the common external tangents. (See figure 3.17.)

Assume that circle C_1 is smaller than C_2, otherwise, reverse the names. Find the centers O_1 and O_2 of the circles, and then subtract the radii such that $M_2C_2 = O_1C_1$. Next, construct a circle concentric with C_2, and passing through M_2. Find the tangents to this smaller circle passing through O_1, the center of C_1 using the technique of construction 3.5.17. The common external tangents will be parallel to these tangents, but pushed out by the radius of C_1. To do this, construct the lines through O_2 and the points of tangency with the circle of radius O_2M_2, and add the radius of C_1 beyond the intersection, giving points T_1 and T_2. The parallels to the tangents O_1N_1 and O_1N_2 passing through T_1 and T_2 are the required common external tangents.

To see why this is true, notice that the common external tangents are perpendicular to the radii of both circles. Thus, the distances along the radii of T_1N_1 and T_2N_2 are both equal to the radius of C_1.

We will use a different trick to find the common internal tangents. See figure 3.18.

The idea is this: if we look at the solutions in the figure, we see that both the common internal tangents pass through a point X somewhere on the line between the centers of the circles. In fact, if we look at $\triangle O_1T_1X$ and $\triangle O_2T_2X$, they are both right triangles, and $\angle O_2XT_2 = \angle O_1XT_1$ since they are vertical angles. That means $\triangle O_1T_1X \sim \triangle O_2T_2X$ when X is at the correct point. The similar triangles will have a ratio of lengths equal to the ratio of the diameters of the two circles, so if we can divide the line O_1O_2 in that ratio at X, that X should be the point we are looking for. Here is how to do that subdivision, and the proof that the final construction is correct:

Find the centers of the circles and connect them with a line. Then pick an arbitrary point E not on the line O_1O_2 and connect it to O_1. O_1E intersects C_1 at P. By drawing a circle centered at P of radius O_2C_2, we can find a point Q on O_1E such that $PQ = O_2C_2$. Connect Q with O_2 and construct a line parallel to that line passing through P. The parallel line through P strikes O_1O_2 at X. Finally, using the technique

of Construction 3.5.17, construct tangents to C_1 passing through X. These lines will be the common tangents of C_1 and C_2.
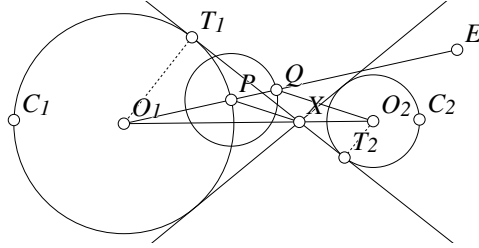
Figure 3.18: Internal Tangents of Two Circles
`Construction/IntTangent.T [P]`

The reason this works is that for the required internal tangent, triangles $\triangle O_1 T_1 X$ and $\triangle O_2 T_2 X$ are similar, and their sizes are in the ratio of the radii of C_1 and C_2. That is because $O_1 T_1$ and $O_2 T_2$ *are* radii of the circles. Therefore the point X divides $O_1 O_2$ in that ratio. By picking points P and Q as we did, the segments $O_1 P$ and PQ are in the ratio, and the parallel lines through P and Q must cut $O_1 O_2$ in exactly the same ratio. Therefore a line through X that is tangent to one circle will be tangent to the other.
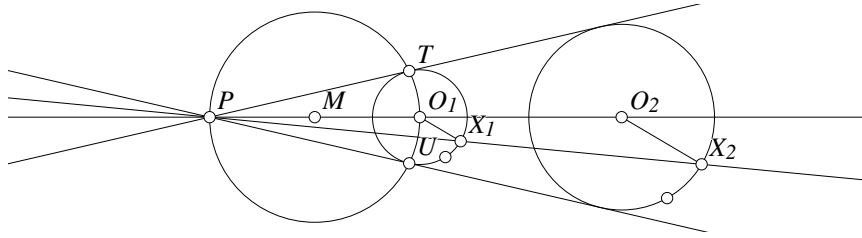
Figure 3.19: External tangents by projection
`Construction/ExtTangent1.T [P]`

Figure 3.19 illustrates that we can also adapt this last method to find the common external tangents of two circles based on the idea of a projection.

Assume the two circles are centered at points O_1 and O_2, as in the figure. A point X_1 is chosen on the circle centered at O_1 at any point except on the $O_1 O_2$ axis. A line parallel to $O_1 X_1$ is drawn through O_2 and it intersects the other circle at a point X_2 in the same direction from O_2 as X_1 is from O_1.

If the lines $O_1 O_2$ and $X_1 X_2$ are drawn, they will intersect at a point P that effectively projects the corresponding points of one circle onto the other, including the points of tangency. The tangents to either circle through P will be tangent to the other circle. In the figure, the tangents through P to the circle centered at O_1 at T and U are determined, and those are the required lines.

It is not difficult to show that this works. Begin by showing that $\triangle PO_1 X_1 \sim PO_2 X_2$. Next construct a triangle with the same similarity ratio to $\triangle PO_1 T$ and so on.

3.6.2 Equilateral Locus of Points

Let A, B an C be three distinct points in the plane. Find the locus of the centers of equilateral triangles that can be drawn having one edge passing through A, another

through B and the third through C.

When problems like this occur in Euclidean geometry texts, it is likely that the solution will be something simple: a line or a circle, or perhaps some conic section like an ellipse or parabola. But even if you know that it is a circle, which circle is it? The easiest way to proceed is to draw a **Geometer** diagram and see which points are generated.

We will not solve this problem here, but will demonstrate how to do the construction, and how to modify the construction to gain a lot more clues about how to approach the problem.

To construct the **Geometer** diagram, the first step is easy: create three free points A, B and C on the plane. Since one of the edges of any equilateral triangle of interest will pass through the point A, add another free point X and draw the line AX. Since we don't know exactly where A will sit on that line in the final diagram, change the line type from "segment" to "line". To construct the locus, we will simply drag the point X around A and thus sweep out all possible lines.

To construct the other two lines of the equilateral triangle is not difficult: we know that they will make $60°$ angles with the line AX and there are two possible ways to do this. The easiest way is to construct an equilateral triangle using AX as one of the segments and the other two sides will make the two possible lines at $60°$ to AX. It's easy to draw an equilateral triangle; draw the circle centered at A passing through X and then draw the circle centered at X passing through A. Either intersection of these circles is a suitable third vertex Y that will form an equilateral triangle $\triangle AXY$. (Note that either choice of circle-circle intersections will generate a pair of additional lines that indicate the directions that any line at $60°$ will to AX.)

Figure 3.20: Equilateral Locus
Construction/EquLocus.T [M]

Now we simply need to construct lines through B and C that are parallel to the edges AY and XY of $\triangle AXY$, but there are two ways to do this since we can either draw the parallel to AY through B or C and then the parallel to XY through the other. If we try both pairings, we find that each of them generates an equilateral triangle, so there will be two centers associated with each possible line AX.

Label the intersections of the lines to form two triangles: $\triangle PQR$ and $\triangle P'Q'R'$ and all that needs to be done is to find the centers of those triangles. The best way to locate the centers is probably to find the intersections of a pair of medians for each or of a pair of altitudes for each.[2] A pair of angle bisectors may give bogus results, since every pair of crossing lines has *two* angle bisectors and we need to assure that

---

[2]With **Geometer** it may be easiest simply to construct the inscribed circle and locate its center.

the one chosen is the one running through the triangle. Even if you choose correctly while drawing the figure, motion of the point X may change this choice, and the other bisector may result, yielding a bogus triangle center.

At this point, it's probably a good idea to change the colors of all the construction lines to the invisible color, and if you're going to present the drawing to others, you may want to erase the infinte lines connecting the vertices of triangles $\triangle PQR$ and $\triangle P'Q'R'$ (by making them the invisible color) and redraw them as the segments PQ, QR, et cetera.

Finally, set the colors of the centers of the triangles to be the smearing color and drag point X around point A to see what results. The answer appears to be two concentric circles. It is often a good idea to change the point type to "dot" for locus problems like this since it is easier to see exactly what shape is swept out by a dot than by one of the other symbols that indicate points. See figure 3.20.

Notice that in the figure, the points are uniformly drawn around the circles. This is because the figure was generated from a script that is a bit more complex than the construction described above. The files `Construction/EquLocus.T` (the original version) and `Construction/EquLocus1.T` (the modified version) illustrate how this is done.

Now that we firmly believe that the locus is a pair of concentric circles, the same figure can be used to find clues about which circles they are. One interesting thing to do is to change the colors of points P, Q and R (or of P', Q' and R') to be the smearing color and see what shapes they sweep out. Try this. Now four circles seem to intersect at a point. What point is it? See Section 7.4 for more clues.

3.6.3   Three Tangent Circles
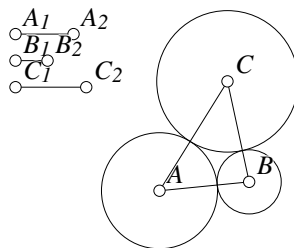
Figure 3.21: Three Tangent Circles
`Construction/TangentCircs.T` [M]

Given three different lengths, construct three mutually tangent circles having those lengths as their radii. See figure 3.21.

Once you see the trick, this problem is not too difficult. Notice that if the radii are a, b and c, then the distances between the centers of the circles will be $a+b$, $b+c$ and $c+a$. Thus to do the construction, all we need to do is to construct line segments having those three lengths and then construct a triangle having those segments as the lengths of its sides. The vertices of the triangles will be the centers of the required circles.

This can be done on a single line. Pick a point on that line and draw circles of radius a and b about it. On opposite sides of the point pick the intersections of those two circles. These points will be a distance of $a+b$ apart. The same line can be used to compute the distance $a+c$ from what will be the center A of the circle of radius a and

the distance $b + c$ from the center of what will be the center B of the circle of radius b. The intersection of the circle with center A and radius $a + c$ and the circle with center B and radius $b + c$ will be the center C of the circle with radius c.

3.6.4 Circle Tangent to Four Circles
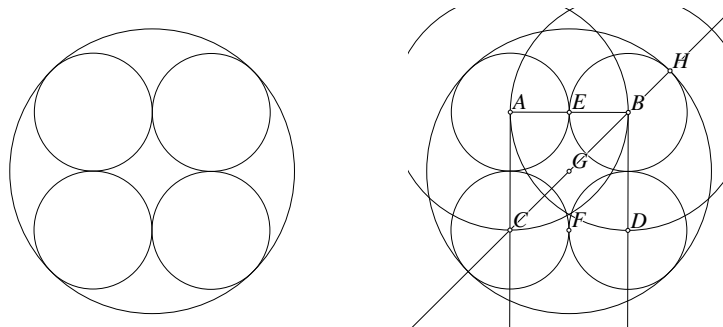
Figure 3.22: Circle Tangent to Four Circles
`Construction/Fourcircs.T` [M]

Suppose you would like to construct a diagram like that displayed on the left in figure 3.22. The four inscribed circles are all of equal size.

To perform the construction (see the diagram on the right of figure 3.22), we note that the centers of the four circles lie on the corners of a square. Those circles will have radii equal to half the length of the side of the square, and the center of the surrounding circle will be at the center of the square. With this information, the construction is fairly straight-forward.
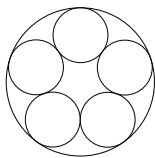
Figure 3.23: Circle Tangent to Five Circles

`Construction/FiveCircs.T` [M]

First we make a square by drawing a segment AB that will be one side of the square. Then construct perpendicular lines to the segment AB at both A and B. We then draw circles about A and B with radii AB, and these circles will intersect the perpendicular lines at points C and D that will be the other two corners of the desired square.

Let E and F be the midpoints of AB and CD and then the four circles centered at A, B, C and D passing through E or F will be the required central mutually-tangent circles. The center of the square is the midpoint of BC (or of AD or EF). If we construct a line through B and C it will intersect the small circle centered at B at H—a point on the desired surrounding circle. Draw that circle and we are done.

To create the final figure that appears on the left side of figure 3.22 we simply set the colors of all the extra construction lines to the invisible color.

How would you draw the diagram in figure 3.23? Hint: remember the construction of the regular pentagon in Section 3.5.15.

3.6.5 Tangent Circle Example

Suppose you would like to draw a diagram where you need to construct a circle that is simultaneously tangent a line and passes through two given points. In figure 3.24 we are given the positions of points A and B as well as the line XY and we would like to construct the two circles illustrated.
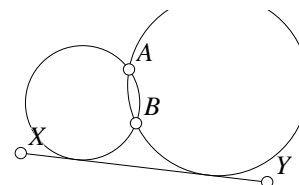
Figure 3.24: Circle Passing through Two Points and Tangent to a Line
Construction/PPL.T [M]
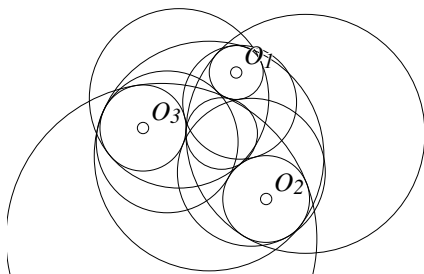
Figure 3.25: Eight Circles Mutually Tangent to Three Given Circles
Construction/Apollonius.T [M]

Note: This is a special case of Apollonius' problem that requires the construction of a circle passing through points or tangent to lines and circles. The problem can be stated and solved for any combination of three points, lines and circles, although there may be multiple solutions. For the problem of finding all circles that are mutually tangent to three given circles, for example, there may be up to eight different solutions. (See figure 3.25.) If we list the three given objects with a P for "point", L for "line" and C for "circle", there are ten different versions of Apollonius' problem: (PPP), (PPL), (PPC), (PLL), (PLC), (PCC), (LLL), (LLC), (LCC) and (CCC). This particular problem is (PPL). The problem (PPP) is equivalent to finding the circumcenter of a triangle, and the problem (LLL) is equivalent to finding a triangle's incenter.

This problem probably seems a bit difficult at first, especially since we do not know how to draw a **Geometer** diagram to experiment with it. (The reader, of course, can manipulate the figure to try to learn something about it.)

But even without a **Geometer** diagram that does the right thing with A, B, X and Y free, you can at least make an exact drawing by making two circles that intersect and then drawing a common external tangent to them. If you do this and label the tangent points as I and J and let P be the intersection of AB with XY, when you test[3] the "fake" **Geometer** diagram in figure 3.26, you will find the following:

```
Ratios:
    [I P] x [J P] = [A P] x [B P]
    [J P] x [J P] = [A P] x [B P]
```

---

[3]See Chapter 2 to learn how to "test" a diagram.

These relationships are obvious, since one of the lines from P cuts a secant and the other is a tangent, so $PI^2 = PJ^2 = PA \cdot PB$. It is easy to find P, so PA and PB are available. If we construct PI to be the geometric mean of those segments as we did in Section 3.5.18 we can find I. The center of the circle will lie on a perpendicular to XY through I and on the perpendicular bisector of the segment AB.

Figure 3.27 shows the complete construction beginning only with A and B and the line XY. **Note:** You will almost certainly find the construction easier to follow if you load the **Geometer** diagram `Construction/LPP2.T` and follow along by pressing the **n** key.

Draw line AB that intersects XY at P. Draw a circle with center P through B and find the other intersection of that circle B' with line AB. Now $PB' = PB$ and PA is on the other side. Let M be the midpoint of $B'A$ and draw a circle through A with center M. This is the circle with AB' as diameter, so the intersection of the perpendicular to AB through P will intersect that circle at D and the length of PD will be the geometric mean of PA and $PB = PB'$.

Figure 3.26: Circle Passing through Two Points and Tangent to a Line
`Construction/PPL1.T [M]`
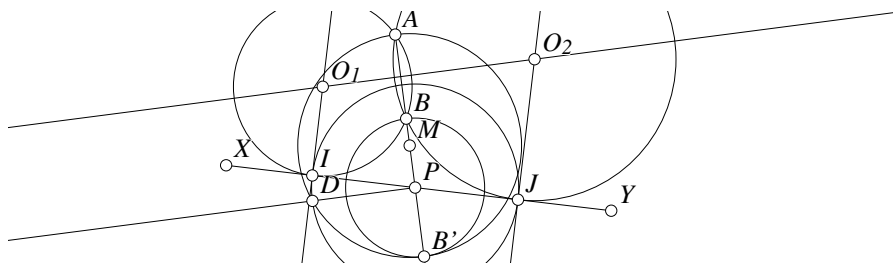
Figure 3.27: Circle through Two Points Tangent to a Line
`Construction/PPL2.T [P]`

A circle centered at P passing through D intersects the line XY at I and J. The centers of the required tangent circles will be on the intersections of lines perpendicular to XY passing though I and J with the perpendicular bisector of the segment AB. From these centers, O_1 and O_2, draw circles passing though A. This completes the construction.

3.6.6 | Another Apollonius Problem

Once we have solved the problem in Section 3.6.5, we can easily solve the (PLL) version of Apollonius' problem: Given a point and two lines, construct the circle that passes through the point and is at the same time tangent to the two lines. There are again two solutions for this problem.

There are two cases to consider. The easy one is when the two lines are parallel. In that case, the center of the circle must lie on a line that is parallel to the given lines and halfway between them. The radius of that circle will be half the distance between the parallel lines, so if we draw a circle about the point with that radius and find the intersections with the middle line, those intersections represent possible centers of the required circle. If the point is

Figure 3.28: Circle Passing through a Point and Tangent to Parallel Lines
`Construction/PLL.T [M]`

between the parallel lines there will be two solutions; if it is outside the parallel lines there will be no solutions, and if it happens to lie exactly on the line, there will be a single solution.
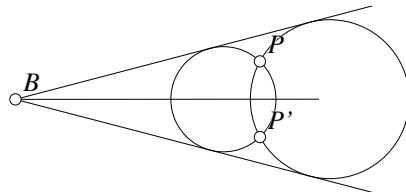
Figure 3.29: Circle Passing through a Point and Tangent to Parallel Lines
`Construction/PLL1.T [M]`

The construction for the case of parallel lines is illustrated in figure 3.28. A point A is selected on one of the lines and a perpendicular is drawn through A intersecting the other line at B. The midpoint M of segment AB is determined and a line is drawn through M parallel to the given lines. Next, a circle of radius AM is drawn about the given point P, and the intersections C_1 and C_2 of that circle with the line through M are the centers of the required circles. The required circles have radii of length AM.

If the lines are not parallel, the solution is actually easier to describe since we already learned how to solve the PPL problem in Section 3.6.5. This is because if the lines are not parallel, they will intersect to form an angle. The centers of the required circles will have to lie on the bisector of that angle, and it is obvious that the required circles passing through the given point P will also have to pass through the reflection P' of P across the angle bisector. (See figure 3.29.) Now we have the same problem that was solved in the previous section—to find a circle passing through two points and tangent to a line.

Obviously if the point P happens to lie on the angle bisector the problem is also easy to solve: The line through P perpendicular to the angle bisector will form a triangle and the two required circles are the incircle and one excircle of that triangle.

What happens if P lies on one of the lines?

3.7   ♦ Advanced Examples

The solutions to most of the examples in this section are specific to **Geometer**. If you use a different application, you will probably need to learn from the user's guide how

to solve these problems with that application. It may be useful to have the **Geometer** reference manual available as you read this section.

3.7.1 Trisect an Angle

The diagram in figure 3.30 shows what we want: $\angle ABC$ is trisected into three equal angles by BD and BE. In other words, $\angle CBD = \angle DBE = \angle EBA = \angle ABC/3$. If we adjust the size of $\angle ABC$ with the mouse, the sizes of the smaller angles will change accordingly.

Although this is impossible using classical methods (see Section 3.12), you can trisect an angle with **Geometer**, but you will have to edit the text version of the diagram.
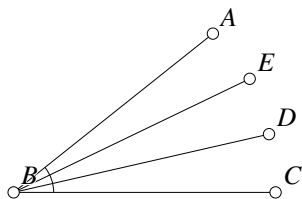
Figure 3.30: Trisect an Angle
Construction/GeomTrisect.T [P]

The easiest way to construct the figure is to use the standard **Geometer** tools to draw the points A, B and C, the line segments AB and AC and the angle $\angle CBA$. Then edit the diagram, and the first seven lines of the listing below will be created by **Geometer**.

At this point you need to add the next four lines. The first one that creates `af` makes a floating point number whose value is $1/3$ of the measure of the angle called ang1. We then create the angle `ang2` having this measure and then add two points (D and E) making appropriate angles first with segment CB and then with segment CE. The final two lines can be typed in the editor or created by geometer with the standard "line through two points" command.

```
.geometry "version 0.60";
v1 = .free(0.403754, 0.4244, "A");
v2 = .free(-0.796153, -0.531501, "B");
v3 = .free(0.781055, -0.531501, "C");
l1 = .l.vv(v2, v1);
l2 = .l.vv(v2, v3);
ang1 = .a.vvv(v3, v2, v1);
af = .f.rpn(ang1, 3.000000, .div);
ang2 = .a.f(af);
v4 = .v.avv(ang2, v3, v2, "D");
v5 = .v.avv(ang2, v4, v2, "E");
l5 = .l.vv(v2, v4);
l6 = .l.vv(v2, v5);
```

3.8 Poncelet's Theorem Demonstration
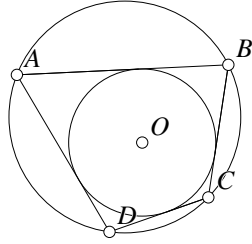
Figure 3.31: Poncelet's Theorem for Quadrilaterals

`Construction/Poncelet4.T [P]`

Poncelet's theorem states that if an n-sided polygon is at the same time inscribed in one circle and circumscribed about another as is the quadrilateral in figure 3.31, then if a different starting position is chosen on the outer circle and lines are successively drawn tangent to the inner circle and extended until they meet the other circle again, this process will terminate after n steps when the final line returns again to the starting point. In the **Geometer** diagram `Construction/Poncelet4.T` you can move the starting point A around the outer circle and see that the quadrilateral will close back on A after four steps no matter where A is located.

In addition, you can move the center point of the inner circle and its size will be adjusted automatically to make an inner circle of the correct diameter that a quadrilateral will continue to satisfy the conditions of Poncelet's theorem.

How is such a **Geometer** diagram created?

The construction is based on a theorem that states that in the special case of a quadrilateral inscribed in a circle of radius R and circumscribed about a circle of radius r, the distance d between the centers of the two circles satisfies:

$$\frac{1}{r^2} = \frac{1}{(R+d)^2} + \frac{1}{(R-d)^2}. \tag{3.1}$$

With this theorem, the construction of the diagram is straightforward. Here is the **Geometer** code that does the trick:

```
.geometry "version 0.60";
R = .f.rpn(0.950000);
V = .v.ff(0.000000, 0.000000, .in);
C = .c.vf(V, R);
O = .free(-0.125749, -0.269461, "O");
d = .f.vv(O, V);
r = .f.rpn(1.000000, R, d, .add, .dup,
        .mul, .div, 1.000000, R, d,
        .sub, .dup, .mul, .div, .add,
        1.000000, .exch, .div, .sqrt);
c = .c.vf(O, r);
v1 = .vonc(C, -0.884999, 0.345365, "A");
l1 = .l.vc(v1, c, 2);
v2 = .v.lcother(l1, C, v1, "B");
l2 = .l.vc(v2, c, 2);
v3 = .v.lcother(l2, C, v2, "C");
l3 = .l.vc(v3, c, 2);
v4 = .v.lcother(l3, C, v3, "D");
l4 = .l.vv(v4, v1);
l5 = .l.vv(v1, v2);
l6 = .l.vv(v2, v3);
l7 = .l.vv(v3, v4);
```

The outer circle is centered at the origin V with coordinates $(0, 0)$ and its radius is R. The radius R is a variable since it would be nice to adjust it later to produce a pleasing diagram where the outer circle is large, but not large enough that the text labels go off the screen. The value 0.95 was selected by trial and error after the rest of the diagram was written and debugged.

The point O is the center of the inner circle and can be moved freely by the user. Once the position of O is determined, the distance d between V and O is calculated.

The next line is a somewhat complex calculation using the reverse-polish calculator within **Geometer** to determine the value of r based on the values of R and d using the equation 3.1. Next, the smaller circle centered at v is drawn with radius r.

At this point, the text editing was terminated and two circles appeared on the **Geometer** window. Standard **Geometer** graphical commands were used to construct tangents and intersections of the points B, C and D based on an initial point A constrained to lie somewhere on the outer circle.

♦ The **Geometer** command .v.lcother was used to find the intersections of successive lines with the outer circle since that command guarantees that the *other* intersection is found—not the one with the previous point. The standard command .v.lc could probably have been used, but would not be certain to work in all configurations. See the **Geometer** manual for details.

3.9 Seven Tangent Circles

Suppose we would like to make a **Geometer** diagram like the one shown in figure 3.32. This seems a lot like the problems in Section 3.6.4 except that it is completely impossible to draw with classical methods, a regular heptagon.

If you know a little trigonometry, however, it is not too hard. A point on the unit circle that makes an angle θ with the x−axis has coordinates $(\cos\theta, \sin\theta)$. We would like to have seven evenly spaced points around the origin (the coordinates of the center of all **Geometer** diagrams are $(0, 0)$). Thus we need angles with measures of $0°, (360/7)°, (2 \cdot 360/7)°, \ldots$ The coordinates of the edges of a square **Geometer** diagram are at $x = \pm1$ and $y = \pm1$, so if the centers of the seven circles are to be visible, we will need them to be somewhat less than 1.

Here is the **Geometer** code that draws the figure:

Figure 3.32: Seven Tangent Circles
Construction/SevenCircs.T [P]

```
.geometry "version 0.60";
r = .f.rpn(0.650000);
a0 = .f.rpn(0.000000);
a1 = .f.rpn(360.000000, 1.000000, .mul, 7.000000, .div);
a2 = .f.rpn(360.000000, 2.000000, .mul, 7.000000, .div);
```

```
a3 = .f.rpn(360.000000, 3.000000, .mul, 7.000000, .div);
a4 = .f.rpn(360.000000, 4.000000, .mul, 7.000000, .div);
a5 = .f.rpn(360.000000, 5.000000, .mul, 7.000000, .div);
a6 = .f.rpn(360.000000, 6.000000, .mul, 7.000000, .div);
x0 = .f.rpn(a0, .cos, r, .mul);
y0 = .f.rpn(a0, .sin, r, .mul);
v0 = .v.ff(x0, y0, .in);
x1 = .f.rpn(a1, .cos, r, .mul);
y1 = .f.rpn(a1, .sin, r, .mul);
v1 = .v.ff(x1, y1, .in);
x2 = .f.rpn(a2, .cos, r, .mul);
y2 = .f.rpn(a2, .sin, r, .mul);
v2 = .v.ff(x2, y2, .in);
x3 = .f.rpn(a3, .cos, r, .mul);
y3 = .f.rpn(a3, .sin, r, .mul);
v3 = .v.ff(x3, y3, .in);
x4 = .f.rpn(a4, .cos, r, .mul);
y4 = .f.rpn(a4, .sin, r, .mul);
v4 = .v.ff(x4, y4, .in);
x5 = .f.rpn(a5, .cos, r, .mul);
y5 = .f.rpn(a5, .sin, r, .mul);
v5 = .v.ff(x5, y5, .in);
x6 = .f.rpn(a6, .cos, r, .mul);
y6 = .f.rpn(a6, .sin, r, .mul);
v6 = .v.ff(x6, y6, .in);
dvv = .f.vv(v0, v1);
rad = .f.rpn(dvv, 2.000000, .div);
c0 = .c.vf(v0, rad);
c1 = .c.vf(v1, rad);
c2 = .c.vf(v2, rad);
c3 = .c.vf(v3, rad);
c4 = .c.vf(v4, rad);
c5 = .c.vf(v5, rad);
c6 = .c.vf(v6, rad);
Rad = .f.rpn(rad, r, .add);
Origin = .v.ff(0.000000, 0.000000, .in);
C = .c.vf(Origin, Rad);
Radin = .f.rpn(rad, r, .sub);
Cin = .c.vf(Origin, Radin);
```

The r is the radius of the centers of the seven circles. It is safest to make it a floating point number called r since if we get it too large or too small, we can just change this one value later to make it right.

The next seven lines define the angles that divide $360°$ into seven equal parts. Next are seven sets of three lines of **Geometer** code that define the points that are at the center of each small circle. All of the points are drawn in the invisible color so they will not appear in the final diagram.

The value dvv is the distance between every pair of centers of those smaller circles, so the radius of those circles will be rad which is half of dvv.

Seven more lines of code draw the small circles with centers at v0, v1 and so on.

Finally, Rad and Radin are the radii of the outer and inner circles, Origin is the center of the diagram, and the other two lines of code draw the outer and inner circles.

3.10 A Tricky Construction

In Section 12.1.11 an interesting problem is presented, and the diagram is based on the situation illustrated in figure 3.33. Given a circle and an arbitrary chord AB of the circle, a point X is selected on that chord. The two circles that are simultaneously tangent to the chord at X and to the original circle are constructed. At the points of tangency, lines tangent to the circles are drawn. The problem in Section 12.1.11 is to show that the ratio of the diameters of the two tangent circles is constant, independent of the location of X, but the problem here is simply to construct the figure. The tangent lines through E are not necessary for the presentation of the problem, but illustrate another construction problem that can be tricky.

The easiest way to construct the figure is to use the method of inversion in a circle covered in Chapter 8. You have probably not read that chapter yet, but reading this section may convince you that the technique of inversion merits study.

The first part is easy. Draw a circle and then use the *P on C* command twice to create points A and B that are constrained to lie on that circle. Draw the segment connecting A and B and then use the *P on L* command to create the point X constrained to lie on AB. So much for the easy part.

We are searching for a line tangent to a line and a circle and passing through a point. If we can somehow invert the problem to send that troublesome point X "to infinity" then we will be searching for a line instead of a circle in the inverted problem. To send a point to infinity, we need to invert in a circle centered at that point.

In this problem, *any* circle centered at X will do, so pick another point and construct a circle centered at X and passing through that point. It may be a good idea

Figure 3.33: A Tricky Construction
Construction/Tricky.T [M]

to change the color of that circle temporarily to something you will recognize so that you do not get mixed up. Since X lies on the line AB, we know that AB will be inverted into itself, so the only object we need to invert is the original circle.

The command to do that is found in the *Primitives* pulldown menu under the *Circle* submenu. The command is *CC=>C Inv* (given two circles, invert the first in the second). After issuing the command, click on the original circle (the one with AB as a chord) and next on the inverting circle. A new circle will appear.

The solutions to this inverted problem will be lines that are parallel to AB and tangent to the inversion of the original circle. They must be parallel since they will meet the line AB at infinity (the inverted image of X, the center of the circle of inversion). There are two such lines.

To find the two solution lines, find the center of the inverted circle, and the easiest way to do that is with the *C=>P Ctr* command (found in the *Primitives* pulldown
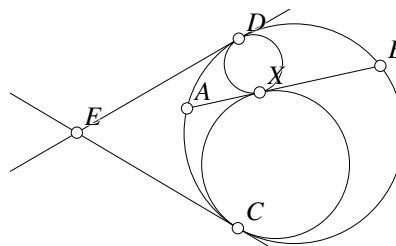
under *Point*). (Remember that the command creates a point, so it will be in the *Point* submenu.) From the point that lies at the center, draw a line perpendicular to AB using the *PL=>L Perp* command in the control area. Then find the intersections of this line with the inverted circle. There will be two solutions. Through both of those points, construct lines parallel to AB using the *PL=>L Par* command in the control area and we will have the two solution lines in the inverted problem.

Finally, to obtain the circles that are mutually tangent to the chord AB at X and to the original circle, invert each of those lines in the circle of inversion to obtain the desired tangent circles. Move X, A, and B to test your construction.

At this point, change the color of all the construction lines, points and circles to the invisible color so that all you can see are the points A, B, X, the segment AB, the original circle and the two tangent circles. (The names of the points in your diagram may be different, of course.)

Next, we would like to construct lines that are mutually tangent to the original circle and to the two tangent circles that you just constructed. The first impulse is to identify points that are at the intersection of those two circles using the *CC=>P* command in the control area, but there are some problems with this. Go ahead and try it and see what happens as you move points A, B and X.

The problem is that the circles are only constructed to the accuracy of the computer, not infinitely accurately. Thus the circles that appear tangent may actually pass through the outer circle twice (very close together), or they may miss it entirely due to numerical round-off error. As you move A, B and X the intersection points will tend to blink on and off which would cause the tangent lines also to blink on and off as the diagram is modified.

The best solution to get a solid diagram is this. The point D in the diagram must lie on the radius of both circles passing through that point. Thus the centers of both circles lie on a line passing through D. Find the centers of both circles. Depending on how you constructed the original circle, its center may be there, but simply colored the invisible color. If that is the case, click on the *Show Invis* button to see it, select it, and change its color temporarily.

When you have both circle centers, connect them with a line and find the intersection of that line with the outer circle. Do the same thing to find the point C. Be careful to find the intersection with the *outer* circle. We will see why shortly.

Finally, construct tangent lines to the outer circle using the *PC=>L* command in the control area, selecting the point and the outer circle. Find the intersection of the two lines at E and you are done.

If you were not careful and selected D to be the intersection with the inner circle, it might be slightly inside the outer circle from the point of view of the computer since the computer's mathematics is not perfect. If it is inside by a millionth of a centimeter, there will still be no tangent line. You could, of course, have selected C and D to be on the inner circles, and then construct tangents through those points to the inner circles. Either way works fine; the only way to have problems is to select the points on one circle and then draw tangents to the other. It may be worth experimenting with the incorrect constructions to see the sorts of problems you have: lines that blink on and off again.

3.11 Classical Construction

The classical constructions use only a straightedge and compass and a precise set of rules about how they may be used. In almost every case, you are given certain lengths, angles, points, lines, or circles, and are asked to construct others using only a straightedge and compass. See [Martin, 1991] for a book completely devoted to geometric construction.

Geometer and other computer geometry programs make this much easier than it is with a physical compass and straightedge—the lines are more precise, any part of the drawing can be adjusted as the construction progresses, and it is easy to undo errors. Besides, using **Geometer**'s "proof" technique, any of the sample constructions in this chapter may be viewed step by step much more easily with **Geometer** than by looking at the diagrams in the text.

The problem statement in the first paragraph of this section is sloppy: what does it mean when we say, "using only a straightedge and compass"? Are you allowed to put marks on the straightedge? Can you make a tangent to a circle by placing the straightedge against it so that it just touches? Even if you are not allowed to mark the straightedge, could you pin the compass to it with your fingers so that you can at least effectively have two marks?

Here is a precise definition of the meaning of construction with a straightedge and compass.

At any stage in the construction, you may do any of the following things to obtain additional points, lines or circles:

1. You may draw a straight line of any length through two existing points.

2. You may find a new point at the intersection of two lines, of two circles, or of a line and a circle. When you are given a segment, of course, you are given the two endpoints, so you can certainly use those.

3. You may construct a circle centered at any existing point having a radius equal to the distance between any two existing points. In other words, you can set the size of the compass from any two points A and B, and then you can move the point of the compass to another point C without changing the setting and draw a circle of radius AB about the point C.[4] (Of course this includes drawing a circle given its center and a point on the edge—you use the center and the edge to set the compass size, and then you re-use the center point as the center of the circle.)

4. You may choose an arbitrary point on a line, circle, or on the plane. (And of course you can also choose a point *not* on a line or circle as in "pick any point not on segment AB.")

---

[4]This is a "modern compass". Originally, most construction problems were stated in terms of a "Euclidean compass" which collapses after each circle is drawn. In other words, a Euclidean compass cannot be used to copy a length. But it is possible to do any construction with a Euclidean compass that can be done with a modern compass. In fact, all constructions that can be done with a straightedge and compass can be done with a compass alone—this is called the Mohr-Mascheroni theorem. See [Martin, 1991].

In the real world of pencils and paper (or computer calculations) none of these constructions can be carried out exactly for arbitrary points. But a mathematical construction assumes perfect tools—that you can find *exactly* where two lines cross, and you can set a compass width to *exactly* the distance between two points.

On the other side of the coin, given these perfect tools, you must find the required point or length *exactly*. Getting the answer accurate to one part in a trillion is not good enough. If all that is required is an answer that is only accurate to, say, one part in a trillion, there is usually a totally mechanical way to proceed, which will be discussed later.

3.12 ♦ Some Constructions Are Impossible

Many, many constructions using only a straightedge and compass are possible. In fact, it seemed to the ancient Greeks that almost any length could be constructed, but from antiquity, there were three famous problems that seemed to defy solution:

Trisection of an angle: Given an arbitrary angle α, construct a new angle having one-third the size: $\alpha/3$.

Doubling a cube: Given the length of the edge of a perfect cube, find the length of an edge of a cube having exactly twice the volume of the original cube.

Squaring the circle: Given a circle, construct a square having exactly the same area as the circle.

One by one (but thousands of years later) all three of these constructions were proved to be impossible. But the problem statements are so easy and the proofs of impossibility so difficult to understand, especially in the case of squaring the circle, that every day letters arrive at mathematics departments throughout the world from crackpots claiming to have solved one of these three classic problems (usually angle trisection)[5].

The constructions that are claimed to solve one of the problems above invariably contain an error. Quite often, however, the construction is correct in a sense—the "error" is that the solver did not understand the valid construction operations stated at the beginning of this chapter[6]. If you can put two marks on your straightedge, it is easy to trisect any angle, for example. Figure 3.34 shows how.

3.12.1 Angle Trisection

See figure 3.34. Construct an arbitrary circle with the angle α that you wish to trisect as a central angle. Then put two marks on your straightedge the same distance apart

---

[5]See [Dudley, 1994] for a huge collection of incorrect methods for angle trisections submitted by these crackpots.

[6]For example, instead of "straightedge and compass" one often hears the expression "ruler and compass" and the word "ruler" may imply that some sort of measurement is possible.
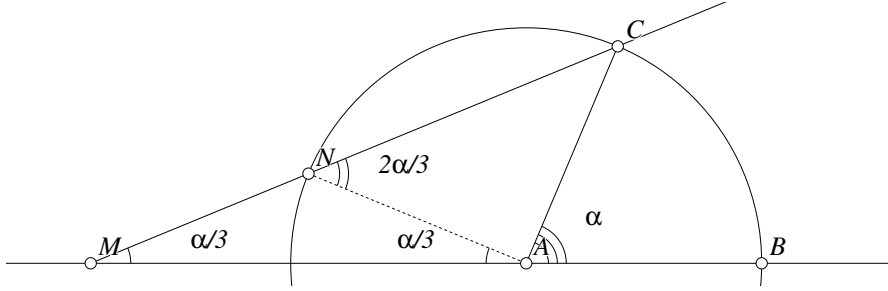
Figure 3.34: Trisecting an Angle
Construction/Trisect.T [M]

as the radius of the circle. Position the straightedge so that one of the marks is on the extension of the line AB (at M), the other is on the circle (at N, so $MN = AB$), and so that the straightedge passes through the point C where the central angle α intersects the circle. Then $\angle AMC = \alpha/3$.

The proof is not difficult. Let β denote the unknown angle at $\angle AMC$, and construct the line AN. $\triangle ANM$ is isosceles, so $\angle NAM = \beta$. Since a straight angle measures $180°$, we have:

$$\beta + \angle NAC + \alpha = 180°. \tag{3.2}$$

Since $AN = AC$, $\triangle NAC$ is also isosceles, with $\angle ANC = 2\beta$ because the exterior angle of a triangle is equal to the sum of the other two interior angles. Therefore, $\angle NAC = 180° - 4\beta$. We can substitute this value for $\angle NAC$ in equation 3.2 above and get

$$\beta + 180° - 4\beta + \alpha = 180°,$$

from which it is easy to conclude that $\alpha = 3\beta$, or $\beta = \alpha/3$ so we have achieved the trisection, but it required marking the straightedge.

3.12.2 ◆ Impossibility Proofs

The proofs that the three problems are impossible to solve using the standard rules for straightedge and compass construction follow the same pattern. First, you show that the given construction operations basically can only generate new points that are related to the old points via arithmetic operations: $(+, -, \times, \text{and} \div)$, or the square root. Then you show that certain numbers cannot be achieved in this way—certain cube roots, for example, or π, since it is transcendental (in other words, it is not the solution to *any* non-zero algebraic polynomial equation with integer coefficients, no matter how complicated). Then doubling the cube and squaring the circle are clearly impossible (doubling the cube requires the calculation of $\sqrt[3]{2}$ and squaring the circle requires the construction of $\sqrt{\pi}$). Angles cannot be trisected in general because the construction of a $20°$ angle implies that you can find the cosine of $20°$, which is the solution of an irreducible cubic equation over the integers. (To see this, expand $\cos 3\alpha$ in terms of $\cos \alpha$:

$$\begin{aligned}
\cos 3\alpha &= \cos(2\alpha + \alpha) = \cos 2\alpha \cos \alpha - \sin 2\alpha \sin \alpha \\
&= \cos(\alpha + \alpha)\cos\alpha - \sin(\alpha + \alpha)\sin\alpha \\
&= (\cos^2 \alpha - \sin^2 \alpha)\cos\alpha - 2\sin\alpha\cos\alpha\sin\alpha \\
&= \cos^3 \alpha - 3\sin^2 \alpha \cos\alpha \\
&= \cos^3 \alpha - 3(1 - \cos^2 \alpha)\cos\alpha \\
&= 4\cos^3 \alpha - 3\cos\alpha.
\end{aligned}$$

If $3\alpha = 60°$, $\cos 3\alpha = 1/2$, so the equation for $\cos\alpha$ is $1/2 = 4\cos^3 \alpha - 3\cos\alpha$, which is a cubic equation whose solution requires taking a cube root. In other words, if $x = \cos 20°$, then x is a solution to the equation

$$8x^3 - 6x - 1 = 0.$$

Therefore, one cannot construct the cosine of $20°$, and hence one cannot construct a $20°$ angle. Since there is one particular angle that cannot be trisected, the problem has no general solution. (Remember that *some* angles can be trisected—for example, it is pretty easy to trisect a $90°$ angle. In fact, it is often the case that the crackpot angle trisection solutions really trisect a single special angle.)

3.12.3 ♦ Constructible Regular Polygons

There are some "oddball" regular polygons that can be constructed. Without proof, here is the general formula. (See [Martin, 1991] for the details.)

If $p_1, p_2, \ldots, p_n$ are distinct Fermat primes and k is any non-negative integer, then a regular polygon with $m = 2^k p_1 p_2 \cdots p_n$ sides can be constructed with a straightedge and compass. A Fermat prime is a prime number of the form $2^{2^i} + 1$, where i is an integer. At the current time, only five different Fermat primes are known: $3, 5, 17, 257$, and 65537, corresponding to i values of $0, 1, 2, 3$, and 4. Thus it is possible to construct the regular 17-gon (called a regular heptadecagon), and such a construction (but not a proof that the construction is correct) is shown in Section 2.1.3. But (in theory, at least), it is also possible to construct a regular 257-gon and a regular 65537-gon.

If you believe that it is possible to construct a regular equilateral triangle, a regular pentagon, et cetera, it is not hard to convince yourself that the other constructions of the form above can also be done. It is not so easy to show that these are the only possible ones.

For example, to construct a regular 60-gon ($60 = 2^2 \cdot 3 \cdot 5$), construct a regular 3-gon (an equilateral triangle) and a regular 5-gon (a regular pentagon), and subtract the central angles. One is a third of a circle and the other is a fifth of a circle, so the difference is an angle that is $1/3 - 1/5 = 2/15$ of a circle. Bisect that angle three times to get an angle that is $1/60$ of a circle.

3.13 139 More Problems

In [Wernick, 1982], William Wernick provided a list of 139 triangle construction problems and gave the solutions for many of them. A more recent update on the state of the problems appears in [Meyers, 1996].

Basically, all require the construction of a triangle given three located points chosen from among the following:

A, B, C, O	the three points, and circumcenter
M_a, M_b, M_c, G	feet of the three medians, and centroid
H_a, H_b, H_c, H	feet of the three altitudes, and orthocenter
T_a, T_b, T_c, I	feet of the internal angle bisectors, and incenter

In other words, a problem like "(A, M_a, G)" means "Construct the triangle $\triangle ABC$ given the points that identify the point A, the foot of the median beginning at A, and the location of the centroid."

At the time of publication of [Meyers, 1996], there were still 20 open problems, among them, the following: (O, H_a, T_b), (O, H_a, I), (O, T_a, T_b), (M_a, M_b, I).

Some of the constructions are impossible (for example (G, H_a, H_b), since it is equivalent to finding the root of an irreducible cubic equation) and some are trivial (for example (A, B, C)).

3.14 Construction Exercises

The following exercises do not have solutions printed in the text. Unless they are marked with a warning ♦ symbol, they are not too difficult. You can find the solutions in **Geometer** files having the same name as the exercise. For example, to find the solution to problem C15, view the file C15.T in the Construction directory.

1. **[C1]** Given an isosceles right triangle $\triangle ABC$ where $\angle BAC = 90°$, find points H, I, J, and K such that H and I lie on BC, J lies on AB, and K lies on AC such that the quadrilateral $HIJK$ is a square.
2. **[C2]** Given a semicircle centered at a point C with diameter AB, find points I and J on AB, and points H and G on the semicircle such that the quadrilateral $GHIJ$ is a square.
3. **[C3]** Given a quadrant of a circle (two radii that make an angle of $90°$ and the included arc), construct a new circle that is inscribed in the quadrant (in other words, the new circle is tangent to both rays and to the quarter arc of the quadrant).
4. **[C4]** Given a point A, a line L that does not pass through A, and a point B on L, construct a circle passing through A that is tangent to L at the point B.
5. **[C5]** ♦ Given a point A, a line L, and a length x, construct a circle passing through A of radius x that is tangent to the line L.

6. **[C6]** ♦ Given a circle K, a line L, and a point A on L, construct a circle that is tangent to K and tangent to L at the point A. (Hint: There may be more than one solution.)

7. **[C7]** Given a point A and two straight lines L_1 and L_2, construct a line L that passes through A, and so that the angle L makes with L_1 is the same angle that it makes with L_2.

8. **[C8]** Given two points A and B and a line L, find a point C on L such that $AC = BC$.

9. **[C9]** Given two points A and B that both lie on the same side of a line L, find a point C on L such that AC and BC make the same angle with L.

10. **[C10]** Given two points A and B that both lie on the same side of line L, find a point C on L such that $AC + BC$ is as small as possible. (Hint: This problem is related to the construction problem 9. Also remember that the shortest distance between two points is a line.)

11. **[C11]** Given a point A on a line L. If B is any point on L, divide the segment AB into a given ratio by another point C (in other words, find C so that $AC : BC = r$ for some fixed ratio r). What is the locus of all possible positions of the point C?

12. **[C12]** Trisect a right angle.

13. **[C13]** ♦ If A and B lie on a line L, the two points C and D on L are said to **harmonically divide** the segment AB if $AC/CB = AD/DB$. Given points A, B, and C on a line L, find the point D such that C and D harmonically divide AB.

14. **[C14]** Given three points, A, B, and C on the plane, but not collinear, find another point X such that if L is *any* line through X, and if perpendiculars are dropped to L from A, B, and C that meet L at A', B', and C', respectively, then $AA' + BB' = CC'$. (Hint: Can you find any particular lines that work?)

15. **[C15]** Given two non-parallel lines L_1 and L_2 and a radius r, construct a circle of radius r that is tangent to both L_1 and L_2.

16. **[C16]** Given a line L and two points A and B on the same side of L and not on L, and given an angle α, construct a point C on L such that $\angle ACB = \alpha$. For certain configurations there may be no solutions.

17. **[C17]** Given a circle K, a point P, and a segment AB whose length is less than the diameter of K, construct a line L through P which intersects K at P_1 and P_2 such that $AB = P_1P_2$.

18. **[C18]** ♦ Given two circles K_1 and K_2 that do not intersect, and such that neither lies inside the other, together with a line segment AB such that the length of AB is smaller than the diameter of the smaller circle, construct a line passing through both circles such that both of the chords cut by the circles have length equal to AB.

19. **[C19]** Given a circle K and two perpendicular radii of K, construct a line that cuts the K twice and each chord so that the three segments of the line between these intersections have equal length.

20. **[C20]** ♦ Given a triangle and a point on one of the sides of the triangle, construct a line through that point that divides the area of the triangle in half.

21. **[C21]** Given three lines in the plane, find all the circles that are simultaneously tangent to all three lines.
22. **[C22]** Given any triangle $\triangle ABC$, construct the point P inside $\triangle ABC$ such that $\mathcal{A}(\triangle APB) = \mathcal{A}(\triangle BPC) = \mathcal{A}(\triangle CPA)$.
23. **[C23]** Given a parallelogram and a point on one of its edges. Construct a line through that point that bisects the area of the parallelogram.
24. **[C24]** ♦ Given a convex quadrilateral $ABCD$, construct a line through A that bisects the area of the quadrilateral.
25. **[C25]** ♦ Given a convex quadrilateral $ABCD$ and a point P on segment AB, construct a line through P that bisects the area of $ABCD$.
26. **[C26]** Given a line L and two points P and Q on the same side of L, construct the point R on L such that the angle $\angle PRQ$ is as large as possible.
27. **[C27]** Given a circle K and two points P and Q on the outside of K, construct the point R on K such that the angle $\angle PRQ$ is as large as possible.
28. **[C28]** ♦ Given an isosceles triangle, construct an equilateral triangle of equal area.
29. **[C29]** ♦ Given the two radii R and r, and a straight line L, construct two circles having those radii that are externally tangent to each other and are externally tangent to L.
30. **[C30]** Construct a circle that passes through two given points and is tangent to a given circle.
31. **[C31]** Given a point A, a circle K, and a point P on K, construct a circle passing through A that is tangent to K at P.
32. **[C32]** Given a circle K, a line L, and an angle α, construct a line tangent to K which makes an angle α with L.
33. **[C33]** Suppose you are given two non-parallel lines that meet so far away that you cannot access the point of intersection. Construct the bisector of the angle formed by those two lines.
34. **[C34]** Given a circle K and two points P and Q. Construct a circle K_1 passing through P and Q such that K_1 intersects K at two diametrically opposite points of K.
35. **[C35]** ♦ Given the hypotenuse and area of a right triangle, construct the triangle.
36. **[C36]** Given two points, construct two equal circles centered at those points such that their common external tangent passes through a third given point.
37. **[C37]** Given two points, construct two equal circles centered at those points such that their common internal tangent passes through a third given point.
38. **[C38]** Given two points and a circle, construct two circles of equal radius with centers at the two points such that the common external tangent of the circles is tangent to the given circle.
39. **[C39]** ♦ Given two concentric circles and a point P between them, construct the circles tangent to both concentric circles and to P. What happens if P is outside the outer circle? Inside the inner circle?
40. **[C40]** ♦ Given three mutually tangent circles, construct the two Soddy circles that are tangent to all three of them. If the triangle whose vertices are the centers of the three circles, the centers of the Soddy circles are called the **Soddy circles, centers**.

41. **[C41]** Given a right triangle and one of its legs, construct a point on that leg that is equidistant from the hypotenuse and the vertex opposite the hypotenuse.

42. **[C42]** Given three points, construct a parallelogram such that the three points are the midpoints of three of its sides.

43. **[C43]** Given a circle, a line, and two points, construct a circle passing through the two points whose common chord with the given circle is parallel to the given line.

44. **[C44]** Construct $\triangle ABC$ given the length of side BC, the length of the bisector of $\angle ABC$, and the size of $\angle ABC$.

45. **[C45]** ♦ Given segments of length x, y, and a, construct a segment of length b such that $x : y = a^2 : b^2$.

46. **[C46]** Construct $\triangle ABC$ given BC, AC, and the length of the median from A to the midpoint of BC.

47. **[C47]** ♦ Construct a quadrilateral $ABCD$ given $\angle A$, $\angle B$, $\angle C$, AB, and AD.

48. **[C48]** Given two equal intersecting circles and an angle, construct two equal chords of the two circles that meet at a point of intersection of the circles and form an angle equal to the given angle.

49. **[C49]** Given $\triangle ABC$ construct its two Brocard points. A **Brocard point** is a point Ω that makes equal angles with the sides of the triangle at the vertices. In other words, find a point Ω such that $\angle CA\Omega = \angle AB\Omega = \angle BC\Omega$. The other Brocard point Ω' makes the other angles equal: $\angle BA\Omega' = \angle AC\Omega' = \angle CB\Omega'$.

50. **[C50]** Construct the midpoint M of segment AB using a compass alone.

51. **[C51]** ♦ Construct a triangle given the lengths of its three altitudes. (Hint: The area of a triangle is half the base multiplied by the altitude. But you can use any altitude.)

52. **[C52]** Construct a triangle given the locations of the feet of its altitudes. (Hint: Look at the classical triangle centers of the triangle connecting the feet of the altitudes.)

53. **[C53]** Construct a triangle given one of its angles together with the radii of the incircle and the circumcircle. (Hint: See Section 6.1.)
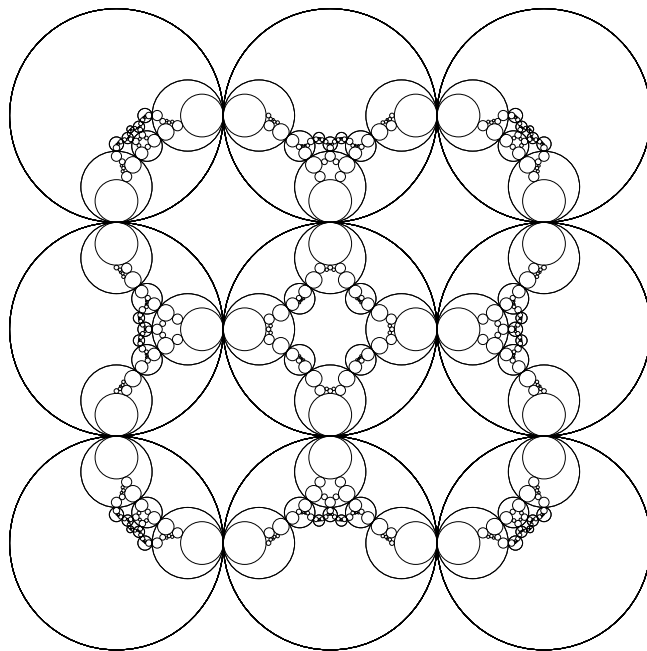
Chapter 4

Computer-Aided Proof

At the time this book is published, commercially available computer geometry programs are not capable of constructing the proof of a geometric theorem, but they can be used to aid in that process. In this chapter we will examine situations where tests and experiments can be done using **Geometer** to help construct a proof of geometric theorems.

Perhaps the most common use of computer geometry programs is to search for theorems and to provide evidence that a theorem or relationship is true. If you are told that the three altitudes of a triangle are concurrent at a point and you do not believe it, you can draw a triangle and its altitudes, and by moving the vertices of the triangle you can see that (at least to the accuracy of the computer screen and your vision) that they always seem meet at a point in thousands of cases. This is not a proof, of course, but when one is convinced that a result is true, the proof is psychologically easier to find[1].

---

[1] A great example of this is "the firing squad synchronization problem" which is described in

Warning: All of the examples in this chapter rely on **Geometer**. Other computer geometry programs may have similar commands, but if you use one of those programs, you may need to have the **Geometer** reference manual available to fully understand all of the examples.

4.1   Finding the Steps in a Proof

At this point it might be good to review the example presented in Section 2.8.

A formal proof in geometry (and in mathematics in general) consists of a series of statements beginning with the given information, and where each statement after that is a logical conclusion made from some combination of the previous statements. Every statement in the proof will be a consequence of the initial information, and the final statement is the theorem which was to be proved.

Proofs that appear in mathematical texts and journals may differ from this somewhat to make them easier to understand, but it should be clear how to convert them into formal proofs. For example, quite often you will see something like this: "If we can show that X is true, then our theorem is true because . . .". Later, there is a line like this: "So since we now know that X is true, we are finished." The formal proof would contain those statements in the opposite order, but the rearrangement above probably makes it easier for the reader to see where the argument is going.

The idea of a formal proof is driven home by the "T" proofs required in most high school geometry classes. Students draw a giant "T" on their paper, and every statement on the left has a corresponding reason on the right why it is true. The final statement on the left is the theorem to be proved.

For any interesting theorem, the proof does not come out of a mathematician's head that way. A bunch of things seem to be true, some depending on others, and the job of constructing a proof sometimes simply amounts to sorting these into a logical order. More often, however, it is difficult to discover many of those "true things".

One of the most powerful features of **Geometer** is that it can often provide a long list of possible entries for the left column of a "T" proof. Of course you then need to sort them into a logical order and provide reasons why each is true.

4.2   Testing a Diagram

Geometer contains a pair of commands, *Test Diagram* and *End Test* that provide a list of relationships that seem to be true.

[Minsky, 1967]. The problem remained an unsolved research problem for years, but as soon as the first solution was found, it became a standard, not too difficult exercise in computer science texts. Another good example is the following problem: "Prove that the number between any pair of twin primes except the pair $(3, 5)$ is divisible by 6." Most mathematicians panic at this point, having no idea how to proceed since so little is known about twin primes. But as soon as you add, "You have 10 seconds to solve it", almost all can prove it instantly.

To use this tool, after a diagram is drawn in **Geometer**, the *Test Diagram* button is pushed. Then the diagram can be modified at will, and the test is ended with the *End Test* command. At the end of the test, **Geometer** presents a list of all the relationships it discovered that seemed to hold throughout the diagram modifications. These include such things as equal segments, equal angles, concurrent lines, collinear points, parallel or perpendicular lines, and various other things.

Geometer makes an effort to avoid listing "obvious" relationships. In other words, if point C is constrained to lie on the line connecting points A and B, **Geometer** will not report that A, B and C are collinear. They are collinear by definition. Similarly, equal vertical angles will not be reported, et cetera. **Geometer** is conservative, however, so some of the things on the list may be "obviously true" to a human, but they were not obvious to **Geometer**.

The command pair that tests diagrams works like this: When **Geometer** enters the testing mode, it examines *all* sets of three points to see which sets all lie on the same line, it looks at *all* sets of four points to see which sets all lie on the same circle, and so on for all the other sorts of relationships it seeks. When any of these occur, and are not "obviously true", **Geometer** adds them to a list.

After making the initial list, as the diagram is modified while in testing mode, those relationships are tested over and over. If a relationship fails, it is removed from the list[2]. This is because a relationship may have held by accident in the initial configuration, but does it does not hold in general. When the user indicates that the test is over, **Geometer** presents everything that remains on the list[3].

4.3 Altitudes as Bisectors
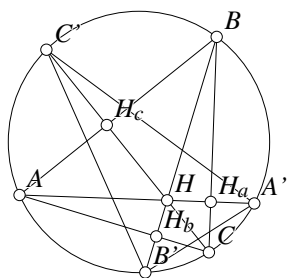
Figure 4.1: Altitudes as Bisectors
Proofaid/AltBisect.D [M]

Let us begin with a very easy example where we will illustrate a relationship between the altitudes of one triangle and the angle bisectors of a related triangle.

Show that if the altitudes of $\triangle ABC$ are extended to intersect that triangle's circumcircle at A', B' and C', forming $\triangle A'B'C'$ as in figure 4.1, then the altitudes of $\triangle ABC$ are the angle bisectors of $\triangle A'B'C'$.

If we test the diagram with **Geometer**, we obtain a list of relationships, but the important ones are these that include not only the angles that interest us, but some others as well:

---

[2]In practice, usually only a tiny motion of a couple of points is needed to eliminate relationships that occur by accident

[3]With suitable editing, of course. For example, if A, B and C are collinear, and so are A, B, and D, then **Geometer** will report that the four points A, B, C, and D are collinear.

```
Equal angles:
    (B A A') (B C C') (B B' A') (B B' C')
Equal angles:
    (C B B') (C A A') (C C' A') (C C' B')
Equal angles:
    (A B B') (A C C') (A A' C') (A A' B')
```

If we simply consider the first set of four equal angles, it is obvious that $\angle BAA' = \angle BB'A'$ and that $\angle BCC' = \angle BB'C'$ since both angles in each pair are inscribed in the same arc of the circle.

There are many ways to proceed, but let us begin with an easy one. We see that $\angle HH_aC = \angle HH_cA = 90°$ since AH and CH are altitudes, and $\angle H_aHC = \angle H_cHA$ because they are vertical angles. Thus $\triangle H_aHC \sim \triangle H_cHA$ because of AA, and so $\angle BCC' = \angle BAA'$. Thus all four angles in the previous paragraph are equal, so $B'B$ bisects $\angle A'B'C'$. Since there is nothing special about the angles we chose (or by symmetry, if you prefer to say it that way) the same is true of all the other angle bisectors.

4.4 Another Bisector

Two circles intersect at points M and N. The common tangent PQ is drawn on the side of the circles nearest N as in figure 4.2. The line PN is extended until it intersects the other circle at R. Show that MQ bisects $\angle PMR$.

This time if we test the diagram, there is absolutely no useful information returned that might help us with the proof, other than reporting that MQ does, in fact, seem to bisect $\angle PMR$.
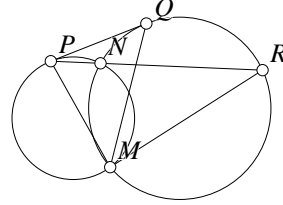
Figure 4.2: Another Bisector
`Proofaid/CircBisect.D` [M]

When this happens a good strategy is to look at the figure to see if drawing some auxiliary lines might help. This is such a simple figure that there are not too many possibilities, so we might as well draw the obvious ones: MN, NQ and QR. In more complex figures it is usually a bad idea just to draw every line you can think of, since the number of reported relationships can easily become unwieldy.

But with the addition of just those three additional lines, the list of relationships now includes a few that might be useful:

```
Equal angles:
    (N R M) (M Q N)
Equal angles:
    (P Q N) (N R Q) (Q M N)
Equal angles:
    (Q P M) (M Q R)
Equal angles:
    (Q P N) (P M N)
```

There are a few ways to complete the proof with this information, but probably the easiest is simply to use the last equality: $\angle QPN = \angle PMN$. The angles are obviously equal since PQ is tangent to the smaller circle in the figure and cuts off exactly the same arc $\overset{\frown}{NP}$ as does $\angle NMP$.

We know that angle $\angle QPN$ is equal to the difference of arcs $\overset{\frown}{RQ} - \overset{\frown}{QN}$. We also know that $\overset{\frown}{QN} = \angle QMN$ and that arc $\overset{\frown}{RQ} = \angle RMQ$, so if we combine all this information we have:

$$\angle NMP = \angle QPN = \angle RMQ - \angle QMN,$$

so $\angle QMP = \angle QPN + \angle QMN = \angle RMQ$ which means that MQ bisects $\angle PMR$.

4.5 Perpendicular Diagonals
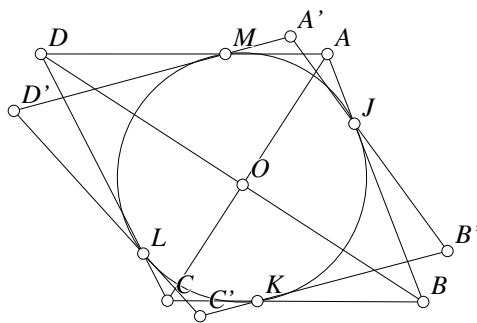
Figure 4.3: Perpendicular Diagonals
`Proofaid/PerpDiags.D [M]`

If a convex quadrilateral $ABCD$ has perpendicular diagonals that meet at a point O, if the quadrilateral is rotated about O by an arbitrary angle θ to form the new quadrilateral $A'B'C'D'$, then the points J, K, L and M that lie, respectively, at the intersections of AB and $A'B'$, BC and $B'C'$, CD and $C'D'$ and finally DA and $D'A'$ all lie on a circle. See figure 4.3.

When the diagram above is drawn in **Geometer** and then is tested for relationships that hold as the figure is manipulated, **Geometer** returns a long list. Many items in the list are obvious, but the following jump out at us.

```
Points on a circle:
    O  D  D'  L  M
Points on a circle:
    O  B  B'  J  K
Points on a circle:
    O  C  C'  K  L
Points on a circle:
    O  A  A'  J  M
```

Now that it is pointed out to us that these sets of five points lie on circles, it is easy to see why. To take the first one as an example, since the two quadrilaterals are rotated by an angle θ relative to each other, $\angle DMD'$, $\angle DOD'$ and $\angle DLD'$ are all equal to θ and therefore lie on the same circle.

Since they lie on the same circles, there are other angles in those circles subtended by equal arcs, so we can write eight equations like: $\angle ODM = \angle OLM$, $\angle ODL = \angle OML$ and so on.

Since the diagonals of the quadrilateral are perpendicular, we know that triangles $\triangle BOA$, $\triangle AOD$, $\triangle DOC$ and $\triangle COB$ are right triangles so their non-right angles are complementary.

Here is a set of equations derived from the observations above:

$$
\begin{aligned}
\angle OJK &= \angle OBC & (4.1)\\
\angle OJM &= \angle OAD & (4.2)\\
\angle OLK &= \angle OCB & (4.3)\\
\angle OLM &= \angle ODA & (4.4)\\
\angle ODA + \angle OAD &= 90° & (4.5)\\
\angle OBC + \angle OCB &= 90° & (4.6)
\end{aligned}
$$

Substituting equations 4.1, 4.2, 4.3 and 4.4 into equations 4.5 and 4.6, and adding those equations together yields:

$$\angle OLM + \angle OJM + \angle OJK + \angle OLK = 180°. \qquad (4.7)$$

But since $\angle OLM + \angle OLK = \angle MLK$ and $\angle OJM + \angle OJK = \angle KJM$, equation 4.7 tells us that $\angle MLK$ and $\angle KJM$ are complementary, so $JKLM$ must be a cyclic quadrilateral.

4.6 Constant Sum
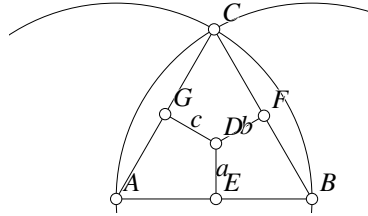
Figure 4.4: Equilateral Projections
`Proofaid/Equilateral.T` [M]

Show that if any point is picked in an equilateral triangle that the sum of the distances to the three sides is constant.

To make a **Geometer** diagram, everything is easy until the final step. We can draw an equilateral triangle beginning with one side by drawing circles centered at each end of the arc and passing through the opposite end. The intersections of the circles are suitable positions for the third point, that can be connected, giving us $\triangle ABC$. Then place a free point D inside and drop perpendiculars

to each side. **Geometer** can only calculate the distances between points, so find the intersections, and you should have a diagram like that in figure 4.4. It would probably be better to give the circles the invisible color to simplify the final diagram.

Now we would like to display the lengths of the three segments DE, DF and DG. This is also not hard. Select one of the lines and then choose the *Display Value* command from the *Edit* pulldown menu, you will get something like "l4 = 0.425" displayed in the upper left of the **Geometer** window. So far you have not given that line a name (points receive names by default but other geometric figures do not). You can do so by selecting it (it should still be selected if you just issued the *Display Value* command) and issue the *Edit Name* command in the *Edit* pulldown. This brings up a dialog box into which you can type the name of the line. In figure 4.4, the lines have been named a, b and c and the display for each has been toggled on. (The keyboard shortcut for those commands is **ctrl-T** to toggle the display and **ctrl-n** to change the name.)

When the lengths are displayed, they appear in the upper left of the diagram like this: "$a = 0.514$".

As you move point D around, the values of the lengths are continuously updated, but it is still a bit of work to add the three each time to assure yourself that the sum is constant. Why not have **Geometer** do that for you? To do this you will need to edit the text version of the file. The entire modified file is listed below since you need to know the internal names that **Geometer** gave to the various points before you can type in the final code.

```
.geometry "version 0.60";
v1 = .free(-0.92, -0.871156, "A");
v2 = .free(0.893333, -0.871156, "B");
c1 = .c.vv(v1, v2);
c2 = .c.vv(v2, v1);
v3 = .v.cc(c1, c2, 2, "C");
l1 = .l.vv(v1, v2);
l2 = .l.vv(v2, v3);
l3 = .l.vv(v3, v1);
v4 = .free(0.026616, -0.372624, "D");
l4 = .l.vlperp(v4, l1, .display, "a");
l5 = .l.vlperp(v4, l2, .display, "b");
l6 = .l.vlperp(v4, l3, .display, "c");
v5 = .v.ll(l1, l4, "E");
v6 = .v.ll(l2, l5, "F");
v7 = .v.ll(l3, l6, "G");
DE = .f.vv(v4, v5);
DF = .f.vv(v4, v6);
DG = .f.vv(v4, v7);
S = .f.rpn(DE, DF, DG, .add, .add, "Sum");
```

The three lines beginning with DE, DF and DG set those names to be the distance between the points in parentheses. The function `.f.vv` makes a floating point number (`.f`) from the distance between two points (`.vv`). The final line performs a floating point calculation with **Geometer**'s built-in rpn[4] calculator. The three lengths are pushed on

---

[4]Reverse Polish Notation.

the stack, and then the top two items are added and the operation repeated. Since the floating point number `S` has a name (`Sum`), its value is also continuously displayed in the upper left of the window.

Notice that the lines `l4`, `l5` and `l6` have the property `.display`. This tells **Geometer** to display their lengths in the window. We could have done this without using the `.display` property simply by giving names to the numbers `DE`, `DF` and `DG`. In that case, the named lines would no longer have the `.display` attribute and the final four lines would look like this, and the view on the screen would remain the same:

```
DE = .f.vv(v4, v5, "a");
DF = .f.vv(v4, v6, "b");
DG = .f.vv(v4, v7, "c");
S = .f.rpn(DE, DF, DG, .add, .add, "Sum");
```

Once you have made the diagram and tested the conjecture by moving the point D inside the triangle to verify that the sum does seem to remain constant, you can begin to look for a proof.

Note: With some effort you could avoid editing the text file to present the sum. Using standard techniques you could construct a line that is the sum of the three lengths and display its length.

When you are trying to show something about a sum of lengths in a geometric way, it is often a good idea to try to make a construction where the three lengths are set end to end. In addition, in this case where you are trying to show that the sum is constant, it would be nice to know what that constant might be.

The first approach seems difficult, but we can get some idea of what the sum might be by looking at some special positions for D. These special positions might include the center, points on the edge, or points at the vertices of the triangle $\triangle ABC$. If D is very close to a vertex, almost the entire sum will be due to one length, and that length is basically the altitude of the triangle.

Why would the sum be equal to an altitude? If it is true, might that lead to a proof? One of the most common uses for the altitude of a triangle is to find its area which is half the base times the altitude. If point D is inside the triangle, the three segments are not altitudes of triangle $\triangle ABC$, but if we connect D to the vertices with line segments, they are the altitudes of triangles $\triangle ADB$, $\triangle BDC$ and $\triangle CDA$.

Since $\triangle ABC$ is equilateral, all three of those smaller triangles have the same length base, so the area of each is the half the product of that base times the altitude. Since the three triangles combine to give the full triangle $\triangle ABC$, their areas must add to the area of the original triangle.

That is basically the proof. Here is a more formal presentation:

We know that

$$\mathcal{A}(\triangle ABC) = \mathcal{A}(\triangle ADB) + \mathcal{A}(\triangle BDC) + \mathcal{A}(\triangle CDA). \tag{4.8}$$

If h is the altitude of $\triangle ABC$ and $s = AB = BC = CA$, then equation 4.8 becomes:

$$\frac{sh}{2} = \frac{s(DE)}{2} + \frac{s(DF)}{2} + \frac{s(DG)}{2} = \frac{s(DE + DF + DG)}{2},$$

so we conclude that $DE + DF + DG = h$.

4.7 Find a Radius

Given a right angle at A and two different points B and C on one of the side of that angle, find the radius (in terms of the lengths of the segments AB and AC) of the circle that passes through B and C and is tangent to the other side of the angle.

This is an easy problem, but if you decide to investigate it with **Geometer**, you apparently immediately encounter a much harder problem, since it is not particularly easy to construct a line through two points and tangent to a line[5].
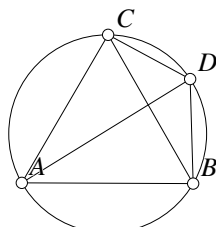

Figure 4.5: Tangent Circle Diagram
`Proofaid/Radius.T [M]`

Figure 4.6: Sum of Lengths
`Proofaid/EqualSum.T [M]`

To draw a useful diagram, however, all you need is to draw something with a circle tangent to one ray of a right angle and passing through the other ray in two points labeled B and C. To do so, draw a right angle, draw a free point O to be the center of the circle, drop a perpendicular line from O to one side of the angle, and the point of intersection will be the point of tangency. The intersections of the circle with the other line are B and C. See figure 4.5.

(Alternatively, you can draw a circle, then a tangent to it, then a perpendicular to the tangent passing through the circle at B and C.)

With the drawing, it is not hard to find the answer. Drop a perpendicular from O to the other side of the angle at X. The point X is clearly the midpoint of segment BC, and AX is clearly the same as OT, the radius of the circle that we are seeking. Clearly $AX = (AB + AC)/2$.

4.8 Sum of Lengths

Show that if $\triangle ABC$ is equilateral and point D lies on its circumcircle on the arc $\overset{\frown}{BC}$, then $AD = BD + CD$. See figure 4.6.

4.8.1 Geometer Analysis of the Problem

---

[5]It can certainly be done; see Section 3.6.5.

It is quite easy to draw the **Geometer** diagram, but when we test it, we seem to obtain almost no information. If you do not believe that it is true, you can help convince yourself by clicking on the segments AD, BD and CD and issuing the *Toggle Display Value* command which causes their lengths to be displayed in the upper left portion of the **Geometer** window. The three lengths will be displayed, but you will need to do some mental arithmetic each time to convince yourself that $AD = BD + CD$ each time.
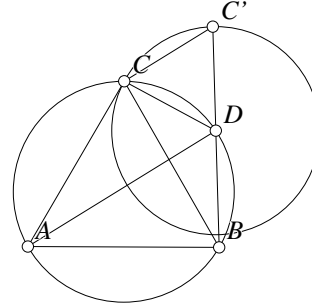
Figure 4.7: Sum of Lengths

`Proofaid/EqualSum.T [T]`

You can add a few lines of code with the text editor to display not only the lengths of the segments, but the result of calculating $S = BD + CD - AD$. If you do that, the lengths and their difference is constantly calculated and you will see that it is always zero.

Here is the code to display these values, assuming the internal point names for A, B, C and D are `v1`, `v2`, `v3` and `v4`, respectively:

```
dbd = .f.vv(v2, v4, "BD");
dcd = .f.vv(v3, v4, "CD");
dad = .f.vv(v1, v4, "AD");
diff = .f.rpn(dbd, dcd, .add, dad, .sub, "BD + CD - AD");
```

To make some progress, why not construct a segment of length $BD + CD$? The easiest way to do that is to extend the segment DB by an additional CD as in figure 4.7. In that figure, C' is on the line BD and $C'D = CD$. We would like to show that $C'B = AD$. It is probably hard to work with just a segment pointing out into the void, so connect C' to C so that at least the new point will be part of a couple of triangles.

With this simple addition, the *Test Diagram* command returns a lot of information, some of which is quite interesting:

```
Equal angles:
   (A C D) (B C C')
Equal length segments:
   [C D] [D C'] [C C']
```

If we can show that these are true, we will know that $\triangle ACD \cong \triangle BCC'$, and we will be able to conclude that $AD = BC'$ which is what we need to show.

It looks like the little triangle $\triangle CDC'$ is equilateral and if we can show that, it will be sufficient to show that the triangles in the previous paragraph are congruent. But $\triangle ABC$ is equilateral, so clearly $\angle CDB = 120°$, so $\angle CDC' = 60°$. Since we constructed $CD = C'D$ we can show that $\triangle CDC'$ is equilateral.

4.8.2 Proof of the Theorem

Extend segment BD to C' so that $DC' = DC$.

Since $\triangle ABC$ is equilateral and $ABDC$ is a cyclic quadrilateral we know that $60° = \angle BAC$ is supplementary to $\angle BDC$, so $\angle BDC = 120°$. Since $\angle BDC$ and $\angle CDC'$ together make a $180°$ angle, $\angle CDC' = 60°$. But $CDC = DC'$, so $\triangle CDC'$ is equilateral and thus $CC' = CD$.

Since $\triangle ABC$ is also equilateral, $AC = BC$, and since $\angle ACD = \angle BCD + 60° = \angle BCC'$ we can conclude that $\triangle ACB \cong \triangle BCC'$. Since the corresponding parts of congruent triangles are congruent, $AD = BC' = BD + CD$.

4.9 Bisector Bisector
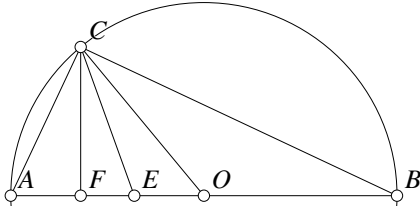
Figure 4.8: Equilateral Projections
`Proofaid/Bisect.T` [M]

Show that if $\triangle ABC$ is a right triangle with $\angle C = 90°$, then the angle bisector at C bisects the angle formed by the median and the altitude from C.

We can easily construct a **Geometer** diagram where it is possible to display all possible right triangle shapes by beginning with the hypotenuse AB and choosing C to lie on the circle whose diameter is the hypotenuse AB. If we then draw in the altitude AF, the angle bisector AE and the median AO we obtain figure 4.8.

4.9.1 Geometer Analysis

If we test the diagram, we obtain quite a few relations, but among them are these:

```
Equal angles:
    (F C E) (E C O)
Equal angles:
    (A B C) (B C O) (A C F)
Equal length segments:
    [A O] [O C]
```

Since CE is the angle bisector, if $\angle ACF = \angle BCO$ we can subtract those from the known equal angles $\angle ACE$ and $\angle BCE$ to obtain the result we want. The last relation above reminds us that all the radii of a circle are equal, and that is plenty of information to build a proof.

4.9.2 Proof of the Result

Since CO is a median, point O is the midpoint of segment AB. We know that if $\angle ACB = 90°$ then C must lie on a circle whose diameter is AB. Thus $AO = BO = CO$. Since CF is an altitude, $\angle CFA = 90°$ and since $\angle FAC$ is equal to itself, by AA, $\triangle ACF \sim \triangle ABC$. Because of that similarity, $\angle ACF = \angle ABC$.

Since $CO = BO$, $\triangle BOC$ is isosceles, so $\angle OBC = \angle OCB$. But $\angle BOC = \angle ACF$, so $\angle ACF = \angle OCB$. Segment CE is the bisector of $\angle ACB$, so $\angle ACE = \angle BCE$. If we subtract the equal angles $\angle OCB$ and $\angle ACF$ from them, we obtain $\angle FCE = \angle ECO$, which means that CE bisects $\angle FCO$.

4.10 The Simson Line

Let M be any point on the circumcircle of an arbitrary $\triangle ABC$. Let A', B' and C' be the perpendicular projections of M on the edges of $\triangle ABC$. (In other words, drop perpendicular lines from M to each of the edges of the triangle and find the points of intersection. At least one of those points has to lie outside the triangle.) Then A', B' and C' lie on a common line called the **Simson line** of the triangle relative to the point M. See figure 4.9.
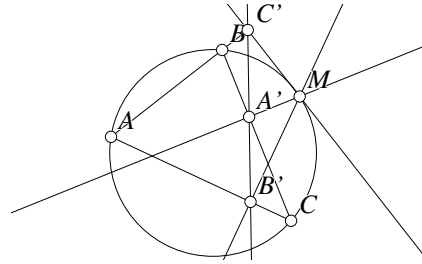
Figure 4.9: The Simson Line
`Proofaid/Simson.T [M]`

4.10.1 Geometer Analysis of the Simson Line

When the diagram above is tested, the following interesting relationships appear:

```
Points on a circle:
    C   M   B'   A'
Points on a circle:
    B   M   C'   A'
Points on a circle:
    A   M   B'   C'
```

That these relationships hold is obvious, since the perpendicular lines that were dropped from M form right angles. But when sets of four points are concyclic, we can find many equal angles, and a good way to show that the points A', B' and C' are collinear is to show that $\angle BA'C' = \angle CA'B'$. Since BC is a line, $B'A'C'$ must also be a line since equal vertical angles are formed.

The nice thing is that $\angle BA'C'$ and $\angle CA'B'$ are both inscribed in the circles with four concyclic points.

Since M is on the circumcircle of the triangle, $ABMC$ are concyclic points so we can relate $\angle BMC$ to $\angle BAC$. But we also know that $B'MC'A$ are concyclic so we can relate $\angle C'MB'$ to the same $\angle BAC$.

From these observations, the proof is easy.
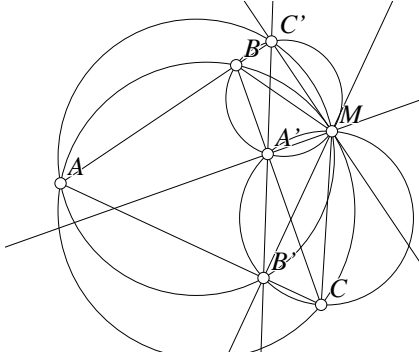
Figure 4.10: The Simson Line
Proofaid/Simson1.T [M]

4.10.2 Proof of the Existence of the Simson Line

The proof depends on the fact that M is on the arc $\overset{\frown}{CB}$ (which will cause A' to be between B' and C'). If this is not the case, simply relabel the figure so that it is.

Since $MA' \perp BC$ and $MC' \perp AB$ right angles are formed at A' and C' so $MA'BC'$ are concyclic. Similar reasoning shows us that $MA'B'C$ and $B'MC'A$ also form concyclic sets of points. Draw the three circles containing those sets of points and in addition, the line segments MB and MC. See figure 4.10.

Since the points $MA'BC'$ are concyclic, $\angle BA'C' = \angle BMC'$. Similarly, since the points $MA'B'C$ are concyclic, $\angle CA'B' = \angle CMB'$.

Since $\angle C'MB' = \angle C'MB + \angle BMB'$ and $\angle BMC = \angle BMB' + \angle B'MC$, we will be done if we can show that $\angle C'MB' = \angle BMC$ since this will imply that $\angle C'MB = \angle B'MC$. Therefore $\angle BA'C' = \angle CA'B'$ which will tell us that equal vertical angles are formed and thus the points A', B' and C' are collinear.

Since $MABC$ form a concyclic set of points, $\angle BMC$ and $\angle BAC$ are supplementary. Since $B'MC'A$ are concyclic, $\angle C'MB'$ and $\angle BAC$ are also supplementary. Since $\angle BMC$ and $\angle C'MB'$ are supplementary to a common angle, they are equal, and the proof is complete.

4.11 A Property of Two Cevians

In figure 4.11, let $\triangle ABC$ be any triangle and choose D and E on the segments BC and AB, respectively. Construct the two circles with diameters AD and CE that intersect at points I and J. Show that IJ passes through the orthocenter H of $\triangle ABC$.

Notice that the line IJ will be the radical axis of the two circles. See Section 5.7 for information about the radical axis.

4.11.1 Geometer Analysis of the Two-Cevian Problem

Since the conjecture involves the orthocenter H, it is only reasonable to draw the altitudes AA', BB' and CC' of the triangle before doing the test since the common intersection of those altitude lines define the orthocenter. When the diagram is tested by **Geometer**, there is not too much of interest, but there are a couple of things:

```
Points on a line:
   H  I  J
Points on a circle:
   C  E  I  J  C'
Points on a circle:
   A  D  I  J  A'
```

We are reminded that the feet of the altitudes AA' and CC' also lie on the circles. This is obvious since the altitudes meet the triangle bases at $90°$ angles. There are no useful ratios—only the obvious ones, but we do notice that **Geometer** reports that H, I and J lie on a line. This seems obvious, but remember, H was defined to be at the intersection of a pair of altitudes, so **Geometer** does not know that it lies on IJ (and neither do we—that is what we are trying to prove).

This does provide an idea. Draw the segments HI and HJ and see what **Geometer** has to say. To show that the three points lie on a line, we may be able to prove that those new segments cause that force vertical angles to be equal or something similar. When we do this, **Geometer** testing yields the following interesting new relationships:

```
Ratios:
   [A H] x [H A'] = [C H] x [H C']
   [B H] x [H B'] = [H I] x [H J]
   [A H] x [H A'] = [H I] x [H J]
```

If these are true, then H does lie on IJ, since IJ is the radical axis (see Section 5.7) of the two circles and H is a point with equal powers relative to those two circles and hence lies on the radical axis.
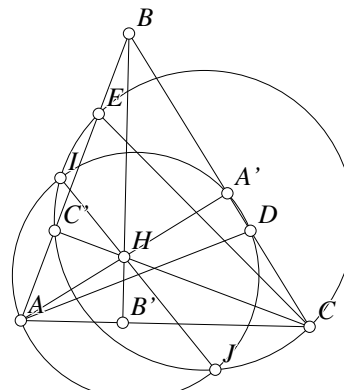
Figure 4.11: Cevians as Diameters
Proofaid/Cevians.T [M]

4.11.2 Proof of the Two-Cevian Problem

Using again figure 4.11, we observe that A', the foot of the altitude from A, lies on the circle with diameter AD since $\angle AA'D$ is a right angle. Similarly, C' lies on the circle with diameter CE.

If I and J are the points of intersection of the two circles, then IJ is the radical axis of those two circles. Since it is the radical axis, we know that the intersection H_1

of this axis with CC' will satisfy $H_1 C' \cdot H_1 C = H_1 I \cdot H_1 J$ and the intersection of H_2 with AA' will satisfy $H_2 A' \cdot H_2 A = H_2 I \cdot H_2 J$.

If we can show that H, the orthocenter, satisfies $HA \cdot HA' = HC \cdot HC'$ we are done since H will be equal to both H_1 and H_2.

Since $\angle AHC' = \angle CHA'$ and $\angle HC'A = \angle HA'C$ we know that $\triangle AHC' \sim \triangle CHA'$ by AA. This similarity tells us that $HA \cdot HA' = HC \cdot HC'$ holds and our theorem about the two cevians is proved.
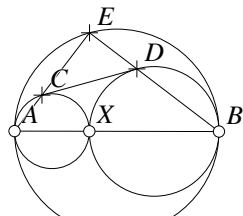
4.12 Tangent Line Problem

Figure 4.12: Tangent Line Problem
Proofaid/Zvezda.T [M]

If AB is the diameter of a circle choose X anywhere on the segment AB. Construct the two circles having AX and XB as diameters, and find the point E on the larger circle that is the intersection of the circle and the perpendicular to AB through X. If AE and BE intersect the smaller circles at C and D, show that CD is the common external tangent of the smaller circles[6]. See figure 4.12.

4.12.1 Geometer Analysis of the Tangent Line Problem

If we test the **Geometer** diagram for relations, we obtain the following possibly interesting relationships:

```
Points on a circle:
    X  E  D  C
Points on a circle:
    A  B  D  C
Right angles:
    (B E A)
```

Since E is on the larger circle, it is obvious why $\angle BEA = 90°$. The cyclic quadrilaterals are also likely to help, since they tell us a number of things about opposite angles.

How does one show that a line is tangent to a circle? One way is this: If $\angle XCD = \angle XAC$ we are done, since they both cut off the same arc $\overset{\frown}{XC}$ of the circle. Clearly we need only show that CD is tangent at C, since a totally symmetric argument will show that it is also tangent at D.

4.12.2 Proof of the Tangent Line Problem

---

[6]Thanks to Zvezdelina Stankova for showing me this problem..

See figure 4.13. Construct segments CX, DX and a perpendicular to AB through X to a point Y on the line CD. Let $\alpha = \angle CAX$. If we can show that $\angle YCX = \alpha$, we will be done.

Since both cut off the same arc $\overset{\frown}{XC}$, $\angle CAX = \angle YXC = \alpha$. In addition to $\angle AEB$, we know that $\angle ECX = \angle EDX = 90°$ so $CEDX$ is a rectangle. The diagonals of a rectangle bisect each other, so $XY = CY$. Thus $\triangle CYX$ is isosceles, so $\angle YCX = \angle YXC = \alpha$, so CY must be tangent to the circle with diameter AX.
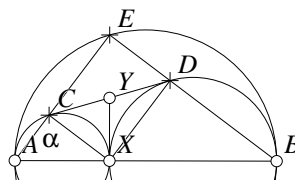
Figure 4.13: Tangent Line Problem
`Proofaid/Zvezda1.T` [M]

4.13 Cyclic Quadrilateral Problem

In figure 4.14, let $ABCD$ be a cyclic quadrilateral. Draw four circles having each of the sides of $ABCD$ as a diameter. Show that the intersections of adjacent sets of circles E, F, G and H also lie on a circle.
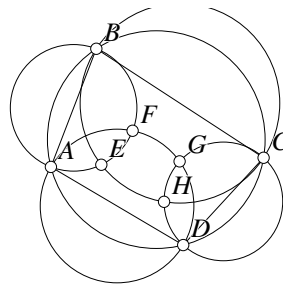
Figure 4.14: Cyclic Quadrilateral Problem
`Proofaid/CyclicQuads.T` [M]

4.13.1 Geometer Analysis of the Problem

A **Geometer** test of the diagram contains the following interesting relationships:

```
Points on a line:
   B  D  F  H
Points on a line:
   A  C  E  G
```

We know that the usual way to show that four points are concyclic is to show that the opposite angles are supplementary, so draw those lines in addition to the lines AC and BD (which appear to pass through E and G or F and H). To get some handle on the interior angles it may be interesting to draw lines like DG so we do that and re-test the diagram (figure 4.15) which yields more interesting information:

```
Parallel lines:
```

```
    (AD) (EH)
Parallel lines:
    (CD) (EF)
Parallel lines:
    (BC) (FG)
Parallel lines:
    (AB) (GH)
Perpendicular lines:
    (AC) (DG)
```

We see that $DG \perp AC$ because they are both inscribed in a semicircle, and that the same thing would hold for any of the other circle diameters.

The parallel lines show us that $\angle GAD = \angle GEH$, et cetera. If we can show that $\angle GEH = \angle GHF$ we will have solved the problem.

By now we have plenty of information to construct the proof.

4.13.2 Proof of the Cyclic Quadrilateral Problem

Since $CGHD$ is a cyclic quadrilateral with CD as diameter, $\angle DGC$ is a right angle. Similarly, $\angle AGD$ is inscribed in the circle with diameter AD, so $\angle AGD$ is also a right angle. These two facts show us that G is on the line AC. Similar arguments about the other points prove that $AEGC$ are all on a line and similarly for $BFHD$.

Because they are both inscribed in the same circle with diameter AD, we know that $\angle GFD = \angle GAD$. Similarly, $\angle CEH = \angle CBH$.
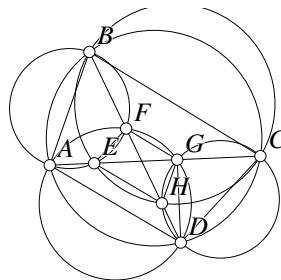
Figure 4.15: Cyclic Quadrilateral Problem
`Proofaid/CyclicQuads1.T [M]`

But $\angle CBD = \angle CAD$ since they are both inscribed on the original circle so we can conclude that $\angle GFH = \angle GEH$ and therefore $EFGH$ is a cyclic quadrilateral.

4.14 A Strange Relationship

Let triangle $\triangle ABC$ have its angle at A bisected by a segment AD, where D is on the line BC. At D construct a perpendicular to AD intersecting AB at E. Find the length of AE in terms of the lengths of AB and AC[7].

---

[7]Thanks to Tatiana Shubin for this problem.

What is surprising and perhaps a bit strange is that the relationship sought does not seem to depend on the angle at A—only on the lengths of the two sides that meet at that angle. As usual, let us call the lengths of the sides AC and AB b and c, respectively.

Figure 4.16: A Strange Relationship
`Proofaid/Strange.T [M]`

4.14.1 A Geometer Analysis of the Problem

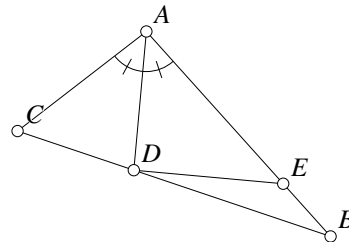If we test the diagram as it stands, **Geometer** returns no interesting information. It does tell us that $AC/CD = AB/BD$, but that is simply because AD is an angle bisector.

We will need to make some construction to proceed. Notice that if we can find the length of either AE or EB in terms of b and c we are done, so it seems that it might be a good idea to construct a line from E that makes some similar triangles that include AE and/or EB as sides. The approach used here is to construct a line through E and parallel to AC and intersecting CD at F. At least this will give us similar triangles $\triangle ABC \sim \triangle EBF$ that involve the unknown lengths.

Figure 4.17: A Strange Relationship
`Proofaid/Strange.T [M]`

With this we can show that $EB/AB = EF/AC$, but we do not know the length of EF in terms of anything else, so it is still not clear what to do.

If we construct this segment EF and again use **Geometer** to test the diagram we obtain these interesting observations:

```
Equal angles:
    (A E D) (D E F)
Equal length segments:
    [C D] [D F]
```

We would like to express the length of EF in terms of some other lengths, and if we can show that CD is indeed equal to DF, perhaps that could be of use. In fact, if we extend segments DE and AC to meet at G, we will have an isosceles triangle $\triangle AGE$, and it looks like $CG = EF$.

These hints provide enough information to find the proof.

4.14.2 Calculation of the Relationship

Label the lengths $AB = c$, $AE = x$ (and therefore $EB = c - x$), $EF = w$, $CG = y$ as shown in figure 4.17. Clearly $\triangle AGE$ is isosceles since its altitude AD bisects the angle at A. Since $EF \parallel AG$, and those parallel lines are cut by transversal CF, we know that $\angle DEF = \angle DGC$. We also know that $\angle CDG = \angle FDE$ since they are vertical angles. Thus $\triangle DCG \cong \triangle DFE$ so $y = w$.

Since $y = w$ and $b + y = AG = AE = x$, $w = x - b$. Since $EF \parallel AC$, $\triangle ABC \sim \triangle EBF$, so we know that $AC/AB = EB/EF$. This gives:

$$\frac{b}{c} = \frac{w}{c - x} = \frac{x - b}{c - x}.$$

Thus $b(c - x) = c(x - b)$ and a little algebra gives us the result we need:

$$x = \frac{2bc}{b + c}.$$

There are of course other ways to do this. For example, you can compare the areas of the triangles: $\mathcal{A}(\triangle ABC) = \mathcal{A}(\triangle CAD) + \mathcal{A}(\triangle BAD)$, and the areas of each can be expressed as $\mathcal{A}(\triangle ABC) = bc \sin(\angle CAB)/2$, et cetera. A brute-force calculation using analytic geometry also works.

4.15 ♦♦ Sum of Powers

If M is a point on a circle, and $ABCD$ is a square inscribed in that circle, show that $MA^2 + MB^2 + MC^2 + MD^2$ is independent of the position of M. Suppose the square is replaced by a regular n-gon $A_1 A_2 \ldots A_n$. Let $S_m = \sum_i MA_i^m$. For what integers m is S_m independent of the position of M?
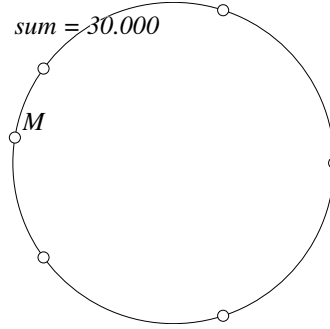
Figure 4.18: Sum of Powers of Lengths
`Proofaid/Pent4.T` [M]

4.15.1 Geometer Analysis

Clearly, it does not matter which square or n-gon is chosen, since we plan to rotate the point M to all possible positions relative to the polygon. Similarly, it does not matter how big a circle is chosen, since if we multiply the size of the circle by a constant α, all of the lengths MA_i will be multiplied by the same α, so the sum S_m will be multiplied by the constant α^m. In case we decide to approach the proof later using analytic geometry, it is probably best to select a circle of radius 1 so that it will be easy to express the vertices of the polygons in terms of sine and cosine without any extraneous constants.

Thus, our **Geometer** diagram might as well have one of the points of the n-gon at $(1, 0)$ and the others equally spaced around the circle. The coordinates of the n points will be $(\cos((i/n)360°), \sin((i/n)360°))$, where $0 \leq i < n$. Since we will want to experiment with different polygons and different powers of m, the **Geometer** diagram should be structured so that it is easy to modify. Here is a reasonable way to do so that generates the diagram in figure 4.18 (where the polygon is a regular pentagon).

```
.geometry "version 0.51";
m = .f.rpn(4.000000);
N = .f.rpn(5.000000);
ang = .f.rpn(360.000000, N, .div);
ctr = .v.ff(0.000000, 0.000000, .in);
pt = .v.ff(1.000000, 0.000000, .in);
c1 = .c.vv(ctr, pt);
M = .vonc(c1, -0.461465, 0.887158, .red, "M");
x0 = .f.rpn(ang, 0.000000, .mul, .cos);
y0 = .f.rpn(ang, 0.000000, .mul, .sin);
v0 = .v.ff(x0, y0);
x1 = .f.rpn(ang, 1.000000, .mul, .cos);
y1 = .f.rpn(ang, 1.000000, .mul, .sin);
v1 = .v.ff(x1, y1);
x2 = .f.rpn(ang, 2.000000, .mul, .cos);
y2 = .f.rpn(ang, 2.000000, .mul, .sin);
v2 = .v.ff(x2, y2);
x3 = .f.rpn(ang, 3.000000, .mul, .cos);
y3 = .f.rpn(ang, 3.000000, .mul, .sin);
v3 = .v.ff(x3, y3);
x4 = .f.rpn(ang, 4.000000, .mul, .cos);
y4 = .f.rpn(ang, 4.000000, .mul, .sin);
v4 = .v.ff(x4, y4);
d0 = .f.vv(v0, M);
d1 = .f.vv(v1, M);
d2 = .f.vv(v2, M);
d3 = .f.vv(v3, M);
d4 = .f.vv(v4, M);
s = .f.rpn(d0, .log, m, .mul, .exp,
        d1, .log, m, .mul, .exp,
        d2, .log, m, .mul, .exp,
        d3, .log, m, .mul, .exp,
        d4, .log, m, .mul, .exp,
        .add, .add, .add, .add, "sum");
```

The **Geometer** code works by calculating the x and y coordinates of each of the regularly-spaced points and then creating those points. The values d0, d1, et cetera, are the distances between each point and the variable point M, and the final calculation of s raises each distance to the power m and adds them all up. Since the s value in the **Geometer** diagram has a name, it is presented on the computer screen and updated with each movement of the point M.

As we view the diagram in **Geometer**, we see that no matter where we move the red point M, the value of $S_4 = \sum_i MA_i^4$ is displayed, and we see that it seems to have the constant value 30.0000.

In this **Geometer** code, m is the exponent, set to 4 in this code fragment. We can try different values of m simply by changing the second line. If $m = 2$, we find that the

sum is a constant 10.0000. If $m = 6$, the sum seems to be a constant 100.0000. In fact, all even powers of m seem to yield constant values of the sum.

The following version of the code (found in `Proofaid/Pent4m.T`) is one line longer, but it uses macro calls to generate points. It is easier to modify, however, than the version above.

```
.geometry "version 0.60";
m = .f.rpn(4.000000);
N = .f.rpn(5.000000);
ang = .f.rpn(360.000000, N, .div);
ctr = .v.ff(0.000000, 0.000000, .in);
pt = .v.ff(1.000000, 0.000000, .in);
c1 = .c.vv(ctr, pt);
M = .vonc(c1, -0.434232, 0.900801, .red, "M");
.macro .vertex v(.flt f)
{
    x = .f.rpn(ang, f, .mul, .cos);
    y = .f.rpn(ang, f, .mul, .sin);
    .return vv = .v.ff(x, y);
}
f0 = .f.rpn(0.000000);
f1 = .f.rpn(1.000000);
f2 = .f.rpn(2.000000);
f3 = .f.rpn(3.000000);
f4 = .f.rpn(4.000000);
v0 = v(f0);
v1 = v(f1);
v2 = v(f2);
v3 = v(f3);
v4 = v(f4);
d0 = .f.vv(v0, M);
d1 = .f.vv(v1, M);
d2 = .f.vv(v2, M);
d3 = .f.vv(v3, M);
d4 = .f.vv(v4, M);
s = .f.rpn(d0, .log, m, .mul, .exp,
        d1, .log, m, .mul, .exp,
        d2, .log, m, .mul, .exp,
        d3, .log, m, .mul, .exp,
        d4, .log, m, .mul, .exp,
        .add, .add, .add, .add, "sum");
```

How about different regular polygons? We will construct a file by modifying the macro-based file we used for the pentagon. The value `N` in the third line is the number of vertices, so if we would like to look at a regular heptagon (7-sided figure), we just change the value of `N` to 7. We also need to add two vertices, `v5` and `v6`, two more distances, `d5` and `v6`, and to modify the formula for the sum at the end of the file.

Here is the **Geometer** code for a regular heptagon (found in `Proofaid/Hept7m.T`, and we find that for $m = 4$, the result seems to be a constant 42.0000. Again, we find that different even values of m generate constant sums and odd values do not. On the CD you can also find a file `Proofaid/hept7.T` which does the same thing, but is based on the non-macro version `Proofaid/pent4.T`.

```
.geometry "version 0.60";
m = .f.rpn(4.000000);
N = .f.rpn(7.000000);
ang = .f.rpn(360.000000, N, .div);
ctr = .v.ff(0.000000, 0.000000, .in);
pt = .v.ff(1.000000, 0.000000, .in);
c1 = .c.vv(ctr, pt);
M = .vonc(c1, -0.603538, 0.797334, .red, "M");
.macro .vertex v(.flt f)
{
    x = .f.rpn(ang, f, .mul, .cos);
    y = .f.rpn(ang, f, .mul, .sin);
    .return vv = .v.ff(x, y);
}
f0 = .f.rpn(0.000000);
f1 = .f.rpn(1.000000);
f2 = .f.rpn(2.000000);
f3 = .f.rpn(3.000000);
f4 = .f.rpn(4.000000);
f5 = .f.rpn(5.000000);
f6 = .f.rpn(6.000000);
v0 = v(f0);
v1 = v(f1);
v2 = v(f2);
v3 = v(f3);
v4 = v(f4);
v5 = v(f5);
v6 = v(f6);
d0 = .f.vv(v0, M);
d1 = .f.vv(v1, M);
d2 = .f.vv(v2, M);
d3 = .f.vv(v3, M);
d4 = .f.vv(v4, M);
d5 = .f.vv(v5, M);
d6 = .f.vv(v6, M);
s = .f.rpn(d0, .log, m, .mul, .exp,
        d1, .log, m, .mul, .exp,
        d2, .log, m, .mul, .exp,
        d3, .log, m, .mul, .exp,
        d4, .log, m, .mul, .exp,
        d5, .log, m, .mul, .exp,
        d6, .log, m, .mul, .exp,
        .add, .add, .add, .add, .add,
        .add, "sum");
```

By making modifications like those above to the **Geometer** file, we can construct a table of values for various regular polygons and various even exponents. In the table below, a "$*$" is placed when the sum is *not* constant.

Notice that the column entries for $m = 0$ are trivial to calculate, so they are included as well. Finally, what seems to be important is not that the inscribed figure is a polygon but rather that it consists of n equally-spaced points, we can add the data for $n = 1$ and $n = 2$, and both of these observations make the pattern clearer.

$n\backslash m$	0	2	4	6	8	10	12	14
1	1	*	*	*	*	*	*	*
2	2	4	*	*	*	*	*	*
3	3	6	18	*	*	*	*	*
4	4	8	24	80	*	*	*	*
5	5	10	30	100	350	*	*	*
6	6	12	36	120	420	1512	*	*
7	7	14	42	140	490	1764	6468	*

Each row appears to consist of the following multiples of n: 1, 2, 6, 20, 70, 252 and 924, when they are constant. These are just the values of $\binom{2k}{k}$ in Pascal's triangle, so the general formula appears to be:

$$S_m = \sum_{i=0}^{n-1} MA_i^m = n\binom{m}{m/2},$$

where the formula works for all even values of m less than $2n$.

4.15.2 Proof of the Result

For the case where $m = 2$, it is easy to prove the result for an arbitrary regular polygon with n sides. On the unit circle, the vertices have coordinates $(\cos\theta_k, \sin\theta_k)$, where $0 \le k < n$ and $\theta_k = (k/n)360°$.

Let the point M have coordinates $(\cos\alpha, \sin\alpha)$. Then the sum S_2 is given by:

$$S_2 = \sum_{k=0}^{n-1} (\cos\theta_k - \cos\alpha)^2 + (\sin\theta_k - \sin\alpha)^2 \qquad (4.9)$$

If we multiply out the right hand side of equation 4.9 and simplify the result, we obtain:

$$
\begin{aligned}
S_2 &= \sum_{k=0}^{n-1} \cos^2\theta_k - 2\cos\theta_k\cos\alpha + \cos^2\alpha + \sin^2\theta_k - 2\sin\theta_k\sin\alpha + \sin^2\alpha \\
&= \sum_{k=0}^{n-1} \cos^2\theta_k + \sin^2\theta_k - 2(\cos\theta_k\cos\alpha + \sin\theta_k\sin\alpha) + \cos^2\alpha + \sin^2\alpha \\
&= \sum_{k=0}^{n-1} 2 - 2(\cos\theta_k\cos\alpha + \sin\theta_k\sin\alpha) \\
&= 2n - 2\left(\cos\alpha\sum_{k=0}^{n-1}\cos\theta_k + \sin\alpha\sum_{k=0}^{n-1}\sin\theta_k\right).
\end{aligned}
$$

If we can show that the sums $\sum_{k=0}^{n-1}\cos\theta_k$ and $\sum_{k=0}^{n-1}\sin\theta_k$ are zero, we are done. If we recall that

$$\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2} \qquad \text{and} \qquad \sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}$$

then the sums that interest us become a pair of geometric series. For example:

$$\sum_{k=0}^{n-1} \cos \theta_k = \frac{1}{2} \left(\sum_{k=0}^{n-1} e^{(k/n)360^\circ} + \sum_{k=0}^{n-1} e^{-(k/n)360^\circ} \right). \tag{4.10}$$

Both series on the right hand side of equation 4.10 begin with 1 and have a ratio of either $e^{(1/n)360^\circ}$ or $e^{-(1/n)360^\circ}$. Both series look like $1 + r + \ldots + r^{n-1} = (1 - r^n)/(1 - r)$, and in both cases $r^n = 1$, so both sums are zero. The same thing happens for the sum of the sines. Therefore the sum S_2 is independent of α.

4.15.3 Alternative Proof with Vectors

The same result can be obtained with vectors. Let A_k for $0 \le k < n$ be the vertices of a regular n-gon inscribed in a circle with center O. If M is any point on the circle, then $\overrightarrow{MA_k} = \overrightarrow{MO} + \overrightarrow{OA_k}$. Then:

$$S_2 = \sum_{k=0}^{n-1} (\overrightarrow{MA_i})^2 = \sum_{k=0}^{n-1} (\overrightarrow{MO} + \overrightarrow{OA_i})^2.$$

Expanding the expression on the right, we obtain:

$$\begin{aligned}
S_2 &= \sum_{k=0}^{n-1} (\overrightarrow{MO}^2 + 2\,\overrightarrow{MO} \cdot \overrightarrow{OA_i} + \overrightarrow{OA_i}^2) \\
&= \sum_{k=0}^{n-1} \overrightarrow{MO}^2 + \sum_{k=0}^{n-1} 2\,\overrightarrow{MO} \cdot \overrightarrow{OA_i} + \sum_{k=0}^{n-1} \overrightarrow{OA_i}^2 \\
&= n\,\overrightarrow{MO}^2 + 2\,\overrightarrow{MO} \cdot \sum_{k=0}^{n-1} \overrightarrow{OA_i} + \sum_{k=0}^{n-1} \overrightarrow{OA_i}^2
\end{aligned}$$

Since $\overrightarrow{MO}^2$ and $\overrightarrow{A_k}^2$ are both constant and equal to r^2, where r is the radius of the circle, this is equivalent to:

$$S_2 = nr^2 + nr^2 + 2\,\overrightarrow{MO} \cdot \sum_{k=0}^{n-1} \overrightarrow{OA_i}$$

Clearly, $\sum_{k=0}^{n-1} \overrightarrow{OA_i} = \overrightarrow{0}$, since if it did point in any direction, by symmetry, it would have to point the same amount in every one of n equally-spaced directions. Thus the sum $S_2 = 2nr^2$ which is what we obtained in the previous section with $r = 1$.
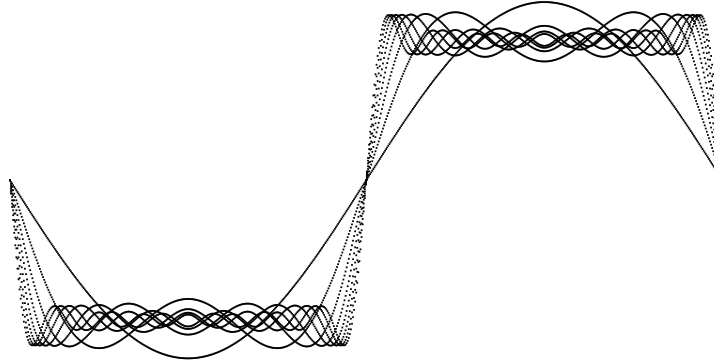
Chapter 5

More Useful Theorems

The results in this chapter are generally not covered in high school geometry courses, but they are all very useful tools for solving other geometric problems. In this chapter we will continue to use computer techniques to solve problems, so if you have not already done so, it may be useful to read Chapter 4 before continuing.

5.1 Ceva's Theorem

Ceva's (pronounced "cheva's") theorem is perhaps the most useful geometric theorem that is not taught in a standard high school course.

Theorem 5.1 (Ceva's Theorem) *Given any triangle* $\triangle ABC$ *with a point P a inside it*[1] *(see figure 5.1), if the lines from A, B and C through P intersect the opposite sides*

---

[1] In fact, it is not difficult to show that P can be outside the triangle as well, as long as it does not lie on any of the (infinite) lines that make up the edges of the triangle.

of the triangle at points X, Y and Z, respectively, then:

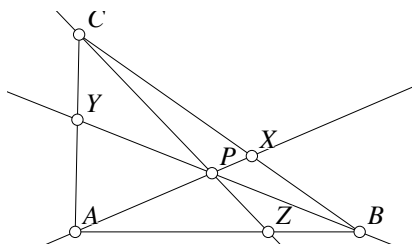$$\frac{AZ}{BZ} \cdot \frac{BX}{CX} \cdot \frac{CY}{AY} = 1. \tag{5.1}$$

The converse is also true (although we do not prove it here): If three lines emanate from the vertices of a triangle and intersect the opposite sides such that the ratios of the lengths into which the sides are divided satisfy equation 5.1 then the three lines are concurrent. Ceva's theorem is also true if the ratios in equation 5.1 are directed ratios (see Section 5.2.1).

Figure 5.1: Ceva's Theorem
Bigtheorems/Ceva.T [M]

This is a fairly amazing theorem, since it applies to *any* triangle and to almost any point. If you have trouble proving this theorem, you may find it instructive before reading on to prove it in certain special cases, such as when P is the centroid, incenter, circumcenter, or orthocenter.

Any line from the vertex of a triangle through a point on the opposite side of that triangle is called a **cevian**, so Ceva's theorem refers to three cevians that happen to pass through the same point.

5.1.1   Geometer Analysis of Ceva's Theorem

If we make no progress on the proof, we can always draw a **Geometer** diagram and see what relationships hold when we test that diagram. After moving around the points, the only relationships that **Geometer** notices are a large number of triangle area ratios that seem to hold:

```
Triangle area ratios:
    [A Z P] x [A C P] = [B Z P] x [B C P]
    [A Z P] x [A Y P] = [B Z C] x [B C Y]
    [A Z C] x [C X A] = [B Z P] x [B X P]
    [A Z C] x [A C P] = [B Z C] x [B C P]
    [A Z C] x [B Z P] = [A Z P] x [B Z C]
    [A B P] x [A C P] = [B X P] x [C X P]
    [A B P] x [C Y P] = [B C P] x [A Y P]
    [A B Y] x [C Y P] = [B C Y] x [A Y P]
    [A B Y] x [B C P] = [A B P] x [B C Y]
    [A B X] x [C X P] = [B X P] x [C X A]
    [A B X] x [A C P] = [A B P] x [C X A]
    [A B X] x [C Y P] = [A B Y] x [C X P]
    [A B C] x [C X P] = [B C P] x [C X A]
    [A B C] x [C Y P] = [B C Y] x [A C P]
    [A B C] x [B Z P] = [A B P] x [B Z C]
    [A B C] x [A Z P] = [A B P] x [A Z C]
    [A B C] x [A Y P] = [A B Y] x [A C P]
    [A B C] x [B X P] = [A B X] x [B C P]
```

After examining a few, we notice that they all are based on the same idea. Relationships like that shown in the fifth line are the easiest to see:

```
[A Z C] x [B Z P] = [A Z P] x [B Z C]
```

This basically says that $\triangle AZC$ and $\triangle BZC$ have the same ratio as $\triangle AZP$ and $\triangle BZP$. This is obvious because the ratios of the areas are the same as the ratio into which Z divides the base AB. Once we see this, it is obvious why the first ratio in the list is true: $\triangle AZC$ is composed of the two triangles $\triangle ACP$ and $\triangle AZP$ and $\triangle BZC$ can be similarly decomposed.

Thus the ratios of the subdivisons of the sides are the same as the ratio of the areas of the triangles $\triangle ACP$, $\triangle BCP$ and $\triangle ABP$. Since Ceva's theorem concerns these ratios, we are probably on the right track. Using this information the proof is simple to construct.

5.1.2 Proof of Ceva's Theorem

Here is a proof, derived from the observations above:

Relative to the base AB, the triangles $\triangle AZP$ and $\triangle BZP$ have the same altitude, their areas are proportional to $AZ : BZ$. The areas of triangles $\triangle AZC$ and $\triangle AZC$ are similarly related so we obtain:

$$\frac{AZ}{BZ} \quad = \quad \frac{\mathcal{A}(\triangle AZP)}{\mathcal{A}(\triangle BZP)} \quad = \quad \frac{\mathcal{A}(\triangle AZC)}{\mathcal{A}(\triangle BZC)} \tag{5.2}$$

Triangle $\triangle AZC$ is composed of triangles $\triangle AZP$ and $\triangle ACP$ and $\triangle BZC$ can be similarly decomposed, so their areas satisfy:

$$\mathcal{A}(\triangle AZC) \quad = \quad \mathcal{A}(\triangle AZP) + \mathcal{A}(\triangle ACP) \tag{5.3}$$
$$\mathcal{A}(\triangle BZC) \quad = \quad \mathcal{A}(\triangle BZP) + \mathcal{A}(\triangle BCP). \tag{5.4}$$

Rearranging equations 5.3 and 5.4, we obtain:

$$\mathcal{A}(\triangle ACP) \quad = \quad \mathcal{A}(\triangle AZC) - \mathcal{A}(\triangle AZP) \tag{5.5}$$
$$\mathcal{A}(\triangle BCP) \quad = \quad \mathcal{A}(\triangle BZC) - \mathcal{A}(\triangle BZP). \tag{5.6}$$

Combining equations 5.2, 5.5 and 5.6 by using the fact that if $a/b = c/d = e$ where b and d are different then $(a - c)/(b - d) = e$, we obtain:

$$\frac{AZ}{BZ} \quad = \quad \frac{(\triangle ACP)}{(\triangle BCP)}. \tag{5.7}$$

There is nothing special about the selection of AB as a base of the triangle, and similar calculations give us:

$$\frac{BX}{CX} \quad = \quad \frac{(\triangle ABP)}{(\triangle ACP)} \tag{5.8}$$

$$\frac{CY}{AY} \quad = \quad \frac{(\triangle BCP)}{(\triangle ABP)}. \tag{5.9}$$

Now multiply together equations 5.7, 5.8 and 5.9 to obtain Ceva's theorem:

$$\frac{AZ}{BZ} \cdot \frac{BX}{CX} \cdot \frac{CY}{AY} = \frac{(\triangle ACP)}{(\triangle BCP)} \cdot \frac{(\triangle ABP)}{(\triangle ACP)} \cdot \frac{(\triangle BCP)}{(\triangle ABP)} = 1. \qquad (5.10)$$

It is a good exercise to show how this proof can be modified to show that Ceva's theorem still holds if P is exterior to the triangle $\triangle ABC$.

5.2 Menelaus' Theorem

Menelaus' theorem is similar to Ceva's theorem (see Section 5.1) but with the roles of points and lines reversed.
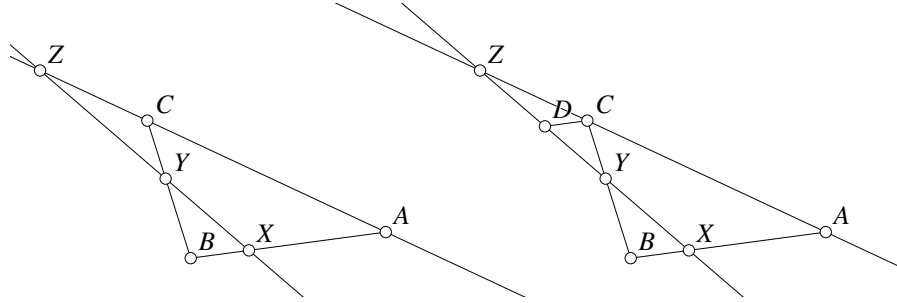

Figure 5.2: Menelaus' Theorem
`Bigtheorems/Menelaus.T [M]`

The theorem is illustrated in the diagram on the left of figure 5.1.

Theorem 5.2 (Menelaus's Theorem) *Let $\triangle ABC$ be an arbitrary triangle. If a line L that does not pass through A, B or C passes through the sides AB, BC and CA at X, Y and Z, respectively, we have:*

$$\frac{AX}{BX} \cdot \frac{BY}{CY} \cdot \frac{CZ}{AZ} = 1. \qquad (5.11)$$

We will not show it here, but like Ceva's theorem, the converse of Menelaus' theorem is also true: Given a triangle and points on the edges of that triangle whose distances satisfy the equation 5.11, then those three points lie on a line.

5.2.1 ♦ Directed Ratios

Notice that no matter how the line L is drawn, it must intersect at least one of the triangle segments (and sometimes all three of them, but never exactly two) outside the triangle. For this reason Menelaus' theorem is often stated with the product of the ratios in equation 5.11 equal to -1, where the ratios are assumed to be directed ratios.

A **directed ratio** of two segments on the same line has the numeric value of the ratio of their lengths, but is positive if they are directed in opposite directions and negative if they are directed in the same direction. In figure 5.2, the directed ratios AX/BX and BY/CY would be positive, but CZ/AZ would be negative.

If all three intersection points of the line L were outside the triangle, all three directed ratios would be negative, again yielding a negative product of ratios.

Notice that Ceva's theorem is also true when the ratios are assumed to be directed: if P is outside the triangle, then two of the directed ratios in equation 5.1 will be negative and one positive, so the statement of Ceva's theorem is correct with $+1$ as the product of the directed ratios.

5.2.2 Geometer Analysis of Menelaus' Theorem

If the diagram on the left of figure 5.2 is subjected to the **Geometer**'s test for relationships, nothing is found. This is not too surprising, since there are so few constraints on the geometry. When this happens, it is likely that the addition of one or more auxiliary lines will help, so a good strategy is to add "likely" lines and rerun **Geometer**'s test.

Recalling the proof of Ceva's theorem, a first approach might be to draw the segments CX, AY and BZ. If you do this, the test will discover a list of triangle area relationships similar to those encountered when we tested the Ceva diagram, but although the set looks promising, it does not seem to yield enough useful ratios to solve the problem.

When ratios play a big part in the statement of a theorem, it is often a good idea to look for similar triangles since every pair of similar triangles in a figure can generate a whole list of preserved ratios. There are no similar triangles in the diagram on the left of figure 5.2, but if we construct the line through C parallel to AB that intersects line XY at D, suddenly we have a few. In particular, $\triangle CDZ \sim \triangle AXZ$ and $\triangle BXY \sim \triangle CDY$. These are nice, since those triangles seem to contain all the lengths of segments that appear in the statement of Menelaus' theorem in equation 5.11. There is nothing special about drawing the line through C parallel to the opposite side; a similar line could have been drawn through A or B.

Testing the **Geometer** diagram with the added segment DC yields, as expected, an interesting set of ratios that are preserved:

```
Ratios:
    [A Z] x [Z D] = [C Z] x [X Z]
    [A C] x [Z D] = [C Z] x [X D]
    [A C] x [X Z] = [A Z] x [X D]
    [B Y] x [Y D] = [C Y] x [X Y]
    [B C] x [Y D] = [C Y] x [X D]
    [B C] x [X Y] = [B Y] x [X D]
    [B X] x [Y D] = [X Y] x [C D]
    [B X] x [C Y] = [B Y] x [C D]
    [A X] x [Z D] = [X Z] x [C D]
    [A X] x [C Z] = [A Z] x [C D]
```

A quick examination shows us that all these ratios can be derived from the pair of similar triangles listed in the previous paragraph. From here it is relatively easy to construct a proof of Menelaus' theorem by fiddling around with the ratios above.

5.2.3 Proof of Menelaus' Theorem

Construct the line parallel to AB through C meeting line XY at D as in figure 5.2. Since $CD \parallel AB$, it is easy to show that $\triangle CDZ \sim \triangle AXZ$ and $\triangle BXY \sim \triangle CDY$. Since those triangles are similar, we have:

$$\frac{CZ}{AZ} \quad = \quad \frac{ZD}{XZ} \tag{5.12}$$

$$\frac{BY}{CY} \quad = \quad \frac{XY}{YD} \tag{5.13}$$

$$\frac{AX}{CD} \quad = \quad \frac{XZ}{ZD} \tag{5.14}$$

$$\frac{CD}{BX} \quad = \quad \frac{YD}{XY} \tag{5.15}$$

If we multiply together equations 5.12 through 5.15, we obtain:

$$\frac{CZ}{AZ} \cdot \frac{BY}{CY} \cdot \frac{AX}{CD} \cdot \frac{CD}{BX} = \frac{ZD}{XZ} \cdot \frac{XY}{YD} \cdot \frac{XZ}{ZD} \cdot \frac{YD}{XY},$$

which reduces to equation 5.11 and proves the theorem.

5.3 Alternate Proofs of Menelaus and Ceva
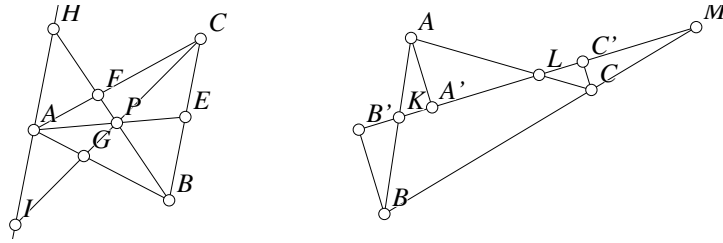
Figure 5.3: Menelaus' and Ceva's Theorems
`Bigtheorems/Ceva1.T, Bigtheorems/Menelaus1.T [P]`

When clever constructions are made, the proofs of Ceva's and Menelaus' theorems are easier to find. Figure 5.3 shows two of those constructions.

On the left, to prove Ceva's theorem, construct a line through point A that is parallel to BC and extend the segments BF and CG to meet that new line. For a quicker proof of Menelaus' theorem, use the diagram on the right. If line KLM cuts the three sides

of the triangle, construct perpendiculars from A, B and C to line KLM meeting it atA', B' and C', respectively.

Try to find alternate proofs of Menelaus' and Ceva's theorems based on these new constructions. Those proofs are not difficult, but the solutions can be found in the files `Ceva1.T` and `Menelaus1.T`

5.4 Using Menelaus' and Ceva's Theorem

If you would like to practice using Ceva's theorem, Chapter 7 on triangle centers contains a surprising number of examples where Ceva's theorem leads to a proof of the existence of some center.

If you would like to practice with Menelaus' theorem, look in Chapter 9 on projective geometry where Menelaus' theorem is used to prove the Euclidean version of a number of projective geometry theorems.

5.5 Ptolemy's Theorem

Ptolemy's theorem provides a powerful method for calculating lengths in a quadrilateral whose vertices happen to be concyclic. Conversely, if you know the lengths of the sides and diagonals, it can be used to determine if the quadrilateral is concyclic.
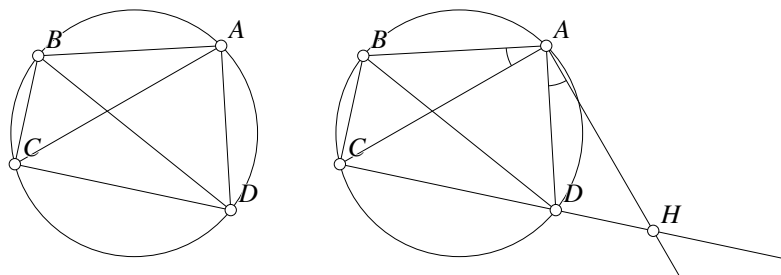
Figure 5.4: Ptolemy's Theorem
`Bigtheorems/Ptolemy.T` [M]

The theorem is illustrated in the diagram on the left of figure 5.4.

Theorem 5.3 (Ptolemy's Theorem) *If A, B, C and D are four concyclic points in clockwise or counterclockwise order on the circle, then the following relationship holds and conversely:*

$$AB \cdot CD + BC \cdot AD = AC \cdot BD. \tag{5.16}$$

Notice that Ptolemy's theorem is a generalization of the pythagorean theorem; if the quadrilateral $ABCD$ is a rectangle, Ptolemy's theorem is equivalent to the pythagorean theorem.

5.5.1 Geometer Analysis of Ptolemy's Theorem

Ptolemy's theorem is not easy to prove, and if we simply draw the basic figure seen on the left of figure 5.4 and test it with **Geometer**'s relationship finding commands, it only tells us about a few obvious angles that are equal (since they are inscribed in the same circle) or supplementary (since they are opposite angles in a cyclic quadrilateral). It seems clear that we will have to use ratios and therefore similar triangles to make any progress, so we need to convert equation 5.16 into ratios, but what are we going to do about that annoying "+" sign?

You probably cannot avoid looking at the answer, since it is drawn on the right of the figure, but here is why it (or any of the 7 other equivalent constructions) is a good idea. To convert the product form of equation 5.4 into a ratio form, we need terms like, for example, AB/AC. Thus we would like to get AC and AD to be corresponding parts of similar triangles. If we imagine rotating $\triangle ABD$ until its angle at A matches a new triangle, it is clear that a rotation by $\angle CAB$ will align AB and AC. Of course this will rotate AD out of the circle by the same angle $\angle CAB$. Therefore, a reasonable way to begin is to construct a line from A making angle $\angle CAB$ with the line AD, as shown in the right side of figure 5.4. The new line must intersect the line CD at some point H outside the circle.

Now when **Geometer** tests the diagram, we obtain a far more interesting set of relations, including:

```
Equal angles:
    (B C A) (C H A) (A D B)
Ratios:
    [C H] x [A D] = [B D] x [A H]
    [B C] x [A H] = [D H] x [A C]
    [A B] x [A H] = [A D] x [A C]
    [A B] x [C H] = [A C] x [B D]
    [A B] x [D H] = [B C] x [A D]
```

The equal angle information is encouraging, since it shows clearly that $\triangle ABD \sim \triangle ACH$, and the ratios, if we can show that they hold due to similarities of triangles, include almost all of the lengths that appear in equation 5.4. We can also see from the ratios that $\triangle ABC$ must be similar to $\triangle ADH$. At first it is a bit annoying to see lengths like CH and DH in the list, until we realize that $CH = CD + DH$! This pulls in the missing CD, *and* gets a "+" sign into the equations.

Using these observations, the theorem is easy to prove.

5.5.2 Proof of Ptolemy's Theorem

This is just a proof of one direction of the theorem, namely, that if the points are concyclic, the relationship in equation 5.16 holds.

Construct a line from A making angle $\angle CAB$ with line AD that intersects the line CD at H.

We know that $\angle ABD = \angle ACD$ since they are inscribed in the same circle, and $\angle DAB = \angle HAC$ since they are both equal to $\angle DAC + \angle CAB$. Therefore by AA similarity, $\triangle DAB \sim \triangle HAC$.

We know that $\angle ABC$ and $\angle ADC$ are supplementary since they are opposite angles in an inscribed quadrilateral, and it is obvious that $\angle ADC$ and $\angle AHD$ are also supplementary, so $\angle ADH = \angle ABC$. This, and the fact that $\angle BAC = \angle DAH$ allows us to conclude by AA similarity that $\triangle ADH \sim \triangle ABC$.

We know that:

$$CD \quad + \quad DH \quad = \quad CH \tag{5.17}$$

and the two sets of similar triangles show us that:

$$\frac{BD}{CH} \quad = \quad \frac{AB}{AC} \quad = \quad \frac{AD}{AH} \tag{5.18}$$

$$\frac{AD}{AB} \quad = \quad \frac{DH}{BC} \quad = \quad \frac{AH}{AC} \tag{5.19}$$

To get rid of the DH and CH in equation 5.17, we can substitute their values obtained from equations 5.18 and 5.19:

$$CD + \frac{BC \cdot AD}{AB} = \frac{AC \cdot BD}{AB} \tag{5.20}$$

If we multiply equation 5.20 by AB, we obtain:

$$AB \cdot CD + BC \cdot AD = AC \cdot BD,$$

which proves the theorem.

5.5.3 ♦ Additional Notes on Ptolemy's Theorem

It is possible to prove the theorem using brute force by listing the lengths of all six sides and relating them with the law of sines. Using the fact that various sets of angles add to $180°$ (because they form a triangle or are opposite angles in the cyclic quadrilateral), we can expand the equations using the formulas for the sine and cosine of a sum of angles. After about ten lines of calculation, Ptolemy's formula drops out.

There is also a beautiful derivation of Ptolemy's theorem based on inversion in a circle. See Section 8.8.

5.6 Heron's Formula

Heron's formula provides an easy way to calculate the area of an arbitrary triangle given the lengths of its sides:

Theorem 5.4 (Heron's Formula) *Given an arbitrary triangle $\triangle ABC$, let a, b and c be the lengths of the sides opposite the vertices A, B and C, respectively. Let $s = (a + b + c)/2$ be the semiperimeter of $\triangle ABC$. Then the area of the triangle is given by the formula:*

$$\mathcal{A}(\triangle ABC) = \sqrt{s(s-a)(s-b)(s-c)}, \tag{5.21}$$

There is a straightforward but fairly messy trigonometric proof. Begin with the observation that $\mathcal{A}(\triangle ABC) = (ab \sin \theta)/2$, where $\theta = \angle ACB$. We also know, from the law of cosines, that $c^2 = a^2 + b^2 - 2bc \cos \theta$. Solve these equations for $\sin \theta$ and $\cos \theta$, and using the fact that $\sin^2 \theta + \cos^2 \theta = 1$, do a lot of algebra, and you will eventually arrive at the result.
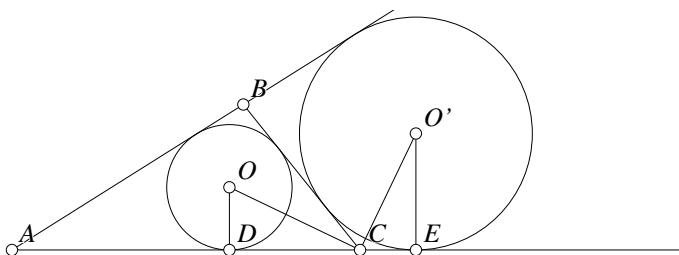
5.6.1 Geometer Analysis of Heron's Formula

Figure 5.5: Heron's Theorem
`Bigtheorems/Heron.T [M]`

A more geometric proof would be nice, but it is a bit difficult to find the correct figure. Recall that Section A.6.8 contains a lot of interesting formulas relating expressions like s, $s - a$, $s - b$ and $s - c$ to the area of a triangle. This observation suggests that perhaps it would be useful to consider a diagram that includes some combination of incircles and excircles. The diagram that does the trick is shown in figure 5.5. The lines OD and $O'E$ are perpendiculars dropped from the centers of the circles to the line AC.

If we use **Geometer** to test that diagram, all sorts of relationships appear, but the ones that are most interesting are these:

```
Right angles:
    (O C O') (A E O') (A D O)
Equal angles:
    (A C O) (B C O) (E O' C)
Ratios:
    [C E] x [C O] = [O D] x [C O']
    [C D] x [C O'] = [O' E] x [C O]
    [C D] x [C E] = [O D] x [O' E]
    [A D] x [O' E] = [A E] x [O D]
```

Since $\angle ACO = \angle EO'C$ and OD and $O'E$ form right angles with AC, $\triangle COD \sim \triangle O'CE$. Using the notation from Section A.6.8, we have $DC = s - c$, $CE = s - b$, $OD = r$ and $O'E = r_a$. This, plus a little algebra with the ratios is enough to yield the result.

5.6.2 Proof of Heron's Formula

If we can show that $\angle OCO' = 90°$ then we will know that $\angle ACO = \angle EO'C$ since both $\angle ACO$ and $\angle EO'C$ are complementary to $\angle O'CE$. But $\angle DCB + \angle BCE = 180°$ and those angles are bisected by CO and CO' so $\angle OCO' = 90°$.

Thus we know by AA that $\triangle COD \sim \triangle O'CE$. From this similarity, we have:

$$\frac{OD}{DC} = \frac{CE}{O'E} \quad \text{or} \quad \frac{\text{r}}{\text{s} - \text{c}} = \frac{\text{s} - \text{b}}{\text{r}_\text{a}}. \tag{5.22}$$

We also know from Section A.6.8 that:

$$\mathcal{A}(\triangle ABC) = r_a(s - a) = rs. \tag{5.23}$$

From equation 5.23 we see that $r = \mathcal{A}(\triangle ABC)/s$ and $r_a = \mathcal{A}(\triangle ABC)/(s - a)$. If we substitute these values into equation 5.22, we obtain:

$$\frac{\mathcal{A}(\triangle ABC)}{s(s - c)} = \frac{(s - b)(s - a)}{\mathcal{A}(\triangle ABC)}$$

Cross-multiply to obtain:

$$(\mathcal{A}(\triangle ABC))^2 = s(s - a)(s - b)(s - c),$$

which is equivalent to Heron's formula.

Note: In a sense, the pythagorean theorem is a special case of Heron's formula. If c is the hypotenuse of a right triangle, then the area of that triangle is $ab/2$. If we substitute $s = (a + b + c)/2$ into Heron's formula, expand, and set it equal to $ab/2$, we obtain:

$$
\begin{aligned}
ab/2 &= (1/4)\sqrt{(2ab)^2 - (a^2 + b^2 - c^2)^2} \\
\frac{a^2b^2}{4} &= \frac{a^2b^2}{4} - \frac{(a^2 + b^2 - c^2)^2}{16} \\
0 &= a^2 + b^2 - c^2.
\end{aligned}
$$

5.6.3 Brahmagupta's Formula

Heron's formula looks a lot like another interesting formula called Brahmagupta's formula that applies to cyclic quadrilaterals. If a quadrilateral is inscribed in a circle and

has sides of lengths a, b, c and d, and if $s = (a + b + c + d)/2$ is its semiperimeter, then the area $\mathcal{A}$ of that quadrilateral is given by the formula:

$$\mathcal{A} = \sqrt{(s - a)(s - b)(s - c)(s - d)}.$$

We will not prove Brahmagupta's formula here, but it can be proved relatively easily with some algebra and the laws of sines and cosines. It can also be proved using Heron's formula on the two triangles formed if you extend two sides of the quadrilateral to meet outside the circle. Heron's formula is a special case of Brahmagupta's formula where two of the points coincide, so effectively $d = 0$.

5.7 The Radical Axis

In this section we will make more sophisticated use of **Geometer**'s features. We will begin by investigating the following question:

Given two circles, what is the set of points whose tangent lines to those two circles have equal length?
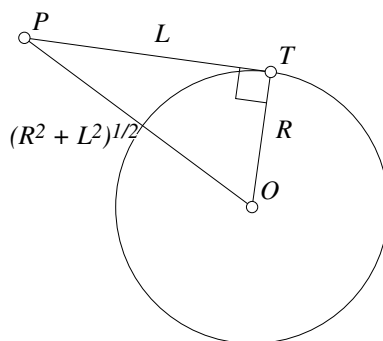

Figure 5.6: pythagorean Theorem
Bigtheorems/Rat1.D [M]

Assuming we do not have a clue what the answer is, we can construct a **Geometer** diagram that will give us a good idea. The diagram is based on the idea presented in figure 5.6. If P is a point outside a circle of radius R whose tangent distance to the circle is L, then the length of the segment PO is $\sqrt{R^2 + L^2}$. Thus, to find *all* points P with a tangent length of L to the circle, we need only construct any right triangle like the one in the figure and then draw a circle about O having the hypotenuse of that triangle as its radius.

To find the points satisfying our condition, we will build a diagram with the two circles but which in addition contains a pair of points the distance between which will be the L. We will copy this distance L on a tangent line to each circle, forming triangles that correspond to $\triangle POT$ in figure 5.6. The circles with the hypotenuses of those
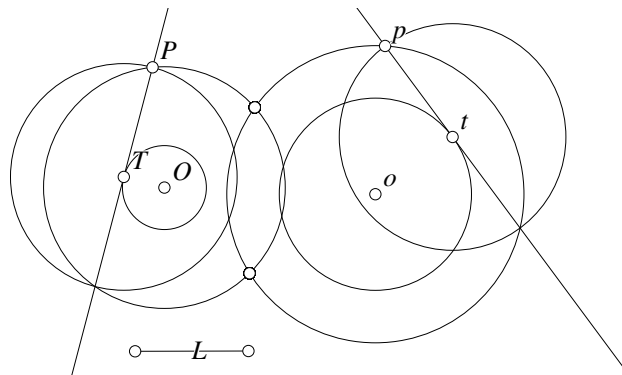
Figure 5.7: Radical Axis Diagram
Bigtheorems/Radical1.T [M]

triangles as radii will represent all points whose tangent distance to the circles is L. The points that interest us are at the intersections of those circles.

Figure 5.7 shows the result of this construction. The sizes and positions of the circles can be changed by adjusting O, o, T and t. The tangent lengths are equal: $TP = tp = L$. Two circles are drawn centered at O and o passing through P and p, respectively. The points of intersection of these last two circles are points without a name.
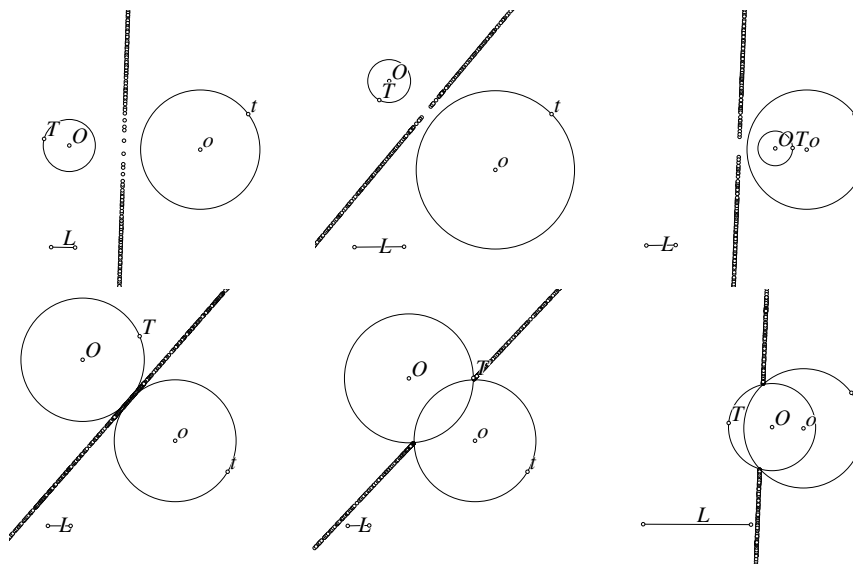

Figure 5.8: The Radical Axis
Bigtheorems/RadicalTest.T [M]

To eliminate much of the clutter for our experiments, change the colors of all the circles except for the original two and of the tangent lines and the points P and p to

be the invisible color. We also change the colors of the points of intersection to be the smear color. Then we move the endpoints of the line segment L and see what happens. We can do this repeatedly with different arrangements of the two original circles to get an idea of what sorts of possibilities occur. The results of six such experiments are illustrated in figure 5.8.

From our experiments, it appears that in every case the locus of points satisfying our conditions is a line or at least a part of a line which will be called the **radical axis** of the two circles. It is not too difficult to extend the definition so that the radical axis will be an entire line even if the two circles intersect. The problem with the definition above is that when P is inside a circle, it is impossible to draw a tangent line.

If we examine figures A.29 and A.30 we see that if a secant of the circle through P intersects the circle at points A and B, then $PA \cdot PB = PT^2$. Instead of looking for locations of P that make PT constant (or equivalently, make PT^2 constant), why not find P such that $PA \cdot PB$ is constant? Then we can use any secant we want. This product, $PA \cdot PB$, is called the **power of a point** P relative to the circle. The radical axis of two circles, then, is best defined as the set of points P that have equal powers relative to both circles.

5.7.1 Proof that the Radical Axis is a Line

So far all we have is a lot of evidence obtained by playing with **Geometer** diagrams that the points satisfying our conditions lie on a line. It is not difficult, now that we know the answer, to construct a proof of that fact.
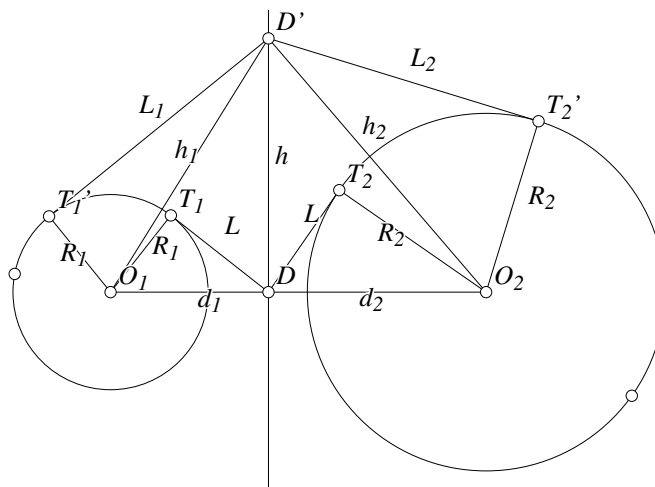

Figure 5.9: Radical Axis Proof
Bigtheorems/RadicalProof.T [M]

It seems that the easiest proof is algebraic. We will prove it only for the case that the circles are disjoint as in figure 5.9. A similar proof can be written when the circles

intersect. We will show that the radical axis is perpendicular to the line connecting the centers of the circles. Pick D on that line of centers such that $DT_1 = DT_2$. We need to show that for any point D' on the line perpendicular to O_1O_2 through D that $L_1 = L_2$, with the labeling as in the figure. To complete the proof, simply apply the pythagorean theorem a few times and combine the results:

$$
\begin{align}
d_1^2 &= R_1^2 + L^2 \tag{5.24}\\
d_2^2 &= R_2^2 + L^2 \tag{5.25}\\
h_1^2 &= d_1^2 + h^2 \tag{5.26}\\
h_2^2 &= d_2^2 + h^2 \tag{5.27}\\
h_1^2 &= R_1^2 + L_1^2 \tag{5.28}\\
h_2^2 &= R_2^2 + L_2^2 \tag{5.29}
\end{align}
$$

Combining equations 5.24, 5.26 and 5.28 and similarly equations 5.25, 5.27 and 5.29we obtain:

$$
\begin{align}
L_1^2 &= h^2 + L^2\\
L_2^2 &= h^2 + L^2
\end{align}
$$

which shows us that $L_1 = L_2$, so the proof is complete.

5.7.2 The Radical Center

Notice that if we have a set of three circles whose centers are not collinear, then each pair will have a radical axis. The point P of intersection of any two of the radical axes will have the same power relative to all three circles, so it must lie on the third radical axis. In other words, the three radical axes meet at a point called the **radical center** of the three circles.
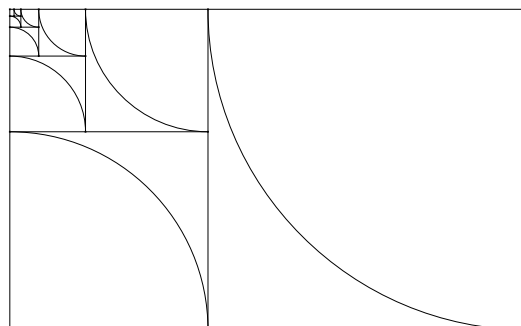
Chapter 6

Locus of Points

Computer geometry programs are particularly useful when studying locus problems whose goal is to find all the points that satisfy certain conditions. The term "satisfy certain conditions" can be very general, and with **Geometer** or a similar program, it is often easy to generate hundreds of such points automatically. This visual display of a large set of solution points often makes it quite easy to guess the answer, even if you have no idea how to prove it.

Here we will examine a number of applications of computer geometry programs to locus problems.

6.1 Unknown Geometric Locus

Many geometric problems are stated directly as locus problems: "Given a circle with points B and C fixed upon it, if the point A is allowed to move on the arc $\overarc{BC}$, what is the locus of points of the incenter of $\triangle ABC$?
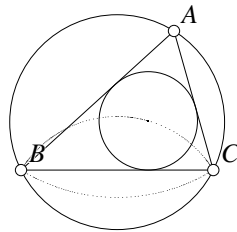
Figure 6.1: Locus of Incenters
`Locus/InLocus.T` [M]

With **Geometer**, it is easy to find out (or at least to make a good guess). Draw the figure described above and change the color of the center of the incircle to the smearing color. Then move point A and see what happens.

The result of one such experiment is shown in figure 6.1. The incenter sweeps out two curves depending upon which side of the chord BC it lies. The two curves look like circles, but if you would like more evidence, try creating three points on the smeared curves and see if the circle through them appears to match the rest of the points. (See Section 6.1.1 for the details on how to do this using **Geometer**.) You can even find the center and radius of that circle, and perhaps that extra evidence will help you to prove your result.

It is also a good idea in situations like this to repeat the experiment with different configurations of the points B and C to see how that affects the locus curve.

6.1.1   Technical Details

If you would like to place a circle on the smeared images of the point that traces the incenter of $\triangle ABC$ in figure 6.1 using **Geometer**, you will find that the instant you create a new geometric object, all the smear points disappear (because **Geometer** thinks you are smearing out a new set). Here is a reasonably easy approach to the problem.

Move A to smear out the curve. Then use the *Free P* command to create a free point. Place the cursor over the smear curve and click down. This creates a point, but the instant the point is created, the smear data disappears. Just smear it again (moving point A only—leave B and C where they are), and then create another new point in the same way. Repeat a third time and you will have three points roughly along the apparent curve. Draw a circle through those three points and smear a final time by moving point A. If the smear seems to stick to your circle, it is likely that the desired locus is, indeed, a circle.

It works better if the three points are as far apart as possible on the smear curve, and if you have clearly missed the curve with one of the points, either drag it closer to the smear curve or erase and re-draw it before you create the circle. In fact, it is obvious that as the point A approaches B or C, so does the center of the inscribed circle, so you can use points B, C, and a single point on the curve to determine the test circle.

In this example, once you have convinced yourself that the smear curve does seem to be a circle you can perform other tests on your **Geometer** circle such as locating its center, looking for tangent lines, et cetera.
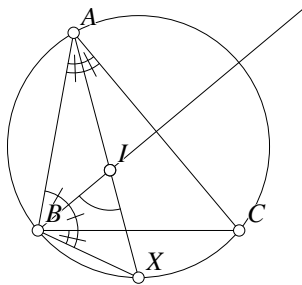
6.1.2   Conjecture and Proof

When you do draw a circle and see that it seems to fit the locus of points very closely, you will find that the center of that circle seems to lie on the arc $\overset{\frown}{BC}$ and in fact seems to bisect the arc. It is also obvious that as A gets very close to B (or C), the incenter will also approach B (or C), so it seems that a good conjecture might be that the incenter I of $\triangle ABC$ satisfies $IX = BX = CX$. See figure 6.2.

Figure 6.2: Locus of Incenters
`Locus/InLocus1.T [P]`

If we test the diagram with **Geometer**, our suspicions are confirmed: We are told that $IX = BX$ and in addition that a bunch of sets of angles are equal. Most of the angle equalities are obvious from the figure, but **Geometer** also reports that $\angle XIB = \angle XBI$.

Using these ideas, it is simple to prove this lemma. AX is an angle bisector so $\angle BAX = \angle CAX$. In addition, AX will bisect the arc $\overset{\frown}{BC}$, so X, the proposed center of our circle of locus points, is in the center of that arc. Of course $\angle CBX$ is also inscribed in the arc CX so it is also equal to $\angle BAX$ and $\angle CAX$.

The line BI is also an angle bisector, so $\angle CBI = \angle IBA$. We also know that $\angle XIB$ is equal to $\angle IAB + \angle IBA$ since it is an exterior angle of $\triangle AIB$. But clearly $\angle XBI = \angle XBC + \angle CBI$, and since both are equal to the sum of congruent angles, $\angle XIB = \angle XBI$. Thus $\triangle XIB$ is isosceles, and hence $IX = BX$.
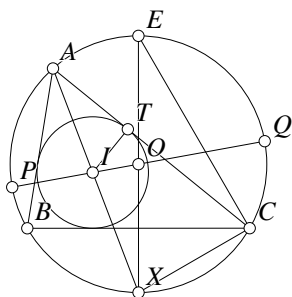
6.2 Euler's Formula

The locus sought in Section 6.1 perhaps seems a bit artificial, but in fact it might be one of the first things you would try if you were working on the following problem:

If R and r are the radii of the inscribed and circumscribed circles of a triangle, show that the distance d between the centers of those circles satisfies the equation:

$$R^2 - d^2 = 2Rr. \tag{6.1}$$

Figure 6.3: Locus of Incenters
`Locus/InLocus2.T [P]`

This is known as Euler's formula[1], and is closely related to the triangular version of Poncelet's theorem. See Section 3.8.

A reasonable approach is to leave R constant and check a variety of inscribed triangles to see how r and d might related. Our experiments with figure 6.1 in the previous section yield some clues.

---

[1] There are probably a dozen formulas called "Euler's formula" so which one is meant generally depends upon the context.

We also notice that $R^2 - d^2 = (R+d)(R-d)$ on the left side of equation 6.1. In figure 6.3 $d = IO$, so it is clear that $IQ = R + d$ and $PI = R - d$. The point I divides a diameter into pieces whose product we are seeking. Any other chord through I will do the same thing, so we can conclude that $R^2 - d^2 = AI \cdot IX$, for example.

On the right side of equation 6.1 we have the product $2Rr$. The diameter of the circumcircle is $2R$, and r is the radius of the incircle.

Figure 6.4: Rectangle of Incenters
`Locus/Circlecenters.T` [P]

If these numbers are going to appear in our calculations, we will need to have segments of those lengths in the diagram. We draw diameter EX and radius IT as in figure 6.3. Other radii and diameters would also work, but EX involves the points O and X that are already in the diagram, and IT connects to the segments AI and IX which appear in formulas we already have. If in addition we draw EC and XC it appears that $\triangle ECX$ is similar to $\triangle ATI$. Of course we can use **Geometer** along the way to test for the various similarities.

With these constructions, the proof of Euler's formula is easy.

Clearly $\angle ATI = 90°$ since IT is a radius of the circle to which AT is tangent, and $\angle ECX = 90°$ since it is inscribed on the diameter EX. By AA, $\triangle ECX \sim \triangle ATI$ so $AI/IT = EX/CX$. But the lemma proved in Section 6.1.2 tells us that $CX = IX$, $IT = r$, and $EX = 2R$, so if we simply substitute and cross-multiply, we find that $2Rr = AI \cdot IX = (R + d)(R - d) = R^2 - d^2$.

Figure 6.5: Three Medians
`Locus/Medians3.T` [P]

If you are interested in using the lemma proved in Section 6.1.2 again, look at the diagram in figure 6.4. Show that the four incenters of the triangles obtained by adding a diagonal to a cyclic quadrilateral lie at the corners of a rectangle. We will not prove it here in the text, but the **Geometer** diagram contains the complete proof.

6.3 Triangle from Three Medians

Given the lengths of the three medians of a triangle, construct the triangle. This seems like a difficult problem at first—none of the sides or angles of the original triangle are given, so it is hard even to know how to begin.

We do know a few facts about medians. All three medians of a triangle meet at the centroid, G, and G is $2/3$ of the distance from each vertex to the opposite side.

In addition, since they are medians, the end of each median meets the midpoint of the opposite side.

The next thing to do is to draw a **Geometer** program to get a feel for the problem. Draw your own, or load the file `Locus/Medians3.T`. The file used here has three lengths, AA', BB' and CC' that can be adjusted on the left of the main diagram. The centroid, G, is fixed, and all three copies of the lengths on the left are copied and constrained to pass through G at $2/3$ of the distance from the vertices-to-be (A, B and C). Since any one of them can be fixed, in the diagram CC' is not movable. The points A and B can be moved, each on a circle centered at G with radius equal to $2/3$ of the lengths of AA' and BB', respectively. Figure 6.5 illustrates a situation where A and B have been manually adjusted to be almost correct—the points A', B' and C' lie close to, but not exactly on, the edges BC, CA and AB.

You can get a feeling for the problem and even obtain approximate solutions by experimenting with a **Geometer** diagram, but this will not provide a construction. So far the diagram has not made use of the fact that the endpoint of each median has to lie on the midpoint of the opposite side. We are leaving C fixed, but what if we allow B to move completely around its circle centered at G and having radius BG? Every location of B on that circle of allowable positions determines where B' lies. The locations of B' and C determine where A would have to be in order that the location of B be a solution. Since B' is the midpoint between A and C in the solution, we can reflect C across B' to find where A would have to lie if B (and hence B') were in the correct position. That theoretical position of A is shown as A^* in figure 6.6 for one particular configuration of B, G and C.
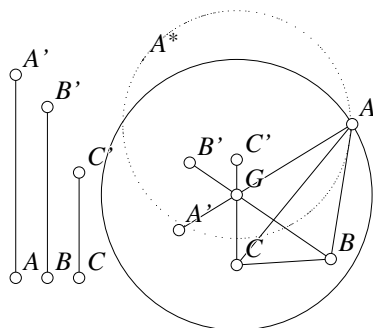

Figure 6.6: Three Medians
`Locus/Medians3.T` [P]

In addition, A must lie on the circle centered at G and having radius $(2/3)AA'$. If A^* is drawn in a smearing color and B is moved around its circle we see the locus of possible locations of A, which appears to be a circle. Since the actual position of A has to be on that circle *and* on the circle of radius $(2/3)AA$; about G, the solution puts A at the intersection of the two circles.

A quick analysis shows us that the point A^* is the reflection of B through the point C'. Thus if B sweeps out a circle of radius r, then A^* will sweep out a circle of the same radius. A brute-force method to construct this circle would be to select three different locations for B on its circle around G and reflect each of them through C'. Then construct the circle passing through those three points. The intersection of that circle and the circle of radius $(2/3)AA'$ about G is the location of vertex A. From there it is easy to find B and the construction is complete.

To review the construction: Draw the segment CC' at any location. Find the centroid G that is $2/3$ of the distance from C to C'. Draw a circle about G of radius $(2/3)BB'$ and select three points on it. Reflect these three points through C' and construct a circle passing through the three reflected points. Draw a circle of radius $(2/3)AA'$ about G and find an intersection of it with the circle through the three reflected points. This is the location of A. Reflect A through C' to obtain the location of B. This completes the construction.

6.4 The Ellipse

An **ellipse** is commonly defined to be the set of points the sum of whose distances to two fixed points (the foci) is constant. Although **Geometer** draws ellipses (and in fact, all conic sections) as primitive objects, it is easy to use the definition above to construct as many points as desired on an ellipse using only line and circle drawing commands.
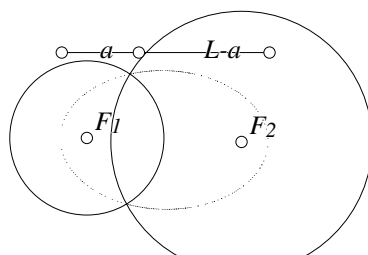

Figure 6.7: Ellipse from Definition
Locus/Ellipse.T [P]

A diagram to draw an ellipse as in figure 6.7 is quite simple. The two foci, F_1 and F_2 are completely free points. Next, a line segment is drawn that has total length L, where L represents the sum of the distances to the foci. A point constrained to be on the line segment divides it into two segments of length a and $L - a$. A circle of radius a is drawn about F_1 and a circle of radius $L - a$ is drawn about F_2. The points at the intersections of the two circles are drawn in a smearing color. As the point separating the two segments is moved, the relative sizes of the circles changes and so do their intersections which smear out the ellipse. Such a diagram makes it easy to adjust the distances between the foci and total length L to see how those changes affect the resulting ellipse.

The ellipse in figure 6.7 is not very pretty since at the left and right ends when the circles approach tangency there are not nearly as many intersections. Remember that the point on the line can only move in units of one pixel on the computer screen so the radii of the circles also change in one-pixel increments. When the circles are close to tangency, making one circle larger and the other smaller by an entire pixel will cause the points of intersection to jump significantly.

Perhaps the easiest way to improve the drawing is to recall that an ellipse is simply a scaled circle. In other words, if every point on a circle has its x-coordinates multiplied by one value and its y-coordinates by another, the resulting points will lie on an ellipse. In figure 6.8 we have done just that. The point A is constrained to be on a circle, and with the editor we added a single line to the end of the file. Here is the **Geometer** code in its entirety, and only the final line was created with the text editor.

```
.geometry "version 0.60";
v1 = .pinned(0, 0);
v2 = .free(0.95, 0, .in);
c1 = .c.vv(v1, v2);
v3 = .vonc(c1, -0.850982, 0.422292, "A");
ve = .v.vscale(v3, 1.700000, 0.400000, .smear, .dot);
```
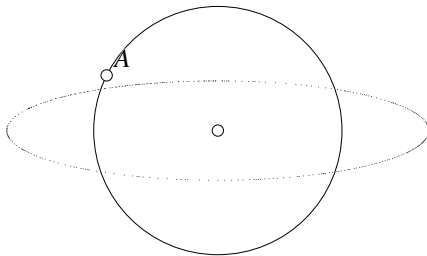
The final line tells **Geometer** to determine the coordinates of `ve` by multiplying the x-coordinate by 1.7 and the y-coordinate by 0.4.

The easiest way to get even spacing is simply to use a parametric definition of the ellipse. If you want an ellipse centered at the origin, aligned with the coordinate axes, and passing through $(\alpha, 0)$, $(-\alpha, 0)$, $(0, \beta)$ and $(0, -\beta)$, then as the parameter t varies from $0°$ to $360°$, the points $(\alpha \cos t, \beta \sin t)$ will sweep out the el-

Figure 6.8: Ellipse with Scaling
`Locus/Ellipse1.T [P]`

lipse. If t takes uniform steps, the output points will not be spaced uniformly, but they will be spaced in a pleasing, symmetric manner.

Using the **Geometer** `.script` command below causes the value of `t` to go from 0 to 360 in steps of 2. When this script is run by pressing the *Run Script* button, the drawing in figure 6.9 is generated.

```
.geometry "version 0.60";
t = .script(0.000000, 360.000000, 2.000000);
x = .f.rpn(t, .cos, 1.700000, .mul);
y = .f.rpn(t, .sin, 0.400000, .mul);
v = .v.ff(x, y, .smear, .dot);
```

6.5 Locus Construction Exercises

Without any explanation of the results, here are a few problems where the goal is to find a locus of points. Read the problems, try to figure out the locus (with or without proof), and then check the associated **Geometer** files for the author's solution. Remember that using the *Edit Geometry* command you can look at the exact order that the commands were issued in **Geometer**.

Figure 6.9: Ellipse with Script
`Locus/Ellipse2.T [P]`

All the examples below were constructed using only **Geometer**'s graphical user interface; it was never necessary to edit the textual version of the diagram.

1. Let A, B, C and D be any four points on a circle. Let P be any point (inside, outside, or on the circle), and draw the lines PC and PD. Let Q and R be the intersections of PC and PD with the circle other than C and D. Find the locus of the intersections of the circles PRA and PQB. (Solution: `Locus/Locus1.T`)

2. Let B and C be two fixed points on a circle, and let A move freely on the circle. Find the locus of the projections of the orthocenter H onto the bisector of $\angle CAB$. Depending on what you mean by the angle bisector (there is an interior and exterior bisector) this could mean different things. The included **Geometer** diagram changes the interpretation from interior to exterior when the point A changes sides of the segment. You may wish to draw both the interior and exterior bisectors to see what happens in both cases. (Solution: `Locus/Locus2.T`)

3. Given a fixed circle K, find the locus of the endpoints of the diameters parallel to a given line of a fixed-radius circle orthogonal to K. (Solution: `Locus/Locus3.T`)

4. Given a fixed circle K and a fixed point P, allow Q to move around K, and construct the circle through P that is tangent to K at Q. What is the locus of the intersection of the lines tangent to that new circle that pass through P and Q? (Solution: `Locus/Locus4.T`)

5. Given two points A and B, find the locus of points P such that $PB : PA = r : 1$, where r is a fixed ratio. (Solution: `Locus/Locus5.T`)

6.6 Constant Ratio

The final exercise in the previous section is important enough to deserve a proof. After constructing the locus using **Geometer**, it is clear that the answer is a circle, but the proof is a little tricky.

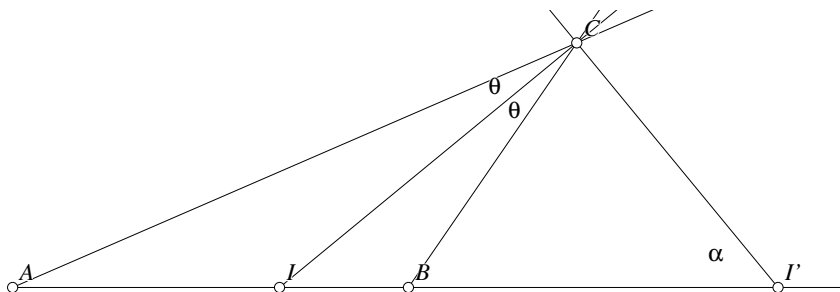Figure 6.10: Interior and Exterior Subdivision
`Locus/InEx.T [P]`

Before stating the theorem we will first prove that the internal and external bisectors of an angle of a triangle divide the side opposite internally and externally in the same ratio. In figure 6.10, if CI amd CI' are the internal and external bisectors of $\angle C$, then $AI/BI = AI'/BI'$. This is just a taste of what is to come in Chapter 10.

The proof is not difficult. Let $\angle ACI = \angle BCI = \theta$. Since CI is the angle bisector of $\angle ACB$, we have $AI/AC = BI/BC$. Notice also that since CI and CI' are both bisectors, $\angle ICI' = 90°$. Next, we apply the law of sines to triangles $\triangle ACI'$ and $\triangle BCI'$:

$$\frac{AI'}{\sin(90° + \theta)} = \frac{AC}{\sin \angle CI'B} \quad \text{and} \quad \frac{BI'}{\sin(90° - \theta)} = \frac{BC}{\sin \angle CI'B}.$$

Since $\sin(90 \circ +\theta) = \sin(90° - \theta) = \cos(\theta)$, the two equations above can be combined to yield the desired result: $AI/BI = AI'/BI'$.

Now we are ready to prove the main theorem.

Theorem 6.1 (Constant Ratio Theorem) *Let A and B be two points on a line and let I lie on the segment between them such that I is not the midpoint of AB. Then the locus of points P such that $PA/PB = IA/IB$ is a circle passing through I whose center is on the line AB. If O is the center of that circle, then[2] $OA : OI = OI : OB$.*

Based on the previous result, the proof is not difficult: see figure 6.11.

Once the locations of the points A, B and I are chosen, find an arbitrary point P such that $PA/AI = PB/BI$. Since this ratio holds, that means that PI is the angle bisector of $\angle APB$. Construct a line perpendicular to PI through P passing through the line AB at I'. PI and PI' are the internal and external angle bisec-

Figure 6.11: Points of Constant Ratio
`Locus/ConstRatio.T` [M]

tors of $\angle APB$ in $\triangle APB$, so the points I and I' divide the segment AB internally and externally in the same ratio. Notice that this makes the position of I' independent of P, so for every such P, the angle $\angle IPI' = 90°$. Thus P lies on a circle whose diameter is II', and the proof is complete.
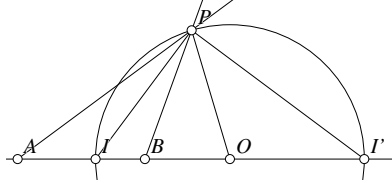
6.7 Higher-Order Curves

In Euclidean geometry, we normally deal with straight lines and circles, and sometimes conic sections: ellipses, hyperbolas and parabolas. All of these are first degree (lines) or second degree (all the rest). By "first degree" we mean that they can be described by equations whose highest power is 1. Second degree curves have equations whose highest power is 2, et cetera. The general equation for any straight line is given by:

$$ax + by + c = 0,$$

---

[2]We will see in Chapter 8 that A and B are inverses of each other relative to the circle. Also note that if I is at the midpoint of the segment AB, then the locus is the perpendicular bisector of AB which can be thought of as a circle of infinite radius.

for suitably-selected constants a, b and c. The general second degree equation is:

$$ax^2 + bxy + cy^2 + dx + ey + f = 0,$$

where a, b, ..., f are constants. Second degree equations describe all the conic sections, where we consider a circle to be a special form of an ellipse.

We have already encountered two examples of curves of higher degree than 2 (the cardioid and the limaçon) in Section 2.10.2.

6.7.1   Example: Strophoids and Freeth's Nephroid

Given a curve C and two points, A and O, let X be an arbitrary point on C, and let P_1 and P_2 be the two points on the line XO such that $AX = AP_1 = AP_2$. The locus of all such P_i is called the **strophoid** of the curve relative to O and A.

In figure 6.12, the curve C is a circle, O is the center of the circle, and A lies on the circle. This particular strophoid is called the nephroid of Freeth.
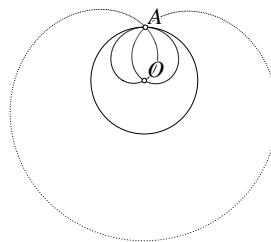
Try drawing strophoids of other curves (do not forget the straight line), and modify Freeth's nephroid to a more general strophoid by allowing the point A to be off the circle, and/or make O different from the center of the circle.

Figure 6.12: Freeth's Nephroid
`Locus/Strophoid.T` [S]

6.7.2   Additional Higher-Order Examples

Following is a collection of interesting looking curves. To save space, the drawings are not included here, but every one is available as a **Geometer** diagram on the CD where each curve is drawn by pressing the *Run Script* button. In every case it is probably worthwhile to try to draw the curve yourself using **Geometer** before loading the file to see the solution.

♦ In many cases in the examples that follow, multiple descriptions may be given for the curve, including some combination of geometric, cartesian, polar, and parametric. If you do not know what these terms mean, do not worry—you can still draw the pictures by pressing the *Run Script* button.

Many additional examples of interesting curves can be found in [Lawrence, 1972].

6.7.2.1   The Cissoid of Diocles

Given a line L and a circle C tangent to it at Q. Let S be the point at the other end of the diameter of C through Q. If X is a point on L and R is the intersecion of the line XS with C, then the **cissoid of Diocles** is the locus of points P on XS such that $PX = RS$.

Cartesian equation: $y^2(a - x) = x^3$.
Polar equation: $r = a \sin \theta \tan \theta$.
Parametric equation: $(a \sin^2 t, a \tan t \sin^2 t)$.
Geometer file: `Locus/Cissoid1.T`

6.7.2.2 The Parabola

Given a line L and a point F (the focus) not on L, a **parabola** is the set of all points whose distance from the focus F and the line L are equal.

Cartesian equation: $y^2 = 4ax$.
Polar equation: $r \sin^2 \theta = 4a \cos \theta$.
Parametric equation: $(at^2, 2at)$.
Geometer file: `Locus/Parabola.T`

6.7.2.3 The Conchoid of Nicomedes

Given a line L and a point O, let the point Q move along L. The **conchoid of Nicomedes** is the locus of points P_1 and P_2 on the line OQ such that $P_1Q = P_2Q$.

Cartesian equation: $(x^2 + y^2)(x - b)^2 = a^2 x^2$.
Polar equation: $r = a + b \sec t$.
Parametric equation: $(b + a \cos t, \tan t(b + a \cos t))$.
Geometer file: `Locus/Conchoid.T`

6.7.2.4 The Hypotrochoid and Rhodonea

If a small circle rolls around the inside of a larger circle and the ratio of their radii is an integer (so that the small circle completes an integral number of turns when it returns to its starting point), then if P is a point connected to the smaller circle, a **hypotrochoid** is the path traversed by P. Note that P may be outside the smaller circle, but if that is the case, it rotates as if it were connected to the smaller circle by a rigid rod.

A **rhodonea** is a special case of the hypotrochoid where if h is the distance of P from the center of the smaller circle, the radii of the smaller and larger circles are $h(n - 1)/(n + 1)$ and $2nh/(n + 1)$, respectively.

Polar equation: $r = a \cos(n\theta)$.
Geometer file: `Locus/Rhodonea.T` (for $n = 7$)

6.7.2.5 Lissijous Figures

If two wheels of the same size turn at different rates such that they eventually return to the starting position after a certain number of turns, and if a point is selected on each, then a **lissijous figure** is the locus of points that are the intersection of a vertical line drawn from one of the points and a horizontal line from the other.

Parametric equation: $(\cos nt, \sin mt)$, $(\cos nt, \cos mt)$ or $(\sin nt, \sin mt)$.
Geometer diagram: `Locus/Lissijous.T`

$\boxed{\text{6.7.2.6}}$ **Dumbbell Curve**

This is simply a curve that looks like a dumbbell.

Parametric equation: $y^2 = x^4 - x^6$.

Geometer file: `Locus/Dumbbell.T`

$\boxed{\text{6.7.2.7}}$ **Bézier Splines**

Choose a set of $n \geq 3$ control points. Given t, where $0 \leq t \leq 1$, between each pair, choose a point that divides the segment connecting them into a ratio of t and $1 - t$, yielding $n - 1$ points. Repeat the process until there is only a single point. As t varies from 0 to 1, this final point sweeps out a curve called a Bézier spline. **Geometer** provides cubic Bézier splines as a drawing primitive.

Parametric equation: $(\sum_{i=0}^{n} x_i \binom{n}{i} t^i (1-t)^{n-i}, \sum_{0}^{n} y_i \binom{n}{i} t^i (1-t)^{n-i})$, where the coordinates of the control points are (x_i, y_i).

Geometer file: `Locus/Bezier.T` (with 5 control points)

$\boxed{\textbf{6.8}}$ ♦ **Plotting Curves of the Form:** $y = f(x)$

By default, **Geometer**'s coordinate system runs from -1 to 1 in the window's smaller dimension, so $(0, 0)$ is always in the center of the window. Suppose we would like to prepare a nice drawing of the function:

$$y = \frac{cos(x)sin(10x)}{2} + sin(3x)$$

from $x = -\pi$ to $x = \pi$, where the angles are measured in radians.
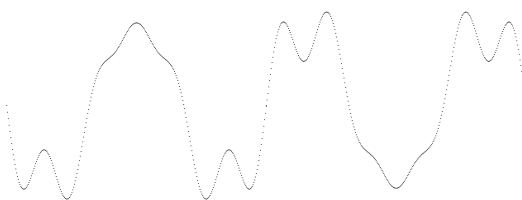

Figure 6.13: $y = cos(x)sin(10x)/2 + sin(3x)$
`Locus/Plot.T` [S]

The final graph will occupy a space $2\pi \approx 6.28$ units long and as much as 3 units high. The smaller dimension will be the 3 units, and it has to fit in a coordinate system running from -1 at the bottom to 1 at the top. Thus we will scale everything by a factor of $2/3$.

Here is the **Geometer** code:

```
.geometry "version 0.60";
.radianmode;
t = .script(-3.141590, 3.141590, 0.010000);
x = .f.rpn(t, 0.666667, .mul);
y = .f.rpn(t, .cos, t, 10.000000, .mul,
        .sin, .mul, 2.000000, .div, t,
        3.000000, .mul, .sin, .add, 0.666666,
        .mul);
v = .v.ff(x, y, .smear, .dot);
```

The `.radianmode` makes **Geometer** calculate in radians instead of degrees. The `t` varies from $-\pi$ to π. The true x value is t, but to fit the curve to the **Geometer** window we need to multiply it and the y-coordinate by $2/3$. The y-coordinate is calculated in a straightforward way, and the result is plotted in a smearing color. The window size is adjusted to have a ratio of about 2π long to 3 high, and the *Run Script* button yields the curve shown in figure 6.13.

6.9 ♦ Plotting Curves in Polar Coordinates

If the points on a curve are defined as a relationship between the distance r from the origin and the angle θ the ray from the origin to the point makes with the positive x-axis, the curve is said to be described in **polar coordinates**. For example, a circle of radius 1 centered at the origin has a radius independent of the angle, so the polar form of that circle is $r = 1$. Usually things are not that simple, however.

There are two ways to plot such curves using **Geometer**. One is for each (r, θ) pair to draw a circle of radius r and a ray at angle θ and plot (with a point having the smear color) the intersection of the ray and circle. If you do not mind editing the textual version of a **Geometer** file, another method is to calculate the cartesian coordinates $(x, y) = (r \cos \theta, r \sin \theta)$ and to plot those points as r and θ vary appropriately.
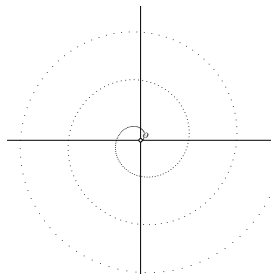

Figure 6.14: Archimedes' Spiral
Locus/Polar.T [S]

Suppose we wish to plot **Archimedes' spiral** given by the polar equation $r = (0.001)\theta$, where the angle θ is measured in degrees. Two solutions to the problem are presented below in **Geometer** code format. The two files are `Locus/Polar.T` and `Locus/Polar1.T` and are listed below. The listing of `Locus/Polar.T` is on the left. Both are scripts, so run them by pressing the *Run Script* button. See figure 6.14.

```
.geometry "version 0.60";              .geometry "version 0.60";
o = .pinned(0, 0, "O");                o = .pinned(0, 0, "O");
one = .pinned(1, 0, .in);              one = .pinned(1, 0, .in);
l1 = .l.vv(o, one, .longline);         l1 = .l.vv(o, one, .longline);
l2 = .l.vlperp(o, l1, .longline);      l2 = .l.vlperp(o, l1, .longline);
theta = .script(0.0, 1080.0, 4.0);     theta = .script(0.0, 1080.0, 4.0);
r = .f.rpn(theta, 0.001, .mul);        r = .f.rpn(theta, 0.001, .mul);
x = .f.rpn(r, theta, .cos, .mul);      c = .c.vf(o, r);
y = .f.rpn(r, theta, .sin, .mul);      ang = .a.f(theta);
p = .v.ff(x, y, .smear, .dot);         v1 = .v.avv(ang, one, o);
                                       l3 = .l.vv(o, v1);
                                       p = .v.lc(l3, c, 2, .smear, .dot);
```

6.10 ♦ Envelopes of Curves: the Epicycloid

Sometimes the shape of a curve is made obvious if you draw a large set of lines that are tangent to the curve. The various types of epicycloids provide some good examples. See figure 6.15.

An epicycloid is path traced by a point on the edge of a circle as it rolls around another circle. If the ratio of the diameters of the circles is rational, the path closes.

To draw the envelope of an N-epicycloid begin by selecting a reference point on a circle. Next, for an evenly-spaced set of points around the circle, connect each of those points by a straight line to the point that is N times as far around the circle as the original point, wrapping around the circle if necessary. Figure 6.15 illustrates the results for $N = 2$, $N = 3$, $N = 5$ and $N = 8$.

The three examples were generated by the following **Geometer** code:

```
.geometry "version 0.61";
m = .script(0.000000, 360.000000, 2.000000);
rad = .f.rpn(0.800000);
vn2 = .pinned(-0.9, 0.92, .dot, "2");
vn3 = .pinned(-0.72, 0.92, .dot, "3");
vn4 = .pinned(-0.54, 0.92, .dot, "4");
vn5 = .pinned(-0.36, 0.92, .dot, "5");
vn6 = .pinned(-0.18, 0.92, .dot, "6");
vn7 = .pinned(0, 0.92, .dot, "7");
vn8 = .pinned(0.18, 0.92, .dot, "8");
vn9 = .pinned(0.36, 0.92, .dot, "9");
vn10 = .pinned(0.54, 0.92, .dot, "10");
vn11 = .pinned(0.72, 0.92, .dot, "11");
vn12 = .pinned(0.9, 0.92, .dot, "12");
l1 = .l.vv(vn2, vn12);
vmid = .vonl(l1, -0.365269, 0.92, .red, "N");
v2x = .f.vxcoord(vn2);
v12x = .f.vxcoord(vn12);
vmidx = .f.vxcoord(vmid);
n = .f.rpn(vmidx, v2x, .sub, v12x, v2x,
        .sub, .div, 10.000000, .mul, 0.500000,
        .add, .truncate, 2.000000, .add);
p0 = .pinned(0, 0, .in);
p1 = .v.ff(rad, 0.000000, .in);
c = .c.vv(p0, p1);
```
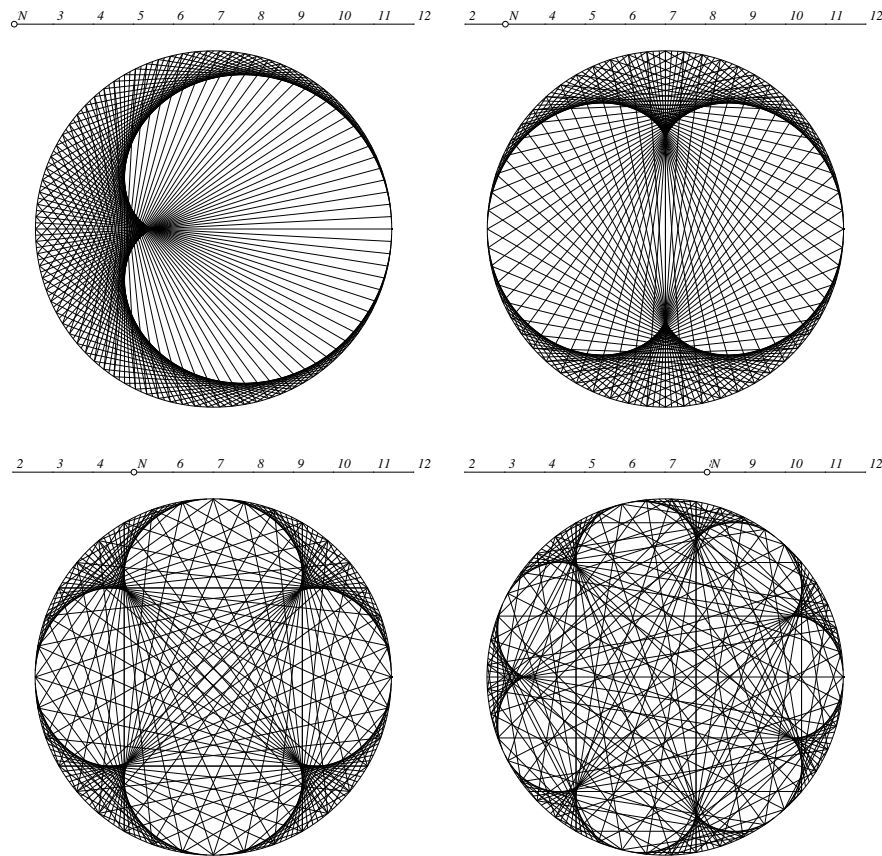
Figure 6.15: Various Epicycloid Envelopes
`Locus/Epicycloid.T [S]`

```
vx0 = .f.rpn(m, .cos, rad, .mul);
vy0 = .f.rpn(m, .sin, rad, .mul);
v0 = .v.ff(vx0, vy0, .in);
m1 = .f.rpn(m, n, .mul);
vx1 = .f.rpn(m1, .cos, rad, .mul);
vy1 = .f.rpn(m1, .sin, rad, .mul);
v1 = .v.ff(vx1, vy1, .in);
l = .l.vv(v0, v1, .smear);
.text("Epicycloid:  Set N above and press 'Run Script'.");
```

Here is how the code works. The variable `m` is the angle around the circle of the initial point, measured in degrees. It varies from $0°$ to $360°$ in steps of $2°$. This can be changed to a smaller number for more tangent lines. It can be any value, but it makes a more uniform diagram if it divides evenly into 360. The value of `rad` is the radius of the circle which is centered at the origin (center = `p0`; point on boundary = `p1`).

The next few lines of code allow the user to choose the value of N for which the N-epicycloid will be generated. The point `vmid` (labeled N in the **Geometer** diagram) can be moved between `vn2` and `vn12` and its position is used to calculate `n`.

To calculate `n` we find the x-coordinates of `vn2`, `vmid` and `vn12`. The expression (`vmidx - v2x`)/(`v12x - v2x`) is the proportion of the distance along the line from `vn2` to `vn12` at which `vmid` lies. It will be between 0.0 and 10.0. When we multiply by ten, add 0.5 (for rounding) and truncate, we obtain an integer between 0 and 10. Adding 2 to that gives us an integer `n` between 2 and 12, inclusive.

The points labeled `vn3`, ..., `vn11` are simply to label the line so the user can tell what value of `n` will be calculated.

Next the circle is drawn and the coordinates of the starting point (determined by angle `m`) and the endpoint (determined by angle `mn`) of the line segment are calculated. That line is then drawn in a smearing color.
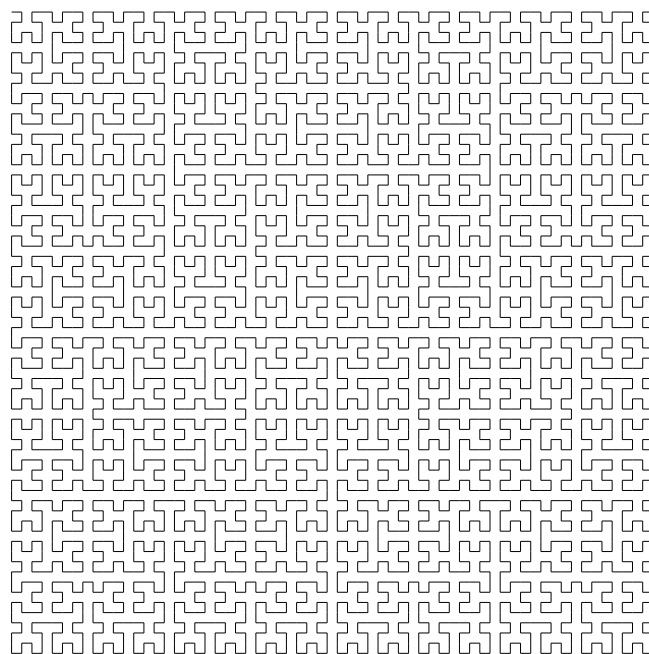
Chapter 7

Triangle Centers

Perhaps the most amazing theorems from high school geometry concern the concurrence of the medians, altitudes, angle bisectors and even the perpendicular bisectors of the sides of a triangle. This occurs not just for particular triangles, but for *every* triangle. These high school examples all have special names that may at first sight seem unrelated to how they are defined. The intersection of the three medians is called the **centroid**, the intersection of the three angle bisectors is called the **incenter**, the intersection of the three altitudes is the **orthocenter**, and the intersection of the three perpendicular bisectors of the sides is the **circumcenter**.

As we shall see, the points of concurrence above just scratch the surface. These, and all the similar points that will be defined later in this chapter will go by the general name **triangle center**. We will learn about more of these centers, about other properties that the centers satisfy, interrelationships among the centers, and how the various centers have applications not only to other areas of geometry, but very practical applications as well.

137

A computer geometry program provides an excellent tool to study these triangle centers because after the geometric definition is made, you can use the computer to investigate how the centers are related to the size and shape of a huge variety of triangles. Without a computer it would be important to draw many careful examples by hand to learn about these relationships.

In fact, if you use **Geometer**, there is a program on the CD called `Finder.T` that provides a tool to help determine the identity of an unknown center. Just construct your center in the diagram and then page through a large collection of common centers (including all those covered in this chapter) to see which one seems to be the same as yours. See Section 7.12 for details.

To prove that various points are centers of one form or another we will often need to show that three lines are concurrent. One of the most useful and powerful ways to do this is with Ceva's theorem, so now might be a good time to review that. See Section 5.1

7.1 General Properties of Triangle Centers

We will be looking at a lot of specific triangle centers in what follows, but since we have already mentioned a few of them, it is a good idea to take a look at the sorts of properties that all triangle centers have. Here is a short list of those properties. You know about the four of them mentioned above (the incenter, circumcenter, orthocenter, and centroid)—check to see that these properties hold at least for those four centers.

- The center will exist for triangles of every shape. The term "center" may be a little misleading in that there is no guarantee that it even lies inside the triangle— the orthocenter and the circumcenter, for example, may lie outside.

- Similar triangles will have similar centers. In other words, if you rotate or translate a triangle, or uniformly scale it to a larger or smaller size, the center should rotate or translate with the triangle, or should scale up and down as the triangle does.

- As a sort of corollary to the statement above, a triangle center must depend "equally" on the three vertices. In other words, it would not be reasonable to define a new kind of center for $\triangle ABC$ to be the midpoint of the segment AB because a simple relabeling of the vertices of the triangle would move this supposed center to lie on a different side of the triangle.

7.2 The Gergonne Point
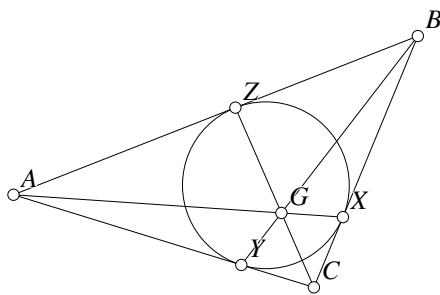
If the points of tangency of the incircle of triangle $\triangle ABC$ are connected to the opposite vertices as in figure 7.1, then they all meet at a point, called the Gergonne point.

Figure 7.1: The Gergonne Point
Tricenters/Gergonne.T [M]

7.2.1 Geometer Analysis of the Gergonne Point

This is not too difficult to prove, but with **Geometer**'s ability to search for relationships, it is extremely easy. When we test the diagram for the figure above, the following results are returned:

```
Points on a line:
    C  Z  G
Concurrent lines:
    (AX) (BY) (CZ)
Equal length segments:
    [C X] [C Y]
Equal length segments:
    [B Z] [B X]
Equal length segments:
    [A Z] [A Y]
Ratios:
    [B Z] x [C X] = [B X] x [C Y]
    [A Z] x [C X] = [A Y] x [C Y]
    [A Z] x [B Z] = [B X] x [A Y]
```

Geometer agrees that the lines AX, BY and CZ are concurrent, but it also reports some interesting information about ratios and equal lengths.

We recall that perhaps the most common method to show that three lines are concurrent is to apply Ceva's theorem. In the problem above, the three lines will be concurrent if

$$\frac{AZ}{BZ} \cdot \frac{BX}{CX} \cdot \frac{CY}{AY} = 1.$$

We notice that in each pair of equal segments reported by **Geometer**, one appears in the numerator above and one in the denominator, and that each segment length appears exactly once. Thus if we can show that the three segment length equalities hold, we will be done. But these are trivial to show, since we know that the lengths of the two external tangents to a circle from a point (like AY and AZ, for example) are of equal length.

7.3   The Nine Point Center

Let $\triangle ABC$ be an arbitrary triangle. If A',
B' and C' are the midpoints of the sides,
X, Y and Z are the feet of the altitudes,
and K, L and M are the midpoints of the
segments HA, HB and HC where H is
the orthocenter (the intersection of the three
altitudes), then the nine points A', B', C',
X, Y, Z, K, L and M all lie on a circle
called the **nine point circle**. By manipu-
lating a diagram, notice that this holds no
matter whether the orthocenter is inside or
outside the triangle. Also notice that the or-
der of the points on the circle may change,
and sometimes two of the points may coin-
cide, but no matter what happens, all nine of

Figure 7.2: The Nine Point Circle
`Tricenters/Ninepoint.T [M]`

them remain locked on a circle. The center of that circle is a triangle center called the
nine point center. See figure 7.2.

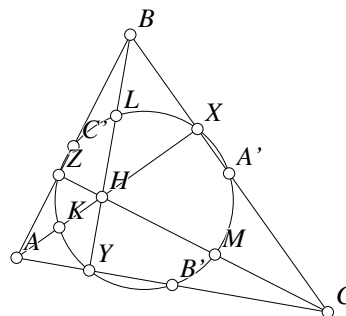### 7.3.1   Geometer Analysis of the Nine Point Circle

If we test the diagram as it is drawn
in figure 7.2, **Geometer** finds many re-
lationships, but nothing too surprising.
There are the obvious sets of points
that lie on circles like $CXYH$ and
$BXZH$, but this is due to the fact that
the altitudes make right angles with the
bases. Similarly, there are a lot of ra-
tios, but most are due to the fact that
we have sets of secants cutting through
the circle. We will probably have to
construct some lines to make progress.

Figure 7.3: The Nine Point Circle
`Tricenters/Ninepoint1.T [M]`

If you erase the nine point circle
from the figure to minimize the clut-
ter and just move around the vertices
A, B and C of the triangle, you may
notice that certain sets of four points
seem always to lie on a rectangle, such
as $A'C'KM$, $A'B'KL$ and $B'C'LM$. This is easy to check—just draw in one of them
in (say $C'A'MK$) and then test the diagram again.

When you do this, **Geometer** does report that the line segments appear to be per-
pendicular, so it probably is a rectangle, but it reports that $C'A'$, KM, and the base
AC are parallel and also that $A'M$, $C'K$ and the altitude BY are also parallel.
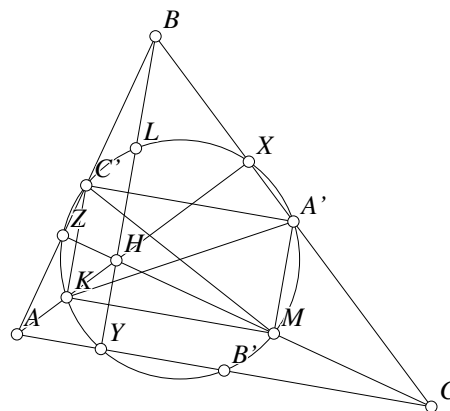
A look at the diagram shows us why: all those segments connect midpoints of triangle edges, so they must be parallel to the bases of those triangles and equal to half their length. In this case, the triangles in question are $\triangle ABC$, $\triangle AHC$, $\triangle BHC$ and $\triangle BHA$. What is more, the altitude BY is perpendicular to the base AC, so instead of merely being a parallelogram, $A'C'KM$ is a rectangle.

Obviously the same method can be used to show that other sets of four points also lie on rectangles.

We can find the center of the rectangle if we wish by finding the intersection of the diagonals $A'K$ and $C'M$, and in fact it should be clear that both of those segments are diameters of the circle we are looking for. But then we notice that if $A'K$ is the diameter of a circle and $\angle A'XK = 90°$ then X must lie on the circle with diameter $A'K$. From here, the proof is obvious.

7.3.2 Proof of the Nine Point Circle

See figure 7.3. Connect $A'C'$ and MK. We know that $A'C' \parallel AC$ since $A'C'$ connects the midpoints of AB and BC in $\triangle ABC$. By exactly the same reasoning, $KM \parallel AC$ using $\triangle AHC$, $C'K \parallel BY$ using $\triangle BAH$ and $A'M \parallel BY$ using $\triangle BCH$. Thus $A'M \parallel C'K$ and $A'C' \parallel KM$.

We know $BY \perp AC$ since BY is an altitude, and since $C'K$ and $A'M$ are parallel to BY and $A'C'$ and CM are parallel to AC, $C'K'$ and $A'M$ must be perpendicular to $A'C'$ and KM. Thus $A'C'KM$ is a rectangle.

Using the other bases and altitudes of $\triangle ABC$ we can conclude in exactly the same way that $A'B'KL$ and $B'C'LM$ are also rectangles.

Consider rectangle $A'C'KM$. Its center is at the intersection of $A'K$ and $C'M$ and the circle passing through its four vertices has $A'K$ and $C'M$ as diameters. But $\angle KXA' = 90°$ since AX is an altitude, so X must lie on the circle having the diameter $C'M$. Similarly Z must lie on the circle having diameter $A'K$. Thus the points A', M, K, C', Z and X all lie on a single circle.

By similar reasoning with the other two rectangles we obtain two other sets of six points lying on circles: $A'B'KLXY$ and $B'C'LMYZ$. Since each of these sets of points shares three points with each of the others, all the points must lie on a single circle—the nine point circle.

7.4 The Fermat Point

Erect equilateral triangles on the outside of each segment of $\triangle ABC$. The line segments connecting the vertices A, B and C to the tips of the opposite equilateral triangles are concurrent at a point O called the **Fermat point** of the triangle. See figure 7.4.

7.4.1 Geometer Analysis of the Fermat Point

If we test the diagram in figure 7.4 there are
a few interesting relationships (many of the
relationships found by **Geometer** have been
omitted from this list):

```
Concurrent lines:
    (AA') (BB') (CC')
Points on a circle:
    B  C  A'  O
Points on a circle:
    A  C  B'  O
Points on a circle:
    A  B  C'  O
Equal angles:
    (B C B') (A C A')
```
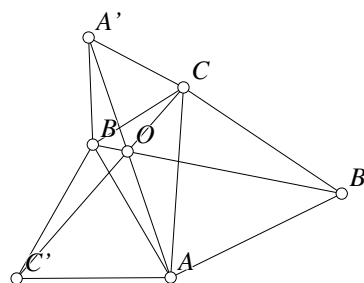

Figure 7.4: The Fermat Point
`Tricenters/Fermat.T [M]`

If we can show that the sets of points lie on a circle, we can conclude that the
angles AOB, BOC and COA are supplementary to $60°$ angles, so each of them must
be $120°$. This will show that the three lines meet at a point.

What we need to show, then, is that the sets of points lie on circles. If we could
show, for example, $\angle OA'C = \angle OBC$ then there would be a circle passing through
$OBCA'$.

The **Geometer** analysis also tells us that $\angle BCB' = \angle ACA'$, and those angles
might be used to show the similarity of $\triangle ACA'$ and $\triangle B'CB$. These facts together
provide plenty of clues to construct the proof.

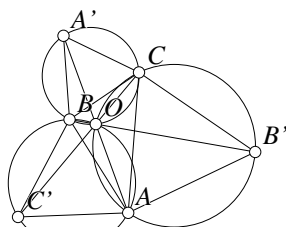## 7.4.2 Proof of the Existence of the Fermat Point

Figure 7.5: The Fermat Point
`Tricenters/Fermat1.T [M]`

Construct the circumcircles of each of the
three equilateral triangles and let O be the
intersection of AA' and BB'. From the fig-
ure it appears that O is also on CC' but this
is what we need to prove to show that the
three lines are concurrent at O and that O is
therefore the Fermat point.

In figure 7.5, $\angle BCA$ is equal to itself,
and since both angles $\angle A'CB$ and $\angle ACB'$
are equal to $60°$, we have:

$$\begin{aligned} \angle A'CA &= 60° + \angle BCA \\ &= \angle BCA + 60° = \angle BCB'. \end{aligned}$$

Since the outer triangles are equilateral, $CB' = CA$ and $A'C = BC$, so from SAS,
we know that $\triangle A'CA \cong \triangle BCB'$. Because the triangles are congruent, $\angle CB'O = \angle CAO$ and $\angle OA'C = \angle OBC$, so $OBA'C$ and $OACB'$ form two sets of four con-
cyclic points. Therefore O lies on the intersections of the circumcircles of the equilat-
eral triangles $\triangle A'BC$ and $\triangle AB'C$.

The same argument can be used to show the congruence of two other pairs of triangles: $\triangle ABA' \cong \triangle C'BC$ and $\triangle B'AB \cong CAC'$ and their common intersections also form sets of concyclic points. If we can show that the three circumcircles of the equilateral triangles pass through a single point we are done since that point must be O.

Since opposite angles in a cyclic quadrilateral are supplementary and the outer triangles are all equilateral, $\angle AOC = \angle COB = 120°$. Thus $\angle BOA = 120°$ since $\angle AOC + \angle COB + \angle BOA = 360°$. Therefore O lies on the circumcircle of $\triangle ABC'$ since it is opposite a $60°$ angle and we are done.

There is a second Fermat point that can be obtained if the three equilateral triangles point inward and not outward. The proof of this is left as an exercise.

7.5 The Nagel Point
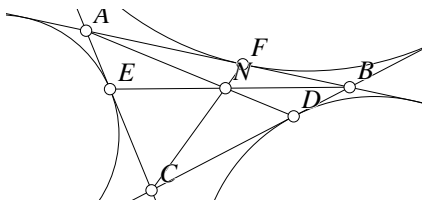
Given an arbitrary triangle $\triangle ABC$, extend the sides and construct the three circles that are exterior to the triangle and tangent to all three edges (the excircles). Connect each vertex with the point of tangency of the excircle opposite it. Those three lines all meet at a single point called the **Nagel point**.

Figure 7.6: The Nagel Point
```
Tricenters/Nagel.T [M]
```

7.5.1 Geometer Analysis of the Nagel Point

Testing the **Geometer** diagram on figure 7.6 almost instantly shows us how to solve the problem. The relationships found by **Geometer** include the following:

```
Equal length segments:
    [B D] [A E]
Equal length segments:
    [B F] [C E]
Equal length segments:
    [A F] [C D]
```

One of the most common methods used to prove that three lines are concurrent is Ceva's theorem which states that the necessary condition for the three lines in this figure to be concurrent is:
$$\frac{AF}{BF} \cdot \frac{BD}{CD} \cdot \frac{CE}{AE} = 1.$$
which is a trivial consequence of the relationships discovered by **Geometer**. The formulas in Section A.6.8 tell us instantly that $BD = AE$, $BF = CE$ and $AF = CD$, so we are done.

7.6 The First Napoleon Point

Given an arbitrary triangle $\triangle ABC$, erect equilateral triangles on all three sides pointing outward. If the centers of these equilateral triangles are connected to the opposite vertices with line segments as in figure 7.7, those three line segments will be concurrent at a point called the **first Napoleon point**. (If the equilateral triangles are all drawn inward from the faces, and the centers connected in the same way, those lines will also be concurrent in what is called the **second Napoleon point**.

   This is highly reminiscent of the Fermat point (see Section 7.4). The difference is that the vertices of the original triangle are connected to the outermost vertex of a different isosceles triangle erected on each side. For the Fermat point, the triangles were equilateral; in this case, they have angles of $30°$, $30°$ and $120°$.
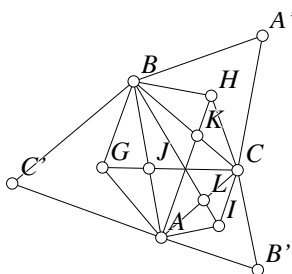
7.6.1 Geometer Analysis of the Napoleon Point

Figure 7.7: The First Napoleon Point
Tricenters/Napoleon.T [M]

In figure 7.7 it seems reasonable to draw a few extra lines and to identify some points since otherwise the bare centers of the triangles will be connected to nothing, and since one of the most common ways to prove the concurrency of lines is with Ceva's theorem, it would be nice to have the points J, K and L identified so that we may be able to work with the cevians in $\triangle ABC$. Thus we draw the segments BH, HC, CI, IA, AG and GB and find points J, K and L that are the intersections of the lines CG, AH and BI with the sides of the original $\triangle ABC$.

   If we test this configuration with **Geometer**, there are quite a few relationships, most of which are trivially true due to the fact that there are three equilateral triangles, and that when we connect the centers of those triangles to the two vertices that are shared with $\triangle ABC$, we form a large set of similar figures. If we look carefully, however, we can find a few relationships that are not so obvious:

```
Equal angles:
    (B C I) (A C H)
Equal area triangles:
    [B C I][A C H]
Triangle area ratios:
    [A B H] x [C H K] = [B K H] x [A C H]
```

   There are equivalent relationships among all the triangles in corresponding positions in the figure, but let us just concentrate for now on this pair.

   From the second relation, we can see that the point K divides the base of $\triangle BCH$ into two triangles whose areas appear in a ratio. But if we think of BC as the base of

$\triangle BCH$, the ratio $BK : CK$ is equal to the ratio of the areas of the smaller triangles $\triangle BKH$ and $\triangle CKH$. If we are going to use Ceva's theorem, we will need to use ratios involving BK and CK, so perhaps we can convert them to ratios of triangle areas.

Using the equal angles found by **Geometer** and the ratios of sides due to the similarity of the equilateral triangles and the included $30° - 30° - 120°$ triangles we can show prove the equality of the triangle areas found by **Geometer**.

$\boxed{7.6.2}$ Proof of the Existence of the First Napoleon Point

Let us begin by showing that $\mathcal{A}(\triangle BCI) = \mathcal{A}(\triangle ACH)$. Since we have equilateral triangles erected on the sides of the triangle as well as the $30° - 30° - 120°$ triangles, it is clear that both $\angle BCI$ and $\angle ACH$ are the sum of $\angle ACB$ and a $30°$ angle, so they are equal.

We know that the area of a triangle is equal to half the product of two sides multiplied by the sine of the included angle, so

$$\mathcal{A}(\triangle BCI) = \frac{BC \cdot CI \sin(\angle BCI)}{2} \text{ and } \mathcal{A}(\triangle ACH) = \frac{AC \cdot CH \sin(\angle ACH)}{2}.$$

Since the two angles above are equal, the ratio of the two areas is therefore:

$$\frac{\mathcal{A}(\triangle BCI)}{\mathcal{A}(\triangle ACH)} = \frac{BC \cdot CI}{AC \cdot CH}. \tag{7.1}$$

But since they are both $30° - 30° - 120°$ triangles, we know that $\triangle ACI \sim \triangle BCH$, and thus $CI/AC = CH/BC$, and this, combined with equation 7.1 tells us that the ratio of the areas is 1, or in other words, $\mathcal{A}(\triangle BCI) = \mathcal{A}(\triangle ACH)$.

Exactly the same technique can be applied to show that $\mathcal{A}(\triangle ABI) = \mathcal{A}(\triangle ACG)$ and $\mathcal{A}(\triangle ABH) = \mathcal{A}(\triangle BCG)$.

If we consider BC to be the base of triangles $\triangle BCH$ and $\triangle BCA$, it is clear that KA and KH divide their areas into equal ratios and the ratio of the areas of $\triangle ACH$ and $\triangle ABH$ will also be the same thing.

From this observation we obtain the following three ratios:

$$\frac{\mathcal{A}(\triangle ACH)}{\mathcal{A}(\triangle ABH)} = \frac{CK}{BK}, \qquad \frac{\mathcal{A}(\triangle ABI)}{\mathcal{A}(\triangle BCI)} = \frac{AL}{CL}, \qquad \frac{\mathcal{A}(\triangle BCG)}{\mathcal{A}(\triangle ACG)} = \frac{BJ}{AJ}$$

If we multiply the three ratios above together and take into account the equality of the pairs of triangle areas, we obtain:

$$\frac{CK}{BK} \cdot \frac{AL}{CL} \cdot \frac{BJ}{AJ} = 1,$$

and this, together with Ceva's theorem, is sufficient to prove that the three lines are concurrent.

7.6.3 A More General Theorem

Notice in the proof above that we did not make use of the fact that the triangles $\triangle ACI$, $\triangle CBH$ and $\triangle BAG$ are $30° - 30° - 120°$ triangles, other than the fact that $30°$ angles are equal to each other. Thus, if *any* similar isosceles triangles are erected on the sides of an arbitrary triangle, the lines connecting the tips of those isosceles triangles with the vertices of the original triangle are coincident. Exactly the same proof will work, so this is an alternative proof of the existence of the Fermat point.

Another commonly-seen triangle center of this type is formed by erecting squares of the sides of a triangle and then connecting the centers of those squares with the opposite vertex of the triangle. This is exactly equivalent to erecting an isosceles $45° - 45° - 90°$ triangle on each edge and connecting the tip to the opposite vertices.

7.7 The Brocard Points

Given a triangle $\triangle ABC$, show that there exists some point Ω such that $\angle \Omega AC = \angle \Omega CB = \angle \Omega BA$. (The character "$\Omega$" is the Greek letter omega, and is the standard notation for the Brocard points.) Find a construction for this point. See figure 7.8.

Figure 7.8: The Brocard Point
`Tricenters/Brocard3.T` [P]

Every triangle has another Brocard point. The other one (called Ω') is obtained by using the opposite angles; in other words, where : $\angle \Omega'AB = \angle \Omega'BC = \angle \Omega'CA$. We will investigate the relationship between the two points Ω and Ω' later in this section.
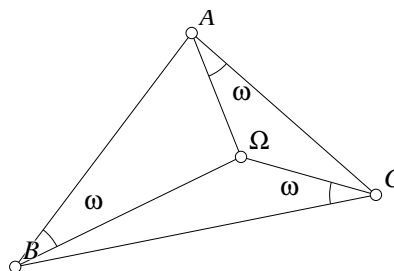
The solution to this problem is similar in a way to the proof that the three angle bisectors meet at a point. The proof of that fact was based on the fact that *any* point on an angle bisector was equidistant from the two rays of the angle, so the point of intersection of two angle bisectors must be equidistant from all three sides, and thus it must lie on the third angle bisector as well.

7.7.1 Geometer Analysis of the Brocard Point

For the Brocard point we can use a similar approach. Suppose we begin by finding all points Ω such that $\angle \Omega CB = \angle \Omega BA$. The set of all solutions will probably lie on some geometric curve. In other words, any point on that curve will satisfy that angle equality. It we do a similar analysis to find all points Y satisfying $\angle YBA = \angle YAC$, those too will lie on another curve. The point (or points) at the intersection(s) of those curves will cause all three angles to be equal, and thus to lie on the curve corresponding to the equality of the third pair of angles.
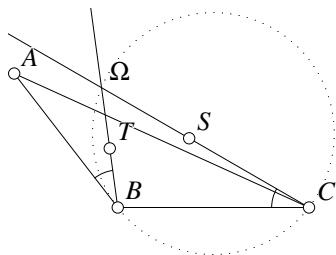
Figure 7.9: Possible Brocard Points
Tricenters/Brocard2.T [P]

Even without any idea how to proceed, we can make a **Geometer** diagram of the triangle with a free point (call it S). Find the angle $\angle SCB$ and copy it to vertex B forming angle $\angle ABT$ as in figure 7.9. Any point Ω lying on the lines CS and BT will cause the two angles $\angle \Omega CB$ and $\angle \Omega BA$ to be equal. As we drag the point S around C we notice that the intersection point Ω seems to sweep out a circle of its own.

In fact, by adjusting the shape of the original triangle (see Section 7.7.2) we notice that the circle of possible locations of Ω passes through B and C and seems to be tangent to the line AB at B.

<div style="margin-left:1em"></div>

$\boxed{7.7.2}$ **Technical Details**

There are three versions of the **Geometer** diagram on the CD: `Brocard1.T`, `Brocard2.T` and `Brocard3.T`. Figure 7.8 was generated with `Brocard3.T` and in fact contains a construction of the Brocard point and a **Geometer** proof that it satisfies the appropriate angle equalities. Figure 7.9 was generated by `Brocard2.T` and is a script that sweeps out a set of evenly-spaced points along the circle of possible solutions.

`Brocard1.T` contains a figure somewhat like the one you would have constructed to study the locus of points Ω that cause the two angles to be the same. In this figure you can move the point S manually and observe how the point Ω moves in response. `Brocard2.T` is a modified version of that diagram that moves S automatically, and in fact the only way to move S in that diagram is with the *Run Script* button in the control area.

Finally, if you try to construct the diagram yourself you may find it difficult to cause the angle at B to have T on the correct side of the ray $\overrightarrow{BA}$. If it is on the wrong side, simply reflect the point T across the line AB (using the *LP=>P Mirror* command under the *Primitives→Point* pulldown menu). This reflected value of T will cause the angle to be on the correct side. After reflecting the point, set the original version of the point T to be the invisible color.

Examine the contents of `Brocard1.T` and `Brocard2.T` with a text editor to learn the details of the construction of those diagrams.

$\boxed{7.7.3}$ **Proof of the Existence of the Brocard Point**

Once we learn that the locus of possible locations of the point Ω relative to one of the triangle's sides is a circle, it is easy to see why it is a circle and how to construct it.
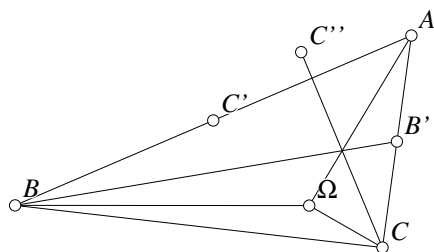
Figure 7.9 shows that $\angle \Omega CB$ is inscribed in the arc ΩB, and that since AB is tangent to the circle and the segment ΩB cuts off the arc ΩB, that $\angle \Omega BA$ must be equal to $\angle \Omega CB$. Thus any point Ω on the circle through C and tangent to AB at B will cause the angle equality $\angle \Omega BA = \angle \Omega CB$ to be satisfied.

Figure 7.10: Brocard, Median and Symmedian Tricenters/Brocard4.T [M]

It is not difficult to construct the circle. We know that its center must lie on the perpendicular bisector of BC since B and C lie on the circle and we know that since the circle is tangent to AB at B its center must lie on the line through B perpendicular to AB. Any point Ω lying on a circle centered at the intersection of those lines and passing through the points B and C satisfies the angle equality. To find the location of the Brocard point, construct two such circles relative to different sides and their intersection will be the Brocard point.

7.7.4 More on the Brocard Points

As we stated at the beginning of Section 7.7, there are two Brocard points depending upon which way the angles are measured. Here are some interesting facts about the two Brocard points Ω and Ω' that you may wish to try to prove, possibly using **Geometer** to help:

- If the lines connecting the vertices of the triangle to Ω form an angle ω with each side, and those connecting the vertices of the triangle to Ω' form an angle ω', then $\omega = \omega'$.

- The line $A\Omega$, when reflected across the angle bisector at $\angle BAC$, is the line $A\Omega'$. Similarly for the other vertices of the triangle[1].

- If O is the circumcenter of $\triangle ABC$, then $O\Omega = O\Omega'$ and $\angle \Omega O\Omega' = 2\omega$.

- The **pedal triangle** of a point relative to a triangle is the triangle that connects the projections of the point on each of the sides of the original triangle. The pedal triangles of Ω and Ω' relative to $\triangle ABC$ are congruent.

- The line passing through point A and the Brocard point Ω, the median from point B and the symmedian from point C all meet at a point. See figure 7.10. The **symmedian** is the median reflected across the angle bisector. In the figure, C' is the midpoint of segment AB and C'' is the reflection of C' across the angle bisector of $\angle BCA$.

---

[1]Two lines through a vertex of a triangle are called **isogonal conjugates** if each is a reflection of the other across the angle bisector. It is an interesting consequence of Ceva's theorem that if three Cevians meet at a point, then so also do their isogonal conjugates.
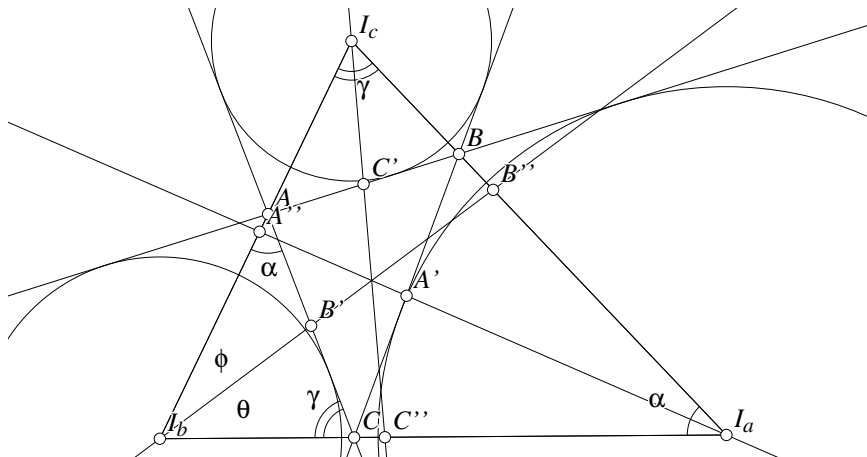
Figure 7.12: The Mittenpunkt
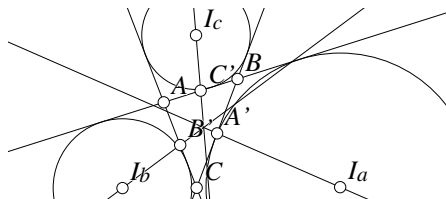`Tricenters/Mittenpunkt.T [M]`

7.8 ♦ The Mittenpunkt

Figure 7.11: The Mittenpunkt
`Tricenters/Mittenpunkt.T [M]`

In an arbitrary triangle $\triangle ABC$, let I_a, I_b and I_c be the centers of the three ex-circles opposite the vertices A, B and C, respectively. Let A', B', and C' be the midpoints of the sides BC, CA and AB, respectively. Then the lines I_aA', I_bB' and I_cC' are concurrent, and meet at a point called the **mittenpunkt**.

7.8.1 Geometer Analysis of the Mittenpunkt

If we test the diagram in figure 7.11 it does not tell us too much, but there are the following interesting relationships:

```
Points on a circle:
   B  C  I\sub{c}  I\sub{b}
Points on a circle:
   A  C  I\sub{c}  I\sub{a}
Points on a circle:
   A  B  I\sub{a}  I\sub{b}
```

We also realize that our best bet to prove the concurrency of three lines is Ceva's theorem, but to use that, we need a triangle where the lines pass through the vertices and that is not the case when we begin. We do know that the centers of the excircles

lie on the exterior angle bisectors, so we know that A, B, and C lie on the lines I_bI_c, I_cI_a and I_aI_b, respectively. A suitable triangle might be $\triangle I_aI_bI_c$, so let A'', B'' and C'' be the points where the lines I_aA', I_bB' and I_cC' intersect I_bI_c, I_cI_a and I_aI_b, respectively. See figure 7.12.

The initial **Geometer** analysis indicated that there are some sets of concyclic points in the initial diagram, which tells us that there are some angle equalities or at the least, pairs of angles that are supplementary. However, even after we have added the lines to make $\triangle I_aI_bI_c$ as in figure 7.12 and tested it, **Geometer** does not report any such relationships among the angles. Why is that?

The reason is that **Geometer** has no way of knowing that the points A, B and C lie on the lines of that triangle. When **Geometer** tests angles, it only tests angles that consist of pairs of line segments that intersect in known points. If we add the segments I_aB, I_aC, I_bA, I_bC, I_cA and I_cB to the diagram and test again, suddenly there is plenty of additional information about equal angles, a small amount of which is listed below. These equalities are symmetric in that each refers to a different angle in the triangle $\triangle I_aI_bI_c$.

```
Equal angles:
    (A B I\sub{c}) (C B I\sub{a}) (I\sub{c} I\sub{b} I\sub{a})
Equal angles:
    (B A I\sub{c}) (C A I\sub{b}) (I\sub{c} I\sub{a} I\sub{b})
Equal angles:
    (B C I\sub{a}) (A C I\sub{b}) (I\sub{b} I\sub{c} I\sub{a})
```

We will also need to make use of the fact that A', B' and C' are the midpoints of the sides of the original triangle, and to use Ceva's theorem, we also have to find equations that involve the sides I_aB, I_cB, and so on.

A look at the diagram shows us that there are six pairs of triangles that are like $\triangle I_bB'C$ and $\triangle I_bB''I_a$ in that they share an angle, they involve the lengths that are equal and that will be needed to apply Ceva's theorem, and they repeatedly make use of other lengths and equal angles in the diagram.

If we apply the law of sines to each, Ceva's theorem can be applied and the result proved.

7.8.2   Proof of the Mittenpunkt's Existence

Consider the two pairs of triangles $(\triangle I_bB'C, \triangle I_bB''I_a)$ and $(\triangle I_bB'A, \triangle I_bB''I_c)$.

We first note that the lines connecting I_a, I_b and I_c bisect the exterior angles of $\triangle ABC$. This, plus a little angle chasing shows us that $\angle I_bI_aI_c = \angle I_bAC = \alpha$ and that $\angle I_bI_cI_a = \angle I_bCA = \gamma$. If we assign the names $\theta = \angle CI_bB'$ and $\phi = \angle AI_bB'$ as in the figure, we can apply the law of sines to each of the four triangles mentioned in the previous paragraph:

$$\frac{\sin\theta}{CB'} = \frac{\sin\gamma}{I_bB'} \quad \text{and} \quad \frac{\sin\theta}{I_aB''} = \frac{\sin\alpha}{I_bB''} \tag{7.2}$$

$$\frac{\sin\phi}{AB'} = \frac{\sin\alpha}{I_bB'} \quad \text{and} \quad \frac{\sin\phi}{I_cB''} = \frac{\sin\gamma}{I_bB''} \tag{7.3}$$

To use Ceva's theorem, we will need to calculate the ratio $I_a B''/I_c B''$. Recalling that $CB' = AB'$ (since B' is the midpoint of AC), we can combine the four equations 7.2 and 7.3 and cancel out the equal terms to obtain:

$$\frac{I_a B''}{I_c B''} = \frac{\sin^2 \gamma}{\sin^2 \alpha}. \tag{7.4}$$

Using completely symmetric methods, and labeling $\angle I_a I_b I_c = \beta$ we get:

$$\frac{I_c A''}{I_b A''} = \frac{\sin^2 \beta}{\sin^2 \gamma} \quad \text{and} \quad \frac{I_b C''}{I_a C''} = \frac{\sin^2 \alpha}{\sin^2 \beta} \tag{7.5}$$

Combining the equations 7.4 and 7.5 yields:

$$\frac{I_a B''}{I_c B''} \frac{I_c A''}{I_b A''} \frac{I_b C''}{I_a C''} = \frac{\sin^2 \gamma}{\sin^2 \alpha} \frac{\sin^2 \beta}{\sin^2 \gamma} \frac{\sin^2 \alpha}{\sin^2 \beta} = 1,$$

which, together with Ceva's theorem, proves the concurrence of the three lines at the mittenpunkt.
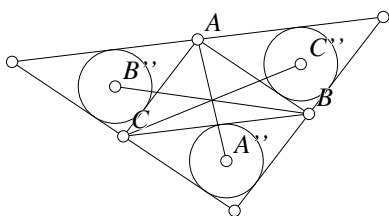
7.9 A New Center

Figure 7.13: A New Center
`Tricenters/Tcenter.T` [M]

It is quite easy to find additional centers. With just a little experimentation using geometer, the following construction appears to generate a triangle center. See figure 7.13.

Given a triangle, construct what is sometimes called its **superior triangle** by drawing lines through each vertex parallel to the opposite side. This will produce an inverted triangle similar to the original with four times the area and in fact, apparently composed of four congruent copies of the original triangle.

In each of the three outer triangles, construct the incircle and connect the three centers of those incircles to the opposite vertex of the original triangle. Those three lines appear to meet at a point. Do they? Testing the diagram with **Geometer** seems to indicate that they do. Unfortunately, other than indicate that the lines seem to be concurrent, **Geometer** does not give much help toward finding a proof.

7.9.1 Geometer Analysis of the Problem

Try drawing some additional line segments and run **Geometer**'s test again. After a few tries, one can find an interesting set of segments: connect A and B with C'', B and C with A'' and A and C with B''. With these additions, **Geometer**'s test of the diagram yields the following interesting relationships:

こ

```
Parallel lines:
    (BA'') (AB'')
Parallel lines:
    (BC'') (CB'')
Parallel lines:
    (AC'') (CA'')
Equal length segments:
    [B A''] [A B'']
Equal length segments:
    [B C''] [C B'']
Equal length segments:
    [A C''] [C A'']
```

If these are true, it is clear that $AC''A''C$, $BC''B''C$ and $AB''A''B$ are all parallelograms whose diagonals bisect each other. Since every diagonal bisects the other, all three must meet at a point.

Since all three exterior triangles are congruent, it is clear that the lengths in the list above are equal. That they are parallel is also clear since, for example, BC'' and CB'' are angle bisectors of corresponding angles in congruent triangles that have parallel sides.

This is a very rough sketch of the proof, but it is not difficult to fill in the details.

7.10 Additional Triangle Centers

There are, in fact, an infinite number of triangle centers, many of which, together with their properties, are listed in [Kimberling, 1994]. Without proof, here is a list of a few more of them, described in geometric terms.

- **The Lemoine Point.** If each of the medians of a triangle is reflected across the corresponding angle bisector, then those three reflections are coincident at the **Lemoine point**.

- **The Spieker Center.** The existence of the **Spieker center** is obvious—it is just the incircle of the triangle whose vertices are the midpoints of the sides of the original triangle (called the **medial triangle**). Physically, it is the center of mass of the perimeter of the triangle. In other words, if a uniform wire were bent into the shape of the triangle, it would balance at the Spieker center.

- **The Feuerbach Point.** This is the point of tangency of the incircle of a triangle with its nine-point circle. The existence of this point is proved by Feuerbach's theorem. See Section 8.17. The Feuerbach point is not precisely a triangle center, since it fails to exist for an equilateral triangle. When the triangle is equilateral, the nine-point circle and the incircle are the same so there are an infinite number of points of tangency.

It is quite easy, especially with a computer geometry program, to experiment with various symmetric combinations of lines and to find other possible triangle centers. Often these centers turn out to be one of the many known centers and if that is the case, it is usually easy to prove that the new center makes sense.

7.11 Trilinear Coordinates

Trilinear coordinates provide a method to locate points on the plane relative to some given fixed triangle $\triangle ABC$ (having sides a, b and c opposite the corresponding angles). The trilinear coordinates of a point are just a ratio of the (directed) distances from the point to the sides of the triangle. If only points inside the triangle are considered there is no need for the "directed" in the definition, but in fact trilinear coordinates can be used to identify any point of the plane, be it inside, outside, or on the triangle. The "directed" in "directed distance" means that the distance is positive if the point is on the same side of the edge as the triangle and negative otherwise[2].

The trilinear coordinates of a point are written "$\alpha : \beta : \gamma$" where α, β and γ are three real numbers, at least one of which is non-zero. The three numbers are proportional to the directed distances of the point from the sides a, b and c, respectively.

The trilinear coordinates of the incenter are $1 : 1 : 1$ since the incenter is equally far from each of the sides, and it is inside the triangle. The trilinear coordinates of the three excenters are similar: $-1 : 1 : 1$, $1 : -1 : 1$ and $1 : 1 : -1$ since those points are also equidistant from the three edges, but they are outside the triangle in one of three directions.

The trilinear coordinates of a point are not unique. If $\alpha : \beta : \gamma$ correspond to a point, then so do $\kappa\alpha : \kappa\beta : \kappa\gamma$, where κ is any non-zero real number. Thus the coordinates of the incenter could equally well have been written: $2 : 2 : 2$, $-1 : -1 : -1$, $2/3 : 2/3 : 2/3$, or even $-\pi : -\pi : -\pi$. This is not such a bizarre concept; we already do something similar with fractions. Although $3/5$ is usually not expressed that way, we all agree that it is equal to $6/10$ or even $(-3\pi)/(-5\pi)$. Coordinates that are unique only up to a real multiple are said to be **homogeneous coordinates**, so trilinear coordinates are sometimes referred to as "homogeneous trilinear coordinates".

At least one of the coordinates must be non-zero, but up to two of them may be zero. The point A has trilinear coordinates $1 : 0 : 0$ (or $\alpha : 0 : 0$), and any point on the line b has trilinear coordinates $\alpha : 0 : \gamma$ since it is a distance 0 from the line b but can be anywhere relative to the other two lines.

As one non-trivial example, consider the centroid of a triangle (the intersection of the three medians). The trilinear coordinates of the centroid are $1/a : 1/b : 1/c$. This is not too hard to see: Remember that the medians divide the triangle into six smaller triangles of equal area, so any two of them will amount to $1/3$ the total triangle area. If we drop perpendiculars to each of the sides of the triangle, these can be thought of as altitudes of triangles $\triangle ABG$, $\triangle BCG$ and $\triangle CAG$, where G is the centroid. If the altitudes to sides a, b and c have lengths h_a, h_b and h_c then from the equality of the areas we obtain: $ah_a = bh_b = ch_c$. Thus the values of h_a, h_b and h_c (which are the distances to the sides) are proportional to $1/a : 1/b : 1/c$, which is what we wished to show.

---

[2]This is not *exactly* correct, since we will allow multiplication of all the coordinates by an arbitrary value, positive or negative. A point inside the triangle will have all its directed distances positive, but if all are multiplied by -1 they will all be negative and no such point exists on the plane. If a set of trilinear coordinates is nonsensical, flip the signs on all of them to obtain a set that makes geometric sense.

Incenter	$1:1:1$
Excenters	$-1:1:1$, $1:-1:1$ and $1:1:-1$
Centroid	$\csc A : \csc B : \csc C$
	$1/a : 1/b : 1/c$
Circumcenter	$\cos A : \cos B : \cos C$
	$a(b^2 + c^2 - a^2) : b(c^2 + a^2 - b^2) : c(a^2 + b^2 - c^2)$
Orthocenter	$\sec A : \sec B : \sec C$
Nine-point center	$\cos(B - C) : \cos(C - A) : \cos(A - B)$
Gergonne point	$\sec^2(A/2) : \sec^2(B/2) : \sec^2(C/2)$
	$1/a(b + c - a) : 1/b(c + a - b) : 1/c(a + b - c)$
Nagel point	$\csc^2(A/2) : \csc^2(B/2) : \csc^2(C/2)$
	$(b + c - a)/a : (c + a - b)/b : (a + b - c)/c$
Mittenpunkt	$\cot(A/2) : \cot(B/2) : \cot(C/2)$
	$b + c - a : c + a - b : a + b - c$
Spieker center	$(b + c)/a : (a + c)/b : (a + b)/c$
Feuerbach point	$1 - \cos(B - C) : 1 - \cos(C - A) : 1 - \cos(A - B)$
Lemoine point	a:b:c
Fermat point	$\csc(A + 60°) : \csc(B + 60°) : \csc(C + 60°)$
1$^{\text{st}}$ Napoleon point	$\csc(A + 30°) : \csc(B + 30°) : \csc(C + 30°)$

Table 7.1: Some Trilinear Coordinates

It is an interesting exercise to determine sets of trilinear coordinates for other special points. Table 7.1 lists the trilinear coordinates for all the circle centers discussed in this chapter. Sometimes two versions of the coordinates are listed, and when that is the case, both are equally valid (one set is a multiple of the other). It is not a difficult exercise to see why the trilinear coordinates for the centroid can written equally well as $1/a : 1/b : 1/c$ or $\csc A : \csc B : \csc C$. Do you think that the $60°$ and the $30°$ in the trilinear coordinates of the Napoleon and Fermat points might have something to do with the fact that they are constructed by erecting isosceles triangles with base angles of $60°$ and $30°$? It might be worth investigating.

7.12   Triangle Center Finder

Since virtually all the triangle centers can be expressed in terms of their trilinear coordinates as displayed in Table 7.1 we can make a **Geometer** diagram that will help us discover if some triangle center we find is the same as one that has already been discovered.

The diagram `Tricenters/Finder.T` does just that.

It uses trilinear coordinates for an arbitrary $\triangle ABC$ to draw many of the more common triangle centers on that triangle, but on different layers.

None of the centers is drawn on layer zero, so to use this diagram, load it into **Geometer** and draw your newly-discovered triangle center on layer zero. Draw the

final center in a different color (let's say you've made it cyan). Now edit the **Geometer** code and find your point. It will be near the end of the file (probably the very last line in the file, in fact), and it will contain text that looks something like this:

```
v27 = .v.ll(112, 141, [.cyan, .in]);
```

(This assumes you made the final point as the intersection of two lines—all that is really important is this part: [.cyan, .in] which says that the point is painted cyan on layer zero and is invisible on all the others. Change "[.cyan, .in]" to ".cyan", and exit the editor. Now your point will be visible on all layers.

To see which triangle center it is (if any), press the *Next* button until your point seems to line up with one of the presented points. Then, to make sure, move the triangle vertices around and make sure that your point stays locked on that center. If not, continue the search.

Of course not all centers are listed—just a few of the more common ones. For a much more exhaustive list, see [Kimberling, 1994]. Almost any of Kimberling's centers could be added to this diagram, and he lists more than 100 of them.

Here's an exercise to test the method. Try to figure out which triangle center this is: Construct the medians of a triangle, but reflect each one across the angle bisector of the angle from which it emanates. Those three reflections of the median lines are also concurrent at a point. (In fact, the point is the isogonal conjugate of the median, but it has a special name. What is it?)

Here is some of the **Geometer** code for the triangle center locator. See the file on the CD for a complete listing.

```
.geometry "version 0.40";
.l0;
.text("Triangle Center Locator: On this figure,
construct the center that interests you. Make
the final point non-white (cyan, for example).
Edit the geometry, and you will find toward the
end your cyan point containing '[.cyan, .in]'.
Change this to '.cyan'. Then step through
the 'proof'", .l0);
.macro .vertex trilinear(.vertex v1, .vertex v2, .vertex v3,
                                    .flt f1, .flt f2, .flt f3)
{
    ang1 = .a.vvv(v3, v1, v2, .in);
    ang2 = .a.vvv(v1, v2, v3, .in);
    ang3 = .a.vvv(v2, v3, v1, .in);
    a1 = .f.rpn(f2, ang1, .sin, .mul, f3,
        f2, ang1, .cos, .mul, .add,
        .atan2);
    a2 = .f.rpn(f3, ang2, .sin, .mul, f1,
        f3, ang2, .cos, .mul, .add,
        .atan2);
    a3 = .f.rpn(f1, ang3, .sin, .mul, f2,
        f1, ang3, .cos, .mul, .add,
        .atan2);
    A1 = .a.f(a1);
```

```
    A2 = .a.f(a2);
    A3 = .a.f(a3);
    va1 = .v.avv(A1, v3, v1, .in);
    va2 = .v.avv(A2, v1, v2, .in);
    va3 = .v.avv(A3, v2, v3, .in);
    l1 = .l.vv(v1, va1, .in);
    l2 = .l.vv(v2, va2, .in);
    l3 = .l.vv(v3, va3, .in);
    .return v4 = .v.ll(l1, l2, .in);
}
v1 = .free(-0.305389, 0.511976, "A");
v2 = .free(0.691617, -0.45509, "B");
v3 = .free(-0.811377, -0.236527, "C");
a1 = .a.vvv(v3, v1, v2, .in, "\alpha");
a2 = .a.vvv(v1, v2, v3, .in, "\beta");
a3 = .a.vvv(v2, v3, v1, .in, "\gamma");
a = .f.vv(v2, v3);
b = .f.vv(v1, v3);
c = .f.vv(v1, v2);
incenter = trilinear(v1, v2, v3, 1.00, 1.00, 1.00, .l1, "I");
l1 = .l.vv(v1, v2);
l2 = .l.vv(v2, v3);
l3 = .l.vv(v3, v1);
excenter1 = trilinear(v1, v2, v3, -1.00, 1.00, 1.00, .l1, "E\sub{1}");
excenter2 = trilinear(v1, v2, v3, 1.00, -1.00, 1.00, .l1, "E\sub{2}");
excenter3 = trilinear(v1, v2, v3, 1.00, 1.00, -1.00, .l1, "E\sub{3}");
ainv = .f.rpn(1.000000, a, .div);
binv = .f.rpn(1.000000, b, .div);
cinv = .f.rpn(1.000000, c, .div);
centroid = trilinear(v1, v2, v3, ainv, binv, cinv, .l1, "M");
circa = .f.rpn(a1, .cos);
circb = .f.rpn(a2, .cos);
circc = .f.rpn(a3, .cos);
circumcenter = trilinear(v1, v2, v3, circa, circb, circc, .l1, "O");
ortha = .f.rpn(1.000000, circa, .div);
orthb = .f.rpn(1.000000, circb, .div);
orthc = .f.rpn(1.000000, circc, .div);
orthocenter = trilinear(v1, v2, v3, ortha, orthb, orthc, .l1, "H");
v4 = .v.vvmid(circumcenter, orthocenter, .l1, "N");
.text("Elementary Centers:
I, E\sub{1}, E\sub{2}, E\sub{3}: Incenter and 3 Excenters
O, M, H: Circumcenter, Centroid, Orthocenter
N: Nine-point center.", .l1);
lemoine = trilinear(v1, v2, v3, a, b, c, .l2, "L");
gerg1 = .f.rpn(1.000000, a1, 0.500000, .mul, .cos,
        .dup, .mul, .div);
gerg2 = .f.rpn(1.000000, a2, 0.500000, .mul, .cos,
        .dup, .mul, .div);
gerg3 = .f.rpn(1.000000, a3, 0.500000, .mul, .cos,
        .dup, .mul, .div);
gergonne = trilinear(v1, v2, v3, gerg1, gerg2, gerg3, .l2, "G");
nag1 = .f.rpn(b, c, .add, a, .sub,
        a, .div);
nag2 = .f.rpn(c, a, .add, b, .sub,
        b, .div);
nag3 = .f.rpn(a, b, .add, c, .sub,
        c, .div);
```

```
nagel = trilinear(v1, v2, v3, nag1, nag2, nag3, .l2, "N");
mit1 = .f.rpn(b, c, .add, a, .sub);
mit2 = .f.rpn(c, a, .add, b, .sub);
mit3 = .f.rpn(a, b, .add, c, .sub);
mittenpunkt = trilinear(v1, v2, v3, mit1, mit2, mit3, .l2, "M");
.text("More centers:
G, L: Gergonne Point, Lemoine Point
N, M: Nagel Point, Mittenpunkt.", .l2);
```

Although there is quite a bit of it (and even more in the actual **Geometer** file), the structure is not too complex. First, a macro is defined that calculates the location of a point given the three vertices of the triangle and the three trilinear coordinates. Later we will discuss how it works.

After the macro is defined, a long series of points are drawn by passing the macro the appropriate trilinear coordinates. For something as simple as the incenter with trilinear coordinates $1 : 1 : 1$, the code simply calls the macro with the vertices A, B and C of the triangle together with 1, 1 and 1. The output of the macro is labeled to indicate which center it is.

After the incenter and the three excenters is the generation of the centroid with trilinear coordinates $1/a : 1/b : 1/c$. First is a calculation (using the rpn calculator inside **Geometer**) of the values of $1/a$, $1/b$ and $1/c$ followed by a call on the macro.

The same procedure is followed for points with more complex descriptions, but although the calculations are a bit messier, the idea is the same.

Now, how does the `trilinear` macro itself work? It is based on the idea that the locus of points whose distance to two fixed lines have a constant ratio is a straight line passing through the intersection of the lines. If we can find a single additional point on the line, we can draw it. These lines of constant ratio corresponding to the trilinear coordinates are drawn from two vertices of the triangle, and their intersection must match the trilinear ratios with all three lines.

7.13 Searching for the "True" Center of a Triangle

We have learned that unless you are speaking about an equilateral triangle, there are many different points that can be considered to be the center of a triangle, and depending upon your application, there are times when each of them is the one that makes sense.

But here is an interesting topic to investigate that will do two things. First, it will teach you something about geometry, and second, if you want, it can teach you about some of **Geometer**'s more advanced features, particularly macros[3].

---

[3]This problem was inspired by a wonderful lecture given in about 1968 by Richard Feynman who wanted to find the "true average" of two numbers. He took the arithmetic mean and the geometric mean, but since those are not necessarily the same he repeated the process with those two means. The pair of numbers obtained in this manner converges extremely rapidly, but unfortunately to a number that can only be expressed as a complete elliptic integral.

Here is the problem: We know that in general, the centroid, nine point center, and incenter of a triangle are different points. But if there is a "true" center of the triangle, perhaps they are close to it. So form the triangle from these three points and find its "true" center. Since we do not know where the true center is, we can perhaps approximate it by finding the centroid, nine point center, and incenter of this new triangle. Repeat the process. Do the sets of three points tend to a limit inside the triangle? Where is it? What if we had used different centers, like the centroid, orthocenter, and Fermat point?

We can investigate the convergence of these sets of points by using **Geometer** to repeatedly calculate the three centers, then the three centers of the triangle formed by those three centers, and so on. Unfortunately, if the centers do get close together, it will become harder and harder to do the construction.

The solution is to write **Geometer** macros and then to call those repeatedly to see if the points converge. The file `Tricenters/TrueCenter.T` does just this. Here is a listing:

```
.geometry "version 0.60";
v1 = .free(-0.320359, -0.227545, "A");
v2 = .free(0.185629, 0.314371, "B");
v3 = .free(0.0479042, -0.505988, "C");
.macro .vertex centroid(.vertex v1, .vertex v2, .vertex v3)
{
    v4 = .v.vvmid(v1, v2, .in);
    v6 = .v.vvmid(v3, v1, .in);
    l1 = .l.vv(v2, v6, .in);
    l2 = .l.vv(v3, v4, .in);
    .return v7 = .v.ll(l1, l2);
}
.macro .vertex ninepointcenter(.vertex v1, .vertex v2, .vertex v3)
{
    v12 = .v.vvmid(v1, v2, .in);
    v23 = .v.vvmid(v2, v3, .in);
    v31 = .v.vvmid(v3, v1, .in);
    c1 = .c.vvv(v12, v23, v31, .in);
    .return v4 = .v.ccenter(c1);
}
.macro .vertex incenter(.vertex v1, .vertex v2, .vertex v3)
{
    l1 = .l.vv(v1, v2, .in);
    l2 = .l.vv(v2, v3, .in);
    l3 = .l.vv(v3, v1, .in);
    c1 = .c.lll(l1, l2, l3, 1, .in);
    .return v4 = .v.ccenter(c1);
}
vg0 = centroid(v1, v2, v3, .magenta);
vo0 = ninepointcenter(v1, v2, v3, .magenta);
vi0 = incenter(v1, v2, v3, .magenta);
vg1 = centroid(vg0, vo0, vi0, .red);
vo1 = ninepointcenter(vg0, vo0, vi0, .red);
vi1 = incenter(vg0, vo0, vi0, .red);
vg2 = centroid(vg1, vo1, vi1, .yellow);
vo2 = ninepointcenter(vg1, vo1, vi1, .yellow);
vi2 = incenter(vg1, vo1, vi1, .yellow);
```

The diagram begins with the three points of the original triangle drawn in white. Next are three macros to construct the three centers that interest us. The vertices `vg0`, `vo0` and `vi0` are the centroid, nine point center and incenter of that triangle. Replace the `0`s with `1`s and we obtain the second level of centers, et cetera. Each set is drawn in a different color to see if the points converge or not. Experiment with the diagram and see what you can learn. See what happens if you go another step or two deeper.

Sometimes bad things happen. Which ones are due to round-off error in the computer, and which are due to features of the original triangle?

If you wish to experiment with different triangle centers, the easiest way to build a macro is this: In the original triangle using the graphical user interface, construct the center in the normal way. For example, if you are just beginning and want to build the macro to draw the incenter, begin with a diagram having three free points and construct the incenter. Your **Geometer** diagram will look something like this:

```
.geometry "version 0.60";
v1 = .free(-0.320359, -0.227545, "A");
v2 = .free(0.185629, 0.314371, "B");
v3 = .free(0.0479042, -0.505988, "C");
v4 = .v.vvmid(v1, v2, .in);
v6 = .v.vvmid(v3, v1, .in);
l1 = .l.vv(v2, v6, .in);
l2 = .l.vv(v3, v4, .in);
v7 = .v.ll(l1, l2);
```

In the text editor it is an easy exercise to convert this code to the macro form in the original listing. The portion of the code above that will go into the macro includes the line beginning with `v4 = .v.vvmid` and continues to the end of the listing. To test it before going on, add a single line of code (that you may wish to erase later) so that your entire program looks like this:

```
.geometry "version 0.60";
v1 = .free(-0.320359, -0.227545, "A");
v2 = .free(0.185629, 0.314371, "B");
v3 = .free(0.0479042, -0.505988, "C");
.macro .vertex centroid(.vertex v1, .vertex v2, .vertex v3)
{
    v4 = .v.vvmid(v1, v2, .in);
    v6 = .v.vvmid(v3, v1, .in);
    l1 = .l.vv(v2, v6, .in);
    l2 = .l.vv(v3, v4, .in);
    .return v7 = .v.ll(l1, l2);
}
testvertex = centroid(v1, v2, v3, .magenta);
```

If it does the right thing, save your file and start working on the definition of the macro for the second triangle center.
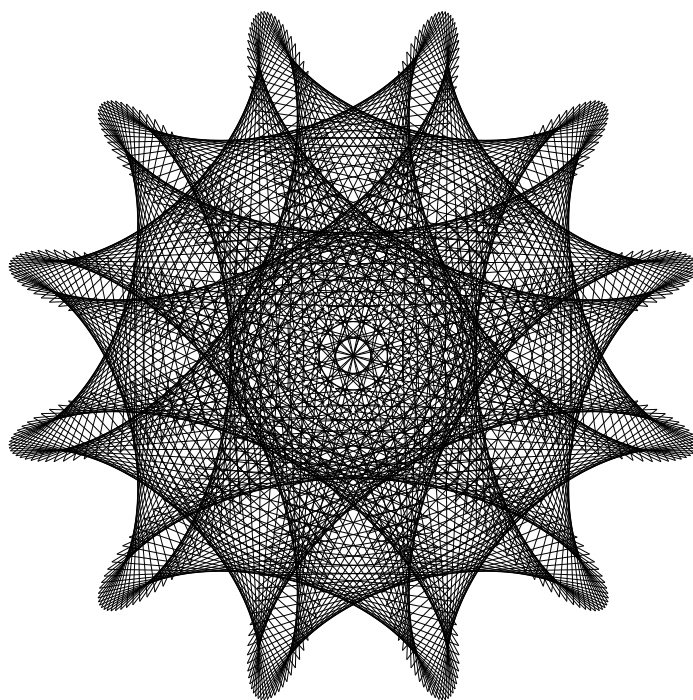
Chapter 8

Inversion in a Circle

Four circles to the kissing come,
The smaller are the better.
The bend is just the inverse of
The distance from the centre.
Though their intrigue left Euclid dumb
There's now no need for rule of thumb.
Since zero bend's a dead straight line
And concave bends have minus sign,
The sum of squares of all four bends
Is half the square of their sum[1].

---

[1]This poem is the Descartes circle theorem, not proved in this book. If $r_1 = 1/a$, $r_2 = 1/b$, and $r_3 = 1/c$ are the radii of three mutually tangent circles, then the radius $r_4 = 1/d$ of the circles tangent to all three satisfies the equation $2(a^2 + b^2 + c^2 + d^2) = (a + b + c + d)^2$ There are two solutions to the equation corresponding to the two circles tangent to the three given circles. The negative solution corresponds to the surrounding circle and the positive one to the surrounded circle. See figure 8.1

Frederick Soddy

8.1 Overview of Inversion

Inversion in a circle is an powerful tool that can be applied to many problems in Euclidean geometry. **Geometer** has built-in commands supporting inversion in a circle. Any time you are faced with the construction of circles that are tangent to a combination of lines and circles, it is wise to consider inversion as a possible technique for finding a solution.
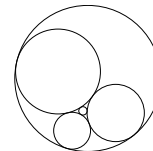
Figure 8.1: Soddy Circles
`Inversion/Soddy.D [D]`

In normal usage, the word "inversion" means "turning inside-out", and this is close to the technical geometrical meaning. The key idea is this: if some figure is inverted (turned inside-out) and then the result is inverted again, the final figure will be identical to the original.

The technique of inversion is very much like changing coordinate systems to simplify a problem. In computer graphics, for example, there is a very simple equation to calculate the position of an object after it is rotated about the origin, but the equations to calculate rotation about an arbitrary point in space are fairly ugly. Thus, to perform rotation about an arbitrary point, the usual technique is to translate all the coordinates so that the point of rotation is at the origin. The desired rotation is then applied to objects expressed in these new coordinates, and finally, the result is translated back by undoing the translation (in other words, by translating in the opposite direction).

Inversion turns the points of the plane inside-out relative to some circle—all the points that were inside the circle wind up on the outside and vice-versa. The points on the circle of inversion itself are not moved. Inversion is a sort of reflection through a circle.

Although inversion is defined simply in terms of its effects on points, what is far more interesting is what inversion does to lines and circles. To invert a line, for example, invert every point on the line through the circle. The resulting set of points is the inverse of the line. The same idea applies to circles.

We will show that the inverted image of a line or circle is another line or circle. By carefully choosing the circle of inversion, you can control when a line is inverted to become a circle, and when it remains a line. Similarly, you can invert circles to become lines, to become other circles, or even make them invert into themselves.

Inversion also supplies the option of sending "troublesome" points to "the point at infinity" where they are so far away that they no longer cause problems.

8.1.1 Inversion of Points

Inversion is always relative to a circle in the plane. In this overview, we will invert through a circle K with center O. Inversion in the circle K (or relative to K) is an

operation on the points of the plane that moves almost every point to a new location. If the point was outside K, its inverse will be inside, and vice-versa. If the point is on the circle K inversion leaves it unchanged (moves it to itself).
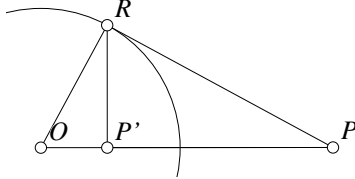
Inversion always moves a point directly toward or away from the center of the circle K. In other words, if you draw a ray connecting the center O of K through the point P, the inverse of P will lie on the same ray $\overrightarrow{OP}$. If K has radius $r = OR$, and if point P is inverted through K to point P', then the formula relating the lengths is this (see figure 8.2:

Figure 8.2: Inversion in a Circle
Inversion/Inversion.T [M]

$$\frac{OP}{OR} = \frac{OR}{OP'}, \quad \text{or} \quad OP \cdot OP' = r^2. \quad (8.1)$$

If you invert a point P in circle K to obtain a point P', then if you invert P' in K, the result will be the original point P. In other words, repeated inversions in the same circle take a point back and forth between two positions. Another way to think about it is that a second identical inversion "undoes" the first. This is clear from equation 8.1. Mathematically, we say that the operation of inversion in a circle is self-inverse.

Inversion is much like the operation of inversion in the real numbers that takes a number x and converts it to the number $1/x$. Inverting twice in the real numbers leaves the number unchanged[2] ($2 \rightarrow 1/2 \rightarrow 1/(1/2) = 2$, or $x \rightarrow 1/x \rightarrow 1/(1/x) = x$).

8.1.2 ♦ Inversion of the Circle Center

Inversion has the property that points near the edge of K are inverted to points also near the edge, but on the other side. Points close to the edge of K, but inside K are inverted to points that are also close to the edge of K, but on the outside, and vice-versa.

Conversely, points very near the center of K (in other words, almost as far from the edge as possible on the inside) are mapped to points very far away from the circle. In fact, if you move a point gradually toward the center of the circle K of inversion, its inverse image moves more and more rapidly away from the center, or "toward infinity".

In the Euclidean plane, the point at the center of K cannot be inverted—it would have to go to a point "infinitely far" from the center. If you imagine adding a "point at infinity" to the Euclidean plane, you can make a mathematically sensible definition for inversion that includes this "point at infinity". The inverse of the center of K is the point at infinity, and the inverse of the point at infinity is the center of the circle of inversion.

---

[2]Inversion in the real numbers is actually *very* similar to geometric inversion in a circle. Numbers inside the range $(-1, 1)$ are inverted to numbers outside that range, -1 and 1 are fixed, and geometric inversion of the center of the circle has exactly the same problem as inversion in the line has with the number zero.

8.1.3 Inversion of Shapes

To invert a shape such as a line or a circle, all that is required is to invert every one of its points. We will show that the inverses of lines and circles are also (almost) lines and circles, but at first it is a bit confusing because inverted lines can become either lines or circles, and similarly for inverted circles.

It is often fruitful to think about this by lumping lines and circles together, and just think of a line as an "infinitely large" circle. In the rest of this paragraph, quotation marks around the word "circle" will indicate that the word means "either a circle or a straight line". This makes some sense, since small parts of very big circles begin to look more and more like straight lines. Since this is the case, why not just think of straight lines as special "circles" that have an infinitely large radius? If you think of it that way, then the theorems about the inversion of lines and circles all look the same. If you invert a "circle" in K, you get another "circle"[3].

8.2 Formal Definition of Inversion

The definition of what it means to invert a point in a circle is simple, but to get an intuitive feeling for inversion, it is a good idea to experiment with the **Geometer** diagram for figure 8.3 that allows you to move the point P and see how its inverse moves in response. Here are some ideas: Try moving P toward the center and away. Place P on the edge of the circle. Try to move P along various straight lines or circles and watch the path of P'.

Given a circle with center O and a point P as in figure 8.3, if P is outside the circle, construct a tangent to the circle at R through P, connect P and O, and then drop a perpendicular from R to PO, intersecting PO at a point P'. This P' is the inverse of the point P relative to that circle. To find the inverse if P is inside the circle, do just the opposite: connect O and P with a line, construct a perpendicular to OP at P that hits the circle at R, then construct a tangent to the circle through R which will hit the line OP at a point

Figure 8.3: Inversion in a Circle
`Inversion/Inversion.T [M]`

P'. Clearly if P' is the inverse of P relative to some circle, then P is the inverse of P' relative to the same circle.

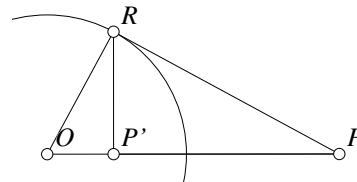There are two other cases to consider where the construction above fails. What if P is on the circle, and what if P is at the center of the circle?

If you experimented with a computer geometry diagram, the answer to the first question should be obvious—the closer P is to the edge of the circle, the closer P' is

---

[3]In fact, there is no reason that K itself needs to be a circle—you can invert in any "circle", even if it is a straight line. Inversion in a line (thought of as an infinitely large circle) turns out to be exactly the same as a mirror reflection of the points across the line.

to P. It only makes sense to define the inverse of a point on the circle to be the point itself, so that is how it is defined.

If P is at O, the center of the circle, it is another story. The closer you move P to O, the further away P' moves. If you look at what happens in our geometrical construction described above, you will see that the closer P is to O, the closer RP is to being parallel to OP. When P is at O, the construction will yield a pair of parallel lines. For this reason, we will simply say that the inverse of the center of the circle is undefined.

"Undefined" is used in the same sense as in arithmetic when we say that 1 divided by 0 is undefined—there is simply no sensible way to define "$1/0$" so that it makes mathematical sense. In the same way that even though it is undefined, you can think of $1/0$ as being ∞ (after all, $1/\epsilon$ is *very* large if ϵ is *very* small), you can think of the inverse of the center of the circle as being a "point at infinity". You cannot use this idea in a formal proof, but nobody can prevent you from thinking about it like that, and thinking about it like that can sometimes be very useful.

We only have to make the point undefined in pure Euclidean geometry where we are only allowed to talk about points in the Euclidean plane. If we are willing to make a new kind of "inversive geometry" where we add a point at infinity to the Euclidean plane, we can formally define inversion of the center of the circle. For now, let us not change geometries in mid-chapter.

8.3 Simple Properties of Inversion

In this section, assume that all inversions will be relative to a circle called K with center O and radius r.

8.3.1 Inversion Expressed as a Ratio

We can express inversion as a ratio. Since the radius of the circle is r, what is the mathematical relationship between the lengths OP' and OP in figure 8.3? There are three similar triangles: $\triangle ORP \sim \triangle OP'R \sim \triangle PP'R$. Using the similarity of the first and last triangles, we obtain:

$$\frac{OP}{OR} = \frac{OR}{OP'}$$

or, as it is more commonly written:

$$OP \cdot OP' = OR \cdot OR = r^2. \tag{8.2}$$

In many books equation 8.2 is used as the definition of inversion. The construction we used to define it is derived as a result. Both definitions are equivalent.

This equation shows that the concept of the inverse of the center as a "point at infinity" is exactly equivalent to the concept that $1/0 = \infty$—as OP gets close to zero, OP' gets "close to" $r^2/0 = \infty$.

8.3.2 ♦ Inversion in the Complex Plane

There is a nice analytic definition of inversion if you know something about complex numbers (see Section A.13).

If o is an imaginary number somewhere on the complex plane, and r is a real radius, then the transformation that inverts all the complex numbers z to their complementary points z' through a circle of radius r centered at o in the complex plane is given by:

$$z' = \frac{r^2}{\bar{z} - \bar{o}} + o.$$

This clearly has the right properties—if z is very close to o (the center of inversion), then $\bar{z}$ will be very close to $\bar{o}$, so the denominator is very small and the result will be very large. Similarly, if we move the o to the other side and multiply both sides by $(\bar{z} - \bar{o})$ to get rid of the fraction, we obtain:

$$(z' - o)(\bar{z} - \bar{o}) = r^2$$

which looks a lot like the geometric definition, since the absolute value of $z' - o$ represents the distance between z' and o.

If inversion in the complex plane is relative to the origin (in other words, if $o = 0$), and if the radius of the circle of inversion is 1 (in other words, if inversion is relative to the unit circle), then the inverse of a (complex) point z is $1/\bar{z}$—very similar to inversion in the real line, since if z happens to be real, $\bar{z} = z$.

8.3.3 The Inverse of a Line

Inverting single points relative to a circle is not too interesting, but what if all the points on a line are inverted? What will be the resulting shape? Using **Geometer**, it is very easy to run experiments like this—take a point that is constrained to be on a line, and invert it through a circle. Then move the point along the line and see which points are swept out by its inverse.

A bit of experimentation should convince you that in almost every case, the inverse of a line seems to be a circle. Can you find the case where it is not? Let us prove the following theorem:

Theorem 8.1 (Inverse of a Line) *If K is a circle and L is a line that does not pass through the center of K, then the locus of the inverses of the points on L is a circle that passes through the center of K. If L lies completely outside of K, its inverse will be completely inside K. If L is tangent to K, its inverse will be a circle inside K, but tangent to K at the point where L is tangent. If L intersects K at two points, its inverse will be a circle intersecting K in those same two points. The inverse of L is not quite the entire circle—it is the circle missing the point that is the center of K. If L passes through the center of K, the points on the inverse of L lie on L itself (although the points of L that were inside K map to points outside, and vice-versa). Of course if L passes through the center of K, that center point can not be inverted, and the inverse of L will be L itself with the exception of the point at the center of the circle.*

It is easy to construct the circle that is the inverse of L. For the case where L does not intersect K, see figure 8.4.

Let O be the center of circle K, and from O, drop a perpendicular line to L, meeting L at U. Let U' be the inverse of U relative to the circle K. We will show that the circle with diameter OU' is the inverse of line L.

To prove this, let P be any point on L, and construct the inverse P' of P. If r is the radius of K, we have:

$$OP \cdot OP' = r^2 = OU \cdot OU'$$

so

$$\frac{OP}{OU'} = \frac{OU}{OP'}.$$

Figure 8.4: Invert a Line in a Circle (Case 1)

Inversion/Inversion1.T [M]

Since they share the angle at O, $\triangle OP'U' \sim \triangle OUP$, and since $\angle OUP = 90°$, we have $\angle OP'U' = 90°$. For every P, $\angle OP'U' = 90°$, so all possible points P' must lie on the circle with diameter OU'.

The second case, where L is tangent to K, is simple. Exactly the same diagram can be used, except that $U = U'$, so the argument works for every point except where $U = U'$, but in that case, the inverse of U, since it lies on the circle, is $U = U'$ itself, and U' is one end of the circle's diameter, so it is on the circle as well. The inverse of a line that is tangent to the circle of inversion is a circle passing through the center of the circle of inversion and tangent to the circle of inversion at the same point that the line is tangent. In other words, it is a circle exactly half the size of the circle of inversion, and it fits snugly between the center and edge of that circle.

In the third case, L passes through K in two points (see figure 8.5). The argument is virtually identical, except when we drop the perpendicular from O to L, the intersection, U, will be inside the circle. Thus the inverse, U' will be outside, and the image circle, which will still be the circle with diameter OU', will pass through the circle in two points which will coincide with the intersections of L and K.

In this situation, there are a few cases to consider. The points on the line that we are inverting can be either inside, outside, or on circle K. If they are on circle K, they are their own inverses, so points where L passes through K are fixed under inversion. If the point is inside (look at P, for example, in figure 8.5), it is easy to show by the inversion
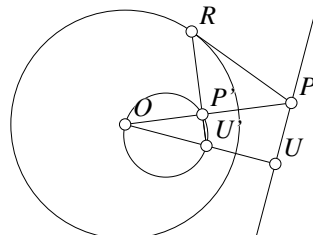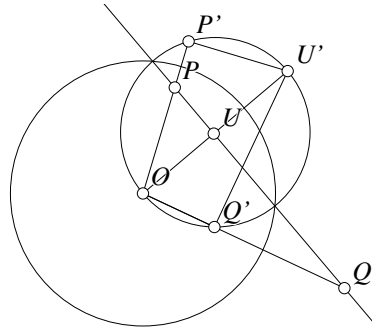
Figure 8.5: Invert a Line in a Circle (Case 3)

Inversion/Inversion2.T [M]

ratios that $\triangle OP'U' \sim \triangle OUP$, so $\angle OP'U' = 90°$, so P' lies on the circle with diameter OU'. If the point on L is outside circle K (see point Q, for example), then since $OU/OQ = OQ'/OU'$ and since they have the included angle $\angle QOU$ in common, we have $\triangle OUQ \sim \triangle OQ'U'$. $\angle OUQ = 90°$, so $\angle OQ'U' = 90°$, so Q' is on the circle with diameter OU' which is what we wanted to prove.

The final case where L passes through O is trivially true—the inverses of points of L are guaranteed to lie on L by definition.

8.3.4 Inversion of a Circle in a Circle

If P inverts to P' in circle K then P' inverts to P in that same circle. From this and the previous section, we know that the inversion of a circle that passes through the center of K of inversion is a straight line, with the same three cases—if the circle is completely within K, the line will be completely outside it. If the circle is tangent to K, the line will be tangent to K at the same point, and finally, if the circle passes through K in two points, the line will also pass through K in those same two points.

But what if the circle to be inverted does not pass through the center of K?

If you experiment with a **Geometer** diagram you can convince yourself that the inverse of a circle not passing through the center of K is another circle. There are again a few cases to consider but let us look at figure 8.6. In that figure, the circle centered at X is being inverted in K, the circle with center O. We will show that if we connect the centers O and X and invert the endpoints D and E of the diameter along the line OX to D' and E', then the inversion of the circle centered at X is the circle whose diameter is $D'E'$.

Clearly D and E go to the right places, but how about an arbitrary point P on the circle centered at X? If we poke around a bit we can find some similar triangles.

From the definition of an inversion we have:

$$OP' \cdot OP = OD' \cdot OD,$$

or $OP'/OD' = OD/OP$. Therefore $\triangle OPD \sim \triangle OD'P'$ by SAS since in addition to the ratio above, the two triangles have $\angle DOP$ in common.

Figure 8.6: Invert a Circle in a Circle
`Inversion/Inversion3.T` [M]

Using exactly the same argument, but using E and E' instead of D and D' we have $\triangle OP'E' \sim \triangle OEP$. By the similarity of these two triangles, we know that $\angle PDE = \angle D'P'O$ and $\angle PEO = \angle OP'E'$. But $\angle OP'E'$ is the supplement of $\angle E'P'P$ and $\angle PEO$ is the supplement of $\angle PED$, so $\angle PED = \angle E'P'P$.

Since they form a triangle,
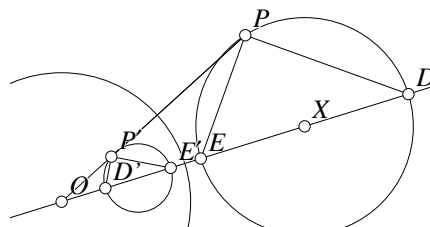
$$\angle PED + \angle PDE + \angle EPD = 180°. \tag{8.3}$$

And since they form a straight line,

$$\angle E'P'P + \angle D'P'O + \angle D'P'E' = 180°. \tag{8.4}$$

Since $\angle PED$ in equation 8.3 is equal to $\angle E'P'P$ in equation 8.4 and similarly for $\angle PDE$ and $\angle D'P'O$, we conclude that $\angle EPD = 90° = \angle D'P'E'$, so P' must lie on the circle with diameter $D'E'$ which is what we wanted to show.

Very similar arguments can be used if the circles intersect, or even if one circle completely surrounds the other.

Do not make the mistake of thinking that because a circle is mapped by inversion into another circle that the center is mapped to the center, or that all the diameters are mapped to other diameters. All you know is that every point on the circle is mapped to other points on the circle, and not necessarily with uniform spacing. Figure 8.7 shows this clearly— the circle on the left with uniformly-spaced points that appear as small circles is inverted in the center circle to make the circle with non-uniformly-spaced points

Figure 8.7: Non-Uniformity of Inversion
`Inversion/Inversion4.D [D]`

indicated as diamonds. The center of the original circle (also a small circle) is inverted to the diamond far to the left of the circle of diamonds. Inversion turns things inside-out, so the inverse of the center may even be outside the inverse of the original circle. The diameters along the line passing through the centers of the circle and its inverse are mapped into each other, but except in very special cases (for example if the circle is concentric with the circle of inversion), the other diameters will almost never be mapped to diameters.
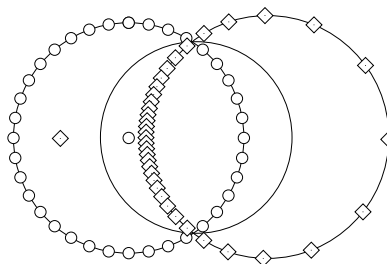
8.3.5 Inversion with Straightedge and Compass

All the results above are theoretical, and can be very useful in constructing a proof, but if you actually have to draw the inverse of a line or a circle, how do you do it?

If the result is a line, simply select two points on the line or circle that is to be inverted and invert them as described in Section 8.2. Then connect those two points with a line.

If the result is a circle, simply invert three different points, and you will be left with three points that must lie on the required circle. The intersection of the perpendicular bisectors of the segments connecting the points is the center of the required circle.

Of course if you are using **Geometer**, just use the built-in inversion commands: *VC=>V Inv*, *LC=>C Inv* and *CC=>C Inv*. Since both lines and circles usually go to circles, **Geometer** assumes that the result is always a circle; if the result is a line, what is drawn is effectively an infinitely large circle.

8.4   Preservation of Angles Under Inversion

When we say that inversion in a circle preserves angles, exactly what is meant? It is clear what is meant by the angle between two lines, but what is the definition of the angle between a line and a circle, or the angle between two circles?

The sensible way to define the angle between curves (where a straight line is a special type of "curve") is to say that the angle between them is the same as the angle between their tangent lines at the point of intersection. Thus if a pair of circles intersect at a point P and you want to measure the angle between them, construct the tangents to each of the circles at P and measure the angle between those tangent lines. If you have a circle and a line, just find the tangent to the circle at the point of intersection and measure its angle relative to the line.

8.4.1   Inversion Preserves Angles Between Lines

Suppose AB and CD are two lines in the plane that intersect at the point P. If they are inverted in a circle with center O, we will show that the resulting figures make the same angle with each other.

The easiest case is if O, the center of inversion, happens to be the same as P. Since both lines pass through the center of inversion, they are transformed into themselves so their images trivially have the same angle between them.

If the center of inversion is not on one of the lines (but it may be on the other), then the line that does not contain the center will be inverted to a circle passing through O. Suppose O is not on AB. Consider the line OX passing through O that is parallel to AB (see figure 8.8). Under inversion, OX is trans-

Figure 8.8: Invert Two Lines in a Circle
`Inversion/Angles1.D` [D]

formed into itself (since it passes through the center of inversion), and AB is inverted into a circle K passing through O. Then OX must be tangent to K, since it touches K at O, and if it intersected K in more than one place, OX and AB would intersect, which is impossible since they are parallel.

If O is on CD, then CD is inverted to itself, and its angle with OX is clearly unchanged (OX is also inverted into itself). But CD makes the same angle with OX as with AB since they are parallel, so the inverses of AB and CD make the same angle.

Finally, if O is on neither line, then both AB and CD are inverted to circles that meet at O, and the lines tangent to those circles at O (OX and OY in the figure) are parallel to the original lines, so they make the same angles as did the original lines.
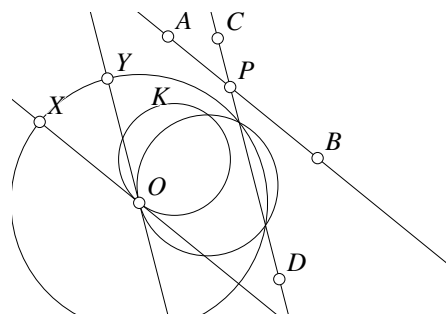
8.4.2 Inversion Preserves Angles Between Circles

If two intersecting circles are inverted, the angle between them is also preserved under inversion. The proof basically uses the result above, but there are a few cases to consider—whether the inversion is through the point of intersection, whether it is on one of the circles, et cetera. We will not present the exhaustive proof here, but the following paragraph gives the general idea.

Where the original circles intersect at a point P, construct lines tangent to them at P, and the angle between those tangent lines is the same as the angle between the circles. Under inversion, we know the inverses of the lines go to circles that intersect at the same angle, and these circles have to be tangent to the images of the circles we were inverting (since they meet the line in exactly one point). Tangent circles have the same tangent line at their point of tangency, so the inverses of the circles meet at the same angle as the inverses of the tangent lines, which meet at the same angle as the original circles.

8.4.3 Orthogonal Circles
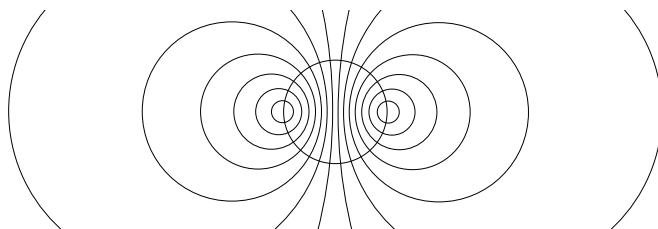
Figure 8.9: A Family of Orthogonal Circles
`Inversion/Orthocircs.T [S]`

An interesting situation to consider is when circles (or lines, or lines and circles) intersect at $90°$. In any of these cases, the two figures are called orthogonal. In figure 8.9 the circle in the center is surrounded by a set of circles that are orthogonal to it.

Since angles are preserved under inversion, what happens if a circle is inverted through another circle to which it is orthogonal? The answer is that the circle is inverted into itself. To see why, remember that the circle of inversion is unchanged by the inversion. The circle orthogonal to it is mapped to a circle that meets it at the same points, and remains orthogonal. Therefore it must be mapped to the same circle. (Note that this does not mean that the points are not moved by the inversion—the ones outside are mapped to points inside and vice-versa, but every point on that circle is inverted to another point on the same circle.)

8.5 Summary of Inversion in a Circle

Here are the key facts about inversion in a circle K centered at O having radius r. Inversion is indicated with primes—the inverse of point P is P'; the inverse of circle ω is ω', et cetera. We will be a bit sloppy in dealing with the inverse of O.

1. Points inside K are inverted to points outside K and vice-versa. Points lying on K invert to themselves. As an important special case, note that $K' = K$.

2. $OP \cdot OP' = r^2$. The points P and P' lie on the same ray originating at O.

3. Circles not passing through the center of K are inverted to other circles. Circles passing through the center are inverted to lines.

4. Lines passing through O are inverted into themselves. Lines not passing through O are inverted to circles passing through O.

5. Lines and circles that intersect in zero, one, or two places are inverted to other lines and circles that intersect in the same number of places. A little care must be taken to interpret this statement correctly if the intersection or tangency is at O. For example, if two circles are tangent at O, then their inverses will be parallel lines (that "meet at infinity"). If a line and a circle are tangent at O, then the inverse of the circle will be parallel to the line (which is inverted into itself).

6. Lines or circles intersecting or tangent to K are inverted to lines or circles intersecting or tangent to K in exactly the same places.

7. If two lines, two circles or a line and a circle meet at an angle α, then the inverses of those two figures will also meet at the same angle α.

8. If a circle K_1 is orthogonal to K, then $K_1' = K_1$.

In most of the examples that follow, we will choose the circle about which we would like to do the inversion, and we can choose it to get the effect you want. For example, suppose we are working with a circle, and we would much rather be working with a straight line. Then choose the circle of inversion to be one whose center lies on the circle you want to convert into a line. After inversion, the circle we wanted to avoid will be converted to a straight line.

If you want a straight line to remain a straight line after the inversion, then you had better invert through a circle whose center lies somewhere on that line. If you want two straight lines to be unchanged after the inversion, then the circle of inversion must have its center at the intersection of those two lines.

If you have three lines that do not meet in a point, and you want them all to be straight lines after an inversion, then, unfortunately, you are out of luck, and you will have to find a different way to solve the problem.

Remember that **Geometer**'s inversion commands are particularly powerful. If you have a problem that you think can be approached with inversion, draw the figure, add

two arbitrary points, make a circle with one as the center and the other on the edge of the circle, and invert all the objects of interest in that arbitrary circle. Then you can move the test circle around to see how inversion relative to circles having various centers and various radii modifies the problem. Try moving the center over all the "interesting" points in the **Geometer** diagram to see what happens.

8.6 Circle through a Point Tangent to Two Circles

Given two circles and a point, find the circle (or circles) that are tangent to both the given circles and pass through the point[4].
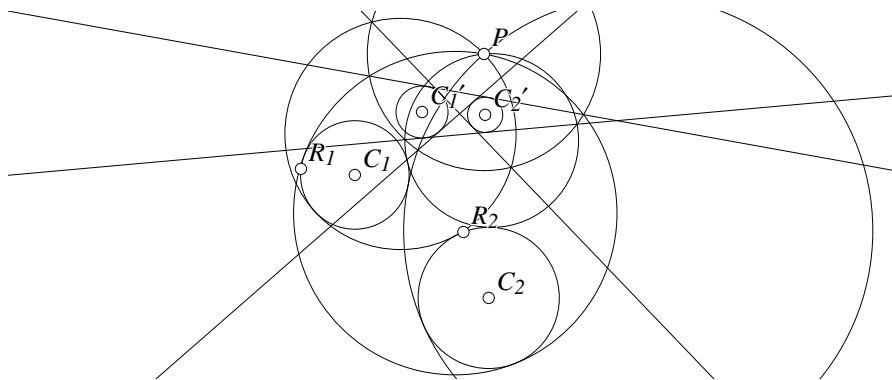
Figure 8.10: Circle through a Point and Tangent to Two Circles
Inversion/Pointcirccirc.T [P]

See figure 8.10. The problem of finding a circle tangent to two other circles seems difficult, to say nothing of the added problem of making it go through a particular point. It is, however, not hard to construct lines tangent to one or two circles. The problem would be much easier if we were searching for a line instead of a circle.

The final solution will be a circle passing through P and tangent to the two circles centered at C_1 and C_2. If we could do an inversion to turn that final solution into a line, the problem might be easier.

If a circle passes through the center of its circle of inversion, it is converted to a line, so why not invert the entire figure through a circle centered at P since the solution has to go through P?

But which circle? There are an infinite number of choices with any possible radius. The answer (in this case) is that it does not matter! Pick *any* circle centered at P and invert everything through it. After inversion the circles centered at C_1 and C_2 become the circles centered at C_1' and C_2'. The solution circle, wherever it was, is inverted to a

---

[4]This is a simple example of what are generally known as Apollonius' Problems—to construct circles satisfying various tangency and concurrency requirements. The most difficult of Apollonius' Problems is this: Given three circles in a plane, construct the circle (or circles) simultaneously tangent to all three of them.

line, but it will still be a line tangent to the inverted copies (C_1' and C_2') of the original circles[5].

We need to find a line tangent to the circles centered at C_1' and C_2', but depending on the relationships and radii of the two original circles, there may be as many as four solutions or as few as zero. The solution lines will be the common external and internal tangents to those circles. If the circles intersect, there are no common internal tangents. If one of the inverted circles is inside the other, there will be no external tangents either.

Invert those lines (however many there are) back through the original circle of inversion to obtain all the possible solutions to the original problem. In the figure there are four common tangents to the inverted circles, so there will be four solutions to the problem, and all are displayed in the figure.

This is an example where a computer geometry diagram is vastly more powerful than a pencil and paper. There are dozens of configurations to try, and it would be very valuable for you to try them yourself. Adjust the positions and radii of the circles centered at C_1 and C_2, and move the point P around to get a feeling for what the solutions look like, and when there are zero, one, two, three, or four solutions (all five situations can occur).

8.7 Circle Tangent to Three Circles
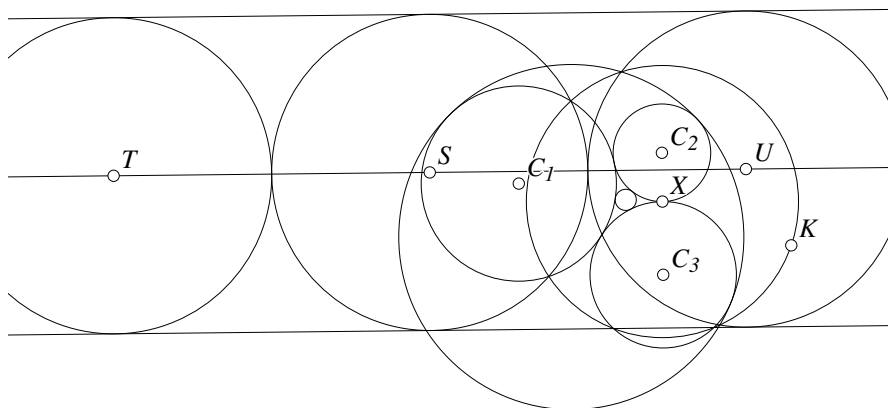
Figure 8.11: Circle Tangent to Three Circles
`Inversion/Threetancirc.T` [P]

---

[5]If you are doing this with a straightedge and compass, it will be slightly easier to choose the circle of inversion so that in addition to having its center at the point, it is also tangent to one of the circles. Do you see why?

This is a very special case of Apollonius' Problem[6]. Suppose that you are given three mutually tangent circles centered at C_1, C_2, and C_3 as in figure 8.11. Construct the circles that are simultaneously tangent to all three of the mutually tangent circles.

In general there will be two solutions—one outside the three circles and one in the space enclosed by them. Since it is hard to find tangents to circles, it would be nice to turn them into lines. We cannot turn all three into lines, but certainly two of them can be by inverting in a circle that is centered at the point of tangency of two of them.

Let X be the point of tangency between the circles centered at C_2 and C_3. Draw an arbitrary circle with X as its center (in the figure it is the circle passing through the point K), and invert all three of the original circles in it. The two circles tangent at its center will be mapped to parallel lines, and those lines will continue to be tangent to the inverted image of the circle centered at C_1 which in the figure is the circle centered at S. The two tangent lines are the roughly horizontal lines touching the circle centered at S above and below. It is easy to find circles centered at T and U that are tangent to the one centered at S and which are also tangent to the two lines. There is one on each side, they have the same radius as the circle centered at S, and their midpoints lie on the line through S and parallel to the lines that are inverses of the other two circles.

Invert the two circles back through the one centered at X and this will yield the two solutions—a tiny one inside and a larger one outside.

There is a purely classical solution to this problem, where "classical" means that it does not use inversion. The complete classical construction can be found in the **Geometer** file `Inversion/ClassicSoddy.T`.

8.8 Ptolemy's Theorem Revisited

Zvezdelina Stankova sparked this author's interest in inversion by showing him this most wonderful proof in the world of Ptolemy's theorem using inversion.

Ptolemy's theorem is the same— given any convex cyclic quadrilateral $ABCD$, show that

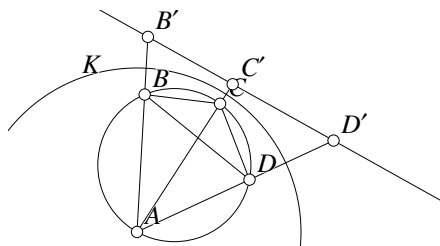$$AB \cdot CD + BC \cdot AD = AC \cdot BD.$$

Figure 8.12: Ptolemy's Theorem Revisited
`Inversion/Ptolemy1.T` [M]

The theorem will be proven by inverting the figure about a circle K with center at A, as shown in figure 8.12.

First we need a simple lemma that is quite useful in its own right.

---

[6]The difficult case of Apollonius' problem solves this for three arbitrary circles that are not necessarily mutually tangent. Also, see the footnote to the poem at the beginning of this chapter.

Lemma 8.1 (Length of an inverted segment) *If the endpoints of segment BC are inverted in a circle with center A, then the length of segment $B'C'$ is related to the length of BC as follows:*

$$B'C' = \frac{BC \cdot AB'}{AC} = \frac{BC \cdot AC'}{AB}.$$

To prove this we can use the same figure 8.12. $AB' \cdot AB = AC' \cdot AC$ because they are inverse points, so $AB'/AC' = AC/AB$. Since $\angle BAC$ is equal to itself, by SAS we have $\triangle ABC \sim \triangle AC'B'$ so

$$\frac{B'C'}{BC} = \frac{AB'}{AC} = \frac{AC'}{AB}.$$

Multiply through by BC to prove the lemma.

If the entire quadrilateral $ABCD$ is inverted through a circle centered at A with an arbitrary radius r, we have:

$$AB \cdot AB' = AC \cdot AC' = AD \cdot AD' = r^2.$$

Since the circle in which $ABCD$ is inscribed passes through the center A of the circle of inversion, its inverse is a line, so B', C', and D' lie on a line, and from the figure, it is clear that $B'D' = B'C' + C'D'$.

Replacing $B'D'$, $B'C'$, and $C'D'$ by their equivalents according to Lemma 8.1, we have:

$$\frac{BD \cdot r^2}{AD \cdot AB} = \frac{BC \cdot r^2}{AC \cdot AB} + \frac{CD \cdot r^2}{AC \cdot AD}.$$

Cancel the r^2 from all terms and multiply through by $AB \cdot AC \cdot AD$ and obtain:

$$AC \cdot BD = BC \cdot AD + CD \cdot AB,$$

which completes the proof of Ptolemy's theorem.

8.9 Fermat's Problem

Ptolemy's theorem has an interesting extension that can be used to solve Fermat's Problem:

Fermat's Problem: Let A, B, and C be any three points in the plane. Find the point X that minimizes the sum of the lengths of the segments connecting A, B, and C to X. In other words, make $AX + BX + CX$ as small as possible.

Ptolemy's theorem tells us that if $ABCD$ is a cyclic quadrilateral, then $AB \cdot CD + BC \cdot DA = AC \cdot BD$, but what if the four points A, B, C, and D do not happen to lie on a circle? is there some inequality that can replace the "=" sign?

The answer is yes. To see why, look at a slight enhancement of Ptolemy's theorem:

Theorem 8.2 (Ptolemy's Theorem Extended) *If A, B, C, and D are any four points in the plane, then*

$$AB \cdot CD + BC \cdot DA \geq AC \cdot BD,$$

and equality holds only if $ABCD$ is a convex cyclic quadrilateral.

Suppose that $ABCD$ is not a cyclic quadrilateral as in figure 8.13. Then we can go through exactly the same steps as we did in the proof of Ptolemy's theorem in the previous section, except it is no longer the case that $B'C' + C'D' = B'D'$—that only holds if B', C' and D' lie on a line, and that happens only for cyclic quadrilaterals. If the three points do not lie on a straight line, the triangle inequality gives us: $B'C' + C'D' \geq B'D'$. Since other than this change of "=" to "$\geq$", the proofs are identical, we have proved the extension to Ptolemy's theorem.
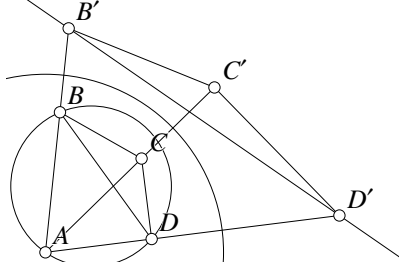
Figure 8.13: Ptolemy's Theorem Extended
`Inversion/Ptolemyext.T [M]`

Now we can look at Fermat's Problem, which is solved by other methods in Section 7.4.

In figure 8.14, let A, B, and C be the points in question, and on side AB construct an equilateral triangle $\triangle ABD$ pointing outward[7]. We claim that the intersection X of the line DC and the circle that circumscribes $\triangle ABD$ is the required point.

Let X be an arbitrary point. Then the extension to Ptolemy's theorem tells us that

$$AX \cdot DB + BX \cdot AD \geq DX \cdot AB,$$

but since $\triangle ABC$ is equilateral, we know that $AD = DB = AB$, so we know that

$$AX + BX \geq DX. \tag{8.5}$$

Add CX to both sides of equation 8.5 (since we are trying to minimize $AX + BX + CX$) to obtain:

$$AX + BX + CX \geq DX + CX. \tag{8.6}$$

If X happens not to be on the line DC, then $DX + CX$ can always be made smaller by moving X on to the line DC by the triangle inequality. Thus to minimize $DC + CX$, X must lie on the line DC between D and C.

---

[7]This choice of a construction may seem as if it comes out of the blue, but in fact, it is similar to the way we solved Fermat's Problem the first time.
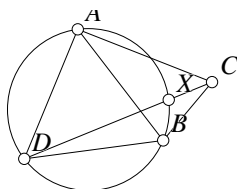
Figure 8.14: Fermat's Problem
`Inversion/Ptolemyfermat.D [D]`

But if X is on DC, then the sum $CX + DX$ on the right of equation 8.6 is constant, and the Ptolemy's theorem extension tells us that the sum on the left can be minimized by choosing X to lie on the circle. Thus we have solved Fermat's Problem.

8.10 Inversion to Concentric Circles

Given two circles in the plane that do not intersect, find an inversion that maps them into a pair of concentric circles.

Concentric circles are perpendicular to their diameters, so it seems clear that the center of inversion we seek should lie on the line connecting the centers of the two original circles. That way, after the inversion, at least both the edges of the images of those circles will remain perpendicular to that line. But we would also like to find a circle center such that after the inversion a line perpendicular to the line connecting the centers will also be orthogonal to both the images of the circles.
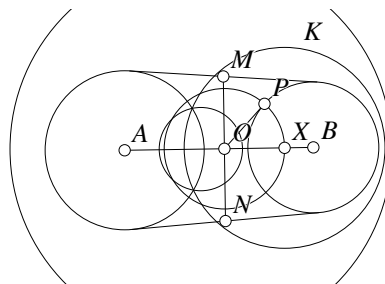
Figure 8.15: Inversion to Concentric Circles
`Inversion/Concentric.T [M]`

From any point on the radical axis (see Section 5.7) we can find a circle centered there that is also perpendicular to both circles. Since the goal is symmetry, choose the circle whose center is on the intersection of the radical axis and the line connecting the diameters of the two circles.

Figure 8.15 makes this clear. The original circles are centered at A and B. To find the radical axis, connect the midpoints M and N of their common external tangents. But MN intersects AB at O. Now find the circle centered at O that is orthogonal to both circles. (To do this, find the point P where a line from O would be tangent to one of the circles, and draw the circle centered at O and passing through P. Note that since O is on the radical axis, the lengths of the tangent segments to both circles will be the same, so a circle of that radius will be perpendicular to both.)

Let this circle intersect the line AB at X, and all we need to do is invert both circles in an arbitrary circle K centered at X. Since AB passes through X and AB is orthogonal to both circles, the line AB will invert to itself and hence will remain perpendicular to the images of both circles. Since the circle centered at O passes through X, under inversion it will go to a line. That line will be perpendicular to AB, since the circle that inverted to it was.

But that line is the image of a circle that was also perpendicular to both the original circles, so the line will be perpendicular to their images. Thus, after inversion, we know of two perpendicular lines that are both perpendicular to the images of both circles, and the only way this can occur is if the two circles are concentric.

8.11 The Steiner Porism

We now have all we need to prove a truly amazing Euclidean geometry theorem.

Theorem 8.3 (Steiner Porism) *Suppose we are given two circles K_0 and K_1 where one is completely contained within the other, and which do not touch. Construct an arbitrary circle between the two that is tangent to both, and call it C_0. Now construct a circle C_1 that is tangent to K_0, K_1, and C_0. Next, continuing away from C_0 construct a circle C_2 tangent to K_0, K_1, and C_1. Continue in this way as shown in figure 8.16.*

It may happen that for some n, C_n is exactly tangent to C_0 as occurs in the figure. If that is true, then it does not matter where you

Figure 8.16: The Steiner Porism
`Inversion/Steiner.T [S]`

chose the initial circle C_0. The sequence of circles will have C_n tangent to C_0 independent of the starting position. Furthermore, if for no n is C_n tangent to C_0 then modifying the starting point will not help—such a tangency will never occur[8].

Based on the previous problem, this one is fairly simple. Find an inversion that takes circles K_0 and K_1 into concentric circles. Then it is clear that the starting position does not matter—all the C_i will be exactly the same size, and they either meet with tangency or they do not. If they do, then they can be inverted back through the circle of inversion to place the image of C_0 wherever is desired. If they do not, then no such inversion will work.

8.11.1 Drawing the Steiner Porism
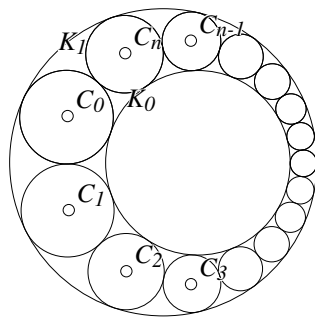
How can figure 8.16 be constructed using **Geometer**?

---

[8]It may even be the case the tangent circles go around the loop more than one time before the nth one is tangent to C_0. The theorem still holds, but from any starting point, a different set of tangent circles will make the same number of loops before the nth one is tangent to the initial one.

Obviously, a set of equally-spaced circles that exactly fill the ring between two concentric circles was constructed, and the result was inverted to obtain a Steiner porism with a lopsided pair of enclosing rings. The inverted circles will exactly fill the space between the lopsided rings.

It is not hard to work out the relative sizes of a pair of concentric circles that will allow for some fixed number n of circles to fit between them. In figure 8.17 we see what we need to begin. In this case, there are nine circles, but let us just call that number n. The central angle between any pair of circle centers is $360°/n$, so angle $\angle AOB = 360°/n$ is a suitable vertex angle of the isosceles triangle in the figure.
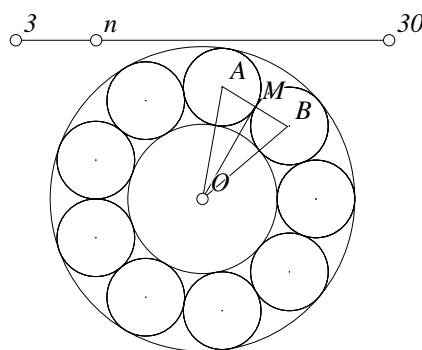
Figure 8.17: A Ball-Bearing Race
`Inversion/Steiner1.T` [M]

If that figure, $\triangle OMA$ is a right triangle where M is the point of tangency of two adjacent circles. Suppose the radius of the inner circle is R_1 and the radius of the small surrounding circles is R_2 (which will make the radius of the larger circle $R_1 + 2R_2$).

Then
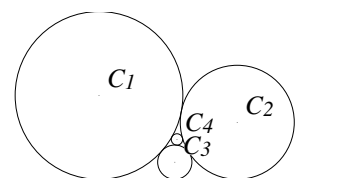
Figure 8.18: Four Circle Problem
`Inversion/Fourcircle.D` [D]

$$\sin\left(\frac{180°}{n}\right) = \frac{R_2}{(R_1 + R_2)},$$

and we can solve for R_2:

$$R_2 = \frac{R_1 \sin(\frac{180°}{n})}{(1 - \sin(\frac{180°}{n}))}.$$

Thus, if you can draw something like figure 8.17, it can be inverted through some circle to produce the required general figure.

The only tricky part of the construction was the method by which the number of circles in the loop could be altered by the **Geometer** program. As the point that controls the number of inner circles is moved on the line, a ratio is calculated of its position relative to the length of the entire segment. That value is truncated to obtain the number of circles to draw. Then the largest number of circles is drawn, each offset from the next by the same angle. If the actual number is smaller than the largest possible number, extra circles are simply drawn on top of earlier ones so only the smaller number appear to be drawn.

8.12   Four Circle Problem

See figure 8.18. Given a line, construct a circle tangent to it centered at C_1 with radius r_1. Next, construct a circle tangent to both the line and the circle centered at C_1. The new circle has center C_2 and radius r_2. As shown in the figure, construct the circle centered at C_3 of radius r_3 tangent to both circles and to the line, as shown. Finally, the circle centered at C_4 of radius $r_4 = 1$ cm is tangent to the first three circles and lies inside as in the figure. Find the perpendicular distance from C_4 to the line in terms of r_1, r_2, r_3, and $r_4 = 1$ cm.
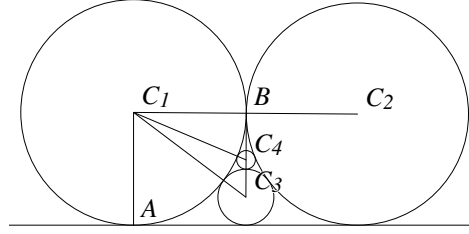
Figure 8.19: The Four Circle Problem
`Inversion/Fourcircle1.D [D]`

Although at first it does not look like it, this is just a special case of the Steiner porism, but in this case, the outer circle is the "circle" with an infinitely large radius—the straight line. Imagine what would happen if this figure were inverted in a random circle that was not special in any way (none of the lines or circles pass though its center). It would become the simplest example of a Steiner porism with an inner and outer circle and three circles filling the ring between them.

If that is the case, it does not matter what r_1, r_2, and r_3 are—the height above the line will only depend on the radius $r_4 = 1$ cm.

So we might as well choose a set of circles with convenient radii as in figure 8.19, where $r_1 = r_2$.

Using the pythagorean theorem on the right triangle $\triangle C_1 B C_3$ we have:

$$
\begin{aligned}
r_1^2 + (r_1 - r_3)^2 &= (r_1 + r_3)^2 \\
r_1^2 + r_1^2 - 2r_1 r_3 + r_3^2 &= r_1^2 + 2r_1 r_3 + r_3^2 \\
r_1^2 - 4r_1 r_3 &= 0 \\
r_1(r_1 - 4r_3) &= 0 \\
r_1 &= 4r_3.
\end{aligned}
$$

Now use the pythagorean theorem again, but this time on $\triangle C_1 B C_4$:

$$
\begin{aligned}
r_1^2 + (r_1 - 2r_3 - r_4)^2 &= (r_1 + r_4)^2 \\
r_1^2 + (r_1 - r_1/2 - r_4)^2 &= (r_1 + r_4)^2 \\
r_1^2 + (r_1/2 - r_4)^2 &= (r_1 + r_4)^2 \\
r_1^2 + r_1^2/4 - r_1 r_4 + r_4^2 &= r_1^2 + 2r_1 r_4 + r_4^2 \\
r_1^2/4 - 3r_1 r_4 &= 0 \\
r_1(r_1 - 12r_4) &= 0 \\
r_1 &= 12r_4.
\end{aligned}
$$

But $r_4 = 1$ cm, so $r_1 = 12$ cm and $r_3 = 3$ cm. The height of C_4 above the line is 7 cm.

The solution above seems pretty good, but there is an even easier way. See figure 8.20. We can invert to a situation where two of the circles become straight lines, and the calculations become even easier. If the radius of the circle centered at C_4 is 1 and the unknown equal radii of the other two circles are both r, we can see that r satisfies:

$$r^2 + (r-1)^2 = (r+1)^2.$$

Figure 8.20: The Four Circle Problem
`Inversion/Fourcircle2.D [D]`

The solution is $r = 4$ and from the figure it is clear that the point C_4 is 7 cm above the lower line.

Note that this problem could also have been solved using the Descartes circle theorem presented in the form of a poem at the beginning of this chapter. In this case, one of the "circles"—the straight line—has infinite radius, so 1 divided by that radius is zero.

8.13 The Arbelos of Pappus

"Arbelos" is the Greek word for a shoemaker's knife. In figure 8.21, ignore everything except for the three circles with diameters OA, OB, and AB, where O, A, and B lie on the same line, and notice that the area inside the larger circle and outside the two smaller circles is divided into two pieces on the left and right. Either of these shapes, which are basically a half-circle with two half-circles removed, look something like an arbelos.

From those original three circles, construct a series of circles on both sides of the circle with diameter AB as shown in the figure. Each circle is tangent to the circles with diameters OA and OB and also tangent to the previous circle in the series. In the figure, the centers of those circles on the right are labeled C_1, C_2, et cetera, and those on the left have centers C_{-1}, C_{-2}, ...
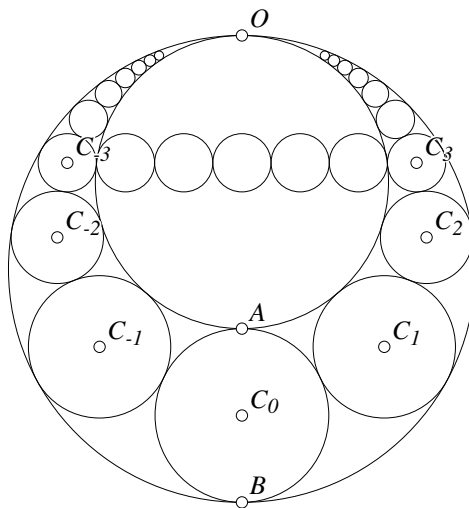

Figure 8.21: The Arbelos of Pappus
`Inversion/Arbelos.T [M]`

We will show that the distance between C_{-1} and C_1 is twice the diameter of the circle centered at C_1, that the distance between C_{-2} and C_2 is four times the diameter of the circle centered at C_2, and in general, that the distance between C_{-n} and C_n is $2n$ times the diameter of the circle centered at C_n. In the figure, this is illustrated for C_{-3} and C_3 and the circles centered there—exactly 5 circles of the same diameter as those circles can be placed on a straight line between them. There would be 1 between the circles at C_{-1} and C_1, 3 of them between the circles at C_{-2} and C_2, et cetera.

The proof is not difficult, and since we have been looking at inversion and the Steiner porism, it is clear that the situation here is very similar. If, for example, we can find an inversion that leaves the circles centered at C_{-3} and C_3 fixed and at the same time maps the circles with diameters OA and OB into parallel lines, we will be done. The circles between those centered at C_{-3} and C_3, namely those centered at C_{-2}, C_{-1}, C_0, C_1, and C_2, under inversion will remain tangent to their neighboring circles, and to the two parallel lines. Clearly, when we look at C_{-4} and C_4, there will be two more circles in the chain between them, so they will have two more circles on the line between them.
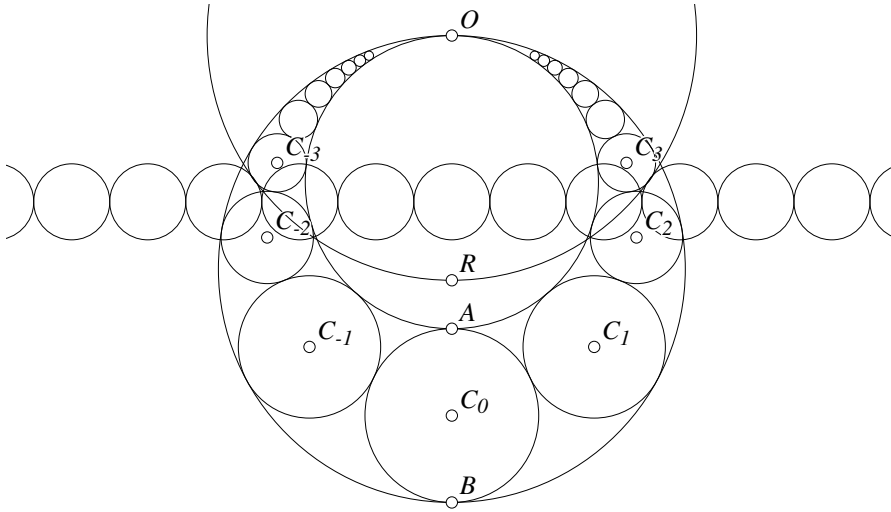
Figure 8.22: Proof of the Arbelos Property
`Inversion/Arbelos1.T` [P]

It is not hard to find such an inversion. Clearly, it will have to send the circles with diameters OA and OB to parallel lines, so the circle of inversion must be centered at O. Figure 8.22 shows the inversion of the circles C_i in such a circle centered at O and passing through a point R.

Since the circles C_{-i} and C_i are symmetric relative to the line OB, as the radius of the circle of inversion centered at O increases, it will expand through each pair of opposite circles in exactly the same way. At some point as it expands through each pair, it will be orthogonal to both and at that point, both of those circles will be inverted into themselves.

This result was known to Pappus sixteen centuries ago, and he certainly did not have the method of inversion available to him. To learn how Pappus might have done it, take a look at "How Did Pappus Do It?", a chapter by Leon Bankoff in [Klarner, 1981].

8.14 Another Arbelos Result

Beginning with the same arbelos (see figure 8.23), construct a perpendicular to the diameter AB at C, and then construct the two circles (shown in the figure with centers at D and E) tangent to that line and to the other pairs of circles. Show that the two circles with centers at D and E have the same radius. They are called **Archimedes' circles**.

A little bit of fiddling around with a computer geometry program leads to an unusual proof. If we draw the common tangent to the circles centered at E and O, we notice that it always seems to pass through the point A. We do not know if that is true, of course, but if it is, we can just write down equations and solve for the radius of the circle centered at E as follows:

Let R_1 be the radius of the circle with diameter CB, R_2 the radius of the circle with diameter AC, and r_1 is the radius of the circle centered at E.

Figure 8.23: Another Arbelos Property

`Inversion/Arbelos2.T [M]`

If the common tangent through T does pass through A, since EO is perpendicular to it we have $\triangle XEO \sim \triangle TAO$ by AA, since there is a common right angle and shared angle $\angle XOE$, so:

$$
\begin{aligned}
\frac{XO}{OT} &= \frac{EO}{AO} \\
\frac{R_1 - r_1}{R_1} &= \frac{R_1 + r_1}{2R_2 + R_1} \\
(2R_2 + R_1)(R_1 - r_1) &= R_1(R_1 + r_1) \\
2R_1R_2 - r_1R_1 - 2r_1R_2 &= r_1R_1 \\
2r_1(R_1 + R_2) &= 2R_1R_2 \\
r_1 &= \frac{R_1R_2}{R_1 + R_2}
\end{aligned}
$$

This is a nice result, since it shows that the radius is a symmetric function of R_1 and R_2. Exactly the same calculations would give the radius of the circle centered at D to be the same thing, since we just need to exchange R_1 and R_2 which leaves the final result unchanged.

We still have not proved anything since we do not know that the tangent at T passes through A. However, we will have a proof if we can show that the distance from E
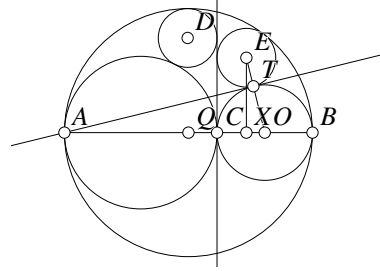
to the large circle is also equal to r_1. To do that, we just need to show that $QE = R_1 + R_2 - r_1$. (The radius of the large circle is $R_1 + R_2$, and if the circle centered at E is tangent to it, then the radius of the large circle from Q through E had better cross E a distance r_1 from the outer circle.) Using the pythagorean theorem twice, we have:

$$EX^2 + XO^2 = EO^2 \tag{8.7}$$
$$EX^2 + QX^2 = QE^2 \tag{8.8}$$

We know that $EO = R_1 + r_1$, $XO = R_1 - r_1$, and $QX = R_2 - R_1 + r_1$. If we just plug these values into equations 8.7 and 8.8 we obtain:

$$EX^2 = (R_1 + r_1)^2 - (R_1 - r_1)^2 = 4R_1 r_1$$

so

$$QE^2 = 4R_1 r_1 + (R_2 - R_1 + r_1)^2$$

We are done if $QE = (R_1 + R_2 - r_1)$ which is true if:

$$(R_1 + R_2 - r_1)^2 = 4R_1 r_1 + (R_2 - R_1 + r_1)^2,$$

which we can check:

$$(R_2 + (R_1 - r_1))^2 - (R_2 - (R_1 - r_1))^2 = 4R_1 r_1$$
$$4R_2(R_1 - r_1) = 4R_1 r_1$$
$$r_1 = \frac{R_1 R_2}{R_1 + R_2},$$

which is true.

8.15 The Problems of Apollonius

Given any combination of three figures chosen from among points, lines, and circles, find all the circles that pass through the selected points and are tangent to the selected lines and circles. There are ten possibilities which we discussed in Section 3.6.5, labeled here as follows: (PPP), (PPL), (PPC), (PLL), (PLC), (PCC), (LLL), (LLC), (LCC) and (CCC). The (PLL) problem, for example requires the construction of all circles passing through a given point and tangent to two given lines.

The (PPP) and (LLL) problems are solved in every high school geometry course. In Sections 3.6.5 and 3.6.6, we solved (PPL) and (PLL), and in Section 8.6 we solved (PCC). In Section 8.7 we even solved a special case (CCC).

Although all ten problems can be solved with classical techniques, the technique of inversion simplifies many of the problems. In this section we will solve (and in some cases, re-solve) various Apollonius' problems using that technique.

In many cases, we can consider lines and circles to be basically the same, from the point of view of inversion, so if you can solve, say, the (PCC) problem, you can probably also solve the (PLL) and (PLC) problems, although you must be careful.

Here is a short discussion of all the cases. The one that is commonly known by itself as "Apollonius' Problem" is the hardest one: (CCC). It has approximately the same difficulty as (LLC), or (LCC). In fact, if you have a rock-solid solution to the three circles problem, you could solve the other two as follows:

First, do an inversion around an arbitrary point not on any of the lines or circles, and convert your problem to (CCC). Solve that problem with your rock-solid method, and then that solution can be re-inverted through the circle around the arbitrary point that you initially selected. Notice that since your "rock-solid" method may itself involve inversion and re-inversion, you may have to do four total inversions to solve the problem.

In what follows, we may ignore the degenerate cases, such as when the points lie on the lines or circles, when the points are coincident, or when the three lines or circles meet at a point. These cases are not hard, and you should be able to figure them out without much difficulty.

8.15.1 Apollonius: (PPP) and (LLL)

The (PPP) and (LLL) problems are equivalent to finding the circumcircle and incenter, respectively, of a triangle.

The (PPP) problem can be solved using inversion—invert in a circle centered at one of the points, sending that point to infinity. If the line connecting the inversions of the other two points is re-inverted in that circle, it will be the required circle. Do you see why?

8.15.2 Apollonius: (PPL)

Do an inversion centered at one of the points. This will send that point to infinity, and will convert the line into a circle. The inverted solutions must pass through the point at infinity, and hence must be lines. They must pass through the inversion of the other point and be tangent to the circle, so if we can find the tangents from the inverted point to the circle, and re-invert that, we will have the solutions. There may be none if the inverted point lies inside the inverted line (which is a circle).

It is not hard to consider the degenerate case where one of the points falls on a line. Invert through the other point, and now there will be only a single solution, since the inverted point will lie on the circle that is the inversion of the line, and there is hence only a single tangent line.

8.15.3 Apollonius: (PPC)

This is exactly the same as the previous problem, except that the circle will be inverted to become a circle instead of the line becoming a circle.

8.15.4 Apollonius: (PLL), (PLC) and (PCC)

The (PLL) problem was solved in Section 8.6, but it, and the other two, can be solved via inversion. Send the point to infinity, and the lines and/or circles will become circles. Find the common exterior tangents, re-invert them, and obtain the solutions. The special cases where a line or circle passes though the point are also easily solved.

8.15.5 ◆◆ Apollonius: (LLC), (LCC) and (CCC)

All three of these problems are equivalent in the sense that a single inversion can convert all of them to (CCC). If the problem initially has any lines, do an inversion relative to a circle K whose center does not lie on any of the lines or circles. All the lines and circles will be inverted to circles. We will demonstrate below how to solve the three circles problem, and once you have done that for the inverted figures, re-invert them in the circle K to obtain your final solution.

There are up to eight possible solutions, and figure 8.24 illustrates a situation where all eight circles tangent to the three given circles exist.

The key idea is this: If the circle of smallest radius r is shrunk it to a point, and at the same time r is either added or subtracted from the radii of the larger circles, then solving Apollonius' problem for a point and two circles will yield a circle whose radius can be increased or decreased by r to yield a solution[9].

Figure 8.24: The Problem of Apollonius

`Inversion/Apollonius.T [M]`

Figure 8.25 demonstrates the general idea. The original circles for which the problem is to be solved are centered at C_1, C_2, and C_3, and they have radii C_1R_1, C_2R_2, and C_3R_3, respectively. Assume that C_1R_1 is the smallest of the three radii. One solution can be obtained by drawing a circle centered at C_2 of radius $C_2R_2 - C_1R_1$ and by drawing a circle about C_3 of radius $C_3R_3 + C_1R_1$. Using techniques we learned earlier in the chapter, we can find a circle that is tangent to those new circles with the modified radii and passing through C_1 as shown in the figure. If the radius of that circle is then increased by C_1R_1, we have one of the eight possible solutions.

This is trickier than it sounds. Remember that there are up to four solutions to the (PCC) problem, possibly having tangencies on both sides of both circles. Only one of these four circles can have its radius increased and still be tangent to the original circles. In fact, if you play with the **Geometer** diagram that generates figure 8.24, you will
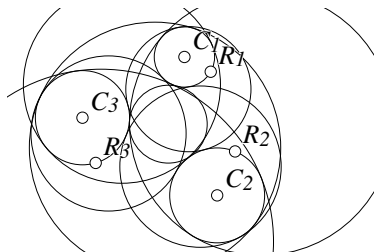
---

[9]There are direct solutions for (LLC) and (LCC) as well that involve moving the line or lines parallel to themselves by a distance equal to the diameter of the smallest circle, just as we expand and shrink the diameters of the larger circles in the three-circle solution.

find that it works only for a limited range of values around the initial configuration—increase or decrease the radii too much and you will find that the solutions jump to the other sides of circles, and on expansion or contraction of those circles, they are no longer solutions to the problem.
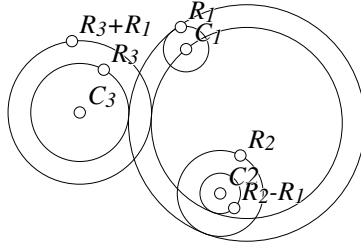
A different solution based on poles and polars appears in Section 10.6.

Figure 8.25: Addition and Subtraction of Radii

Inversion/Apollonius1.T [M]

8.16 Peaucellier's Linkage

For many years, it was unknown whether it was possible to construct a mechanical linkage that would turn perfect circular motion into perfect linear motion. Peaucellier's Linkage (see figure 8.26) shows a linkage that achieves this conversion.

In the linkage, point O is fixed on the circle, point A is constrained to move on the circle, and segments OC and OD are two bars of the length l, while segments AC, CA', $A'D$, and DA are four bars of length r. The bars are all hooked together with flexible joints at points O, A, C, A', and D. (The lines OA' and CD in the figure are solely for the proof—they are not part of the linkage.)
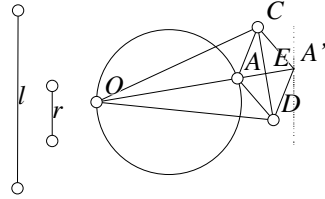
Figure 8.26: Peaucellier's Linkage

Inversion/Peaucellier.T [P]

We can show that the point A' will lie on a straight line if we can show that $OA \cdot OA'$ is constant. If that is the case, then A and A' are inverse points with respect to a circle centered at O. As the point A moves on a circle that passes through O, its inverse, A' must move along the inverse of that circle, which is a straight line since O lies on the circle upon which A is constrained to lie. (If O is not on the circle, A and A' will still be inverse points relative to a circle centered at O, but A' will merely move on a different circle as A traces out the first one.

To show this, construct the lines OA' and CD. Since $ACA'D$ is a rhombus, $CE \perp OA'$ and E bisects AA'. Thus we have

$$OA \cdot OA' = (OE - AE) \cdot (OE + EA') = (OE - AE) \cdot (OE + AE) = OE^2 - AE^2.$$

Using the pythagorean theorem on $\triangle AEC$ and on $\triangle OEC$, we have

$$
\begin{aligned}
OE^2 + EC^2 &= OC^2 = l^2 \\
AE^2 + EC^2 &= AC^2 = r^2.
\end{aligned}
$$

Subtracting, we obtain:

$$OE^2 - AE^2 = l^2 - r^2 = OA \cdot OA'.$$

Since $l^2 - r^2$ is constant, so is $OA \cdot OA'$ and we are done.
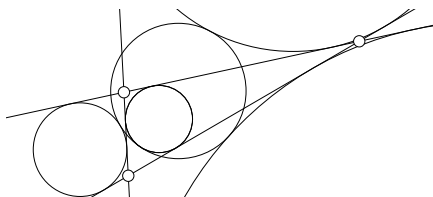
8.17 ♦♦♦ Feuerbach's Theorem

Figure 8.27: Feuerbach's Theorem
Inversion/Feuer.T [M]

Feuerbach's theorem has a fairly difficult proof, both because there are a lot of steps, and because a couple of those steps depend on "big" theorems themselves.

The theorem is not difficult to understand; figure 8.27 shows the nine-point circle and the four circles that are tangent to the three sides of the triangle—the incircle and the three excircles. Remember that the nine-point circle passes through the midpoints of the sides of the triangle.

Theorem 8.4 (Feuerbach's Theorem) *Given an arbitrary $\triangle ABC$, the incircle and the three excircles are all tangent to the nine-point circle. The point of tangency between the incircle and the nine-point circle is known as the* Feuerbach Point.

8.17.1 Analysis of the Problem

If we can show that the nine-point circle is tangent to the incircle and an arbitrary one of the excircles we are done, since all of the excircles are effectively equivalent.

Testing figure 8.27 yields zero relationships, interesting or otherwise, so we will need to add a few things to it to get started. The nine-point circle passes through the midpoints of the sides, so we might as well identify them, and it is probably useful to indicate the points of tangency of the incircle and the excircle with the original triangle. The result is illustrated in figure 8.28.

When **Geometer** is used to test this slightly more complex diagram, suddenly, there are many relationships. Since $\triangle A'B'C'$ connecting the midpoints is similar to the original triangle there will obviously be a lot of equal and supplementary angles, but there is one other very interesting equality:

```
Equal length segments:
   [D C'] [E C']
```

Recalling that the subject of this chapter is inversion in a circle and that the incircle and excircle are both tangent to AB at D and E, since C' is equidistant from D and E, if we invert through a circle K centered at C' passing through D and E it will
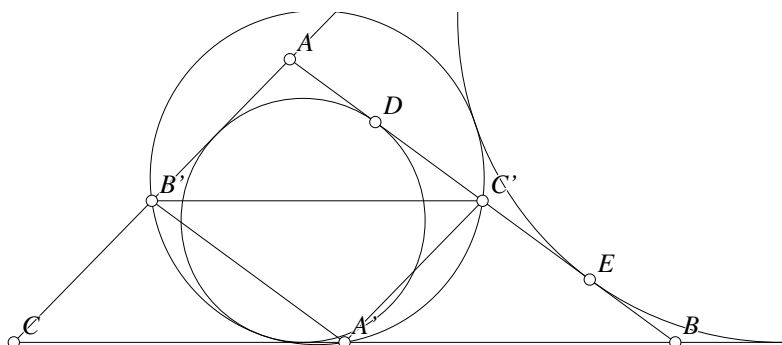
Figure 8.28: Feuerbach's Theorem
Inversion/Feuerbach.T [M]

leave both the incircle and excircle fixed since both will be orthogonal to it. Since the nine-point circle passes through C', the nine-point circle will invert to a line through C'. If, in fact, the nine-point circle is tangent to both the incircle and excircle, so will be the line to which it inverts and that line thus must be a common internal tangent to the incircle and excircle (AB is the other common internal tangent).

If we draw that circle of inversion, the inverted nine-point circle, and identify a few important points we arrive at figure 8.29. In that figure (if the theorem is true), GX is the inversion of the nine-point circle where X is the intersection of the common internal tangents.
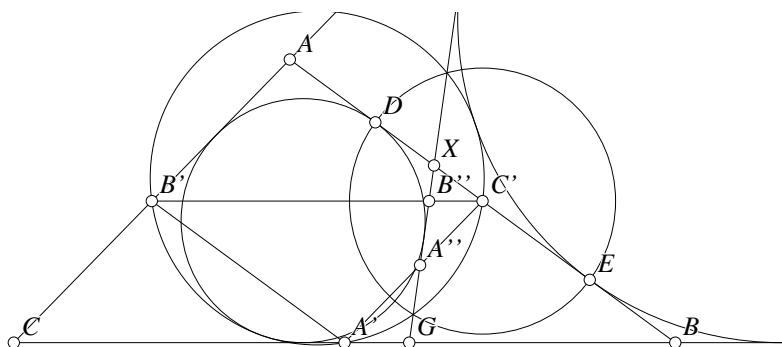
Figure 8.29: Feuerbach's Theorem
Inversion/Feuerbach.T [M]

The easiest way to proceed is simply to draw the other external tangent and prove that it is the inversion of the nine-point circle. To do that, we just need to show that a pair of points on the nine-point circle invert to be on that line. Since inversion takes points directly toward or directly away from the center C', let B'' and A'' be the intersections of the lines $B'C'$ and $A'C'$ with the common tangent CX. If we can show that $B'C' \cdot C'B'' = A'C' \cdot A''C' = C'D \cdot C'D$ we will be done.

Here we will just show that $B'C' \cdot C'B'' = C'D \cdot C'D$ since the proof for the other pair is completely analogous.

If we test the new diagram in figure 8.29 there is a great deal of interesting information, a small amount of which is listed below which certainly shows us that we are on the right track.

```
Supplementary angles:
   (C A B) (C B' A') (C G F) (C' A' B')
 are supplementary to
   (A C' A') (F B'' C')
Points B' and B'' are inverse relative to circle c4.
Points A' and A'' are inverse relative to circle c4.
```

In what follows, we will use the standard notation that $AB = c$, $BC = a$ and $CA = b$. We will also make use of the relationships illustrated in figure A.19 in Appendix A.

We can find the lengths of $C'D = C'E$, of $C'B''$ and of $C'B'$ in terms of s, a, b and c and show that B'' is the inversion of B' relative to the circle K centered at C'.

8.17.2 Proof of Feuerbach's Theorem

As stated above, we want to show that:

$$B'C' \cdot C'B'' = C'D \cdot C'D$$

so we would like to find these lengths in terms of a, b and c.

From figure A.19 we know that $AD = BE = s - a$, and since $AC' = C'B = c/2$, $C'D = C'E = (a-b)/2$. We also know that $B'C' = a/2$, but obtaining an expression for $C'B''$ is a bit trickier.

By AA similarity, $\triangle C'XB'' \sim \triangle BXG$, so:

$$\frac{C'B''}{C'X} = \frac{BG}{BX}. \tag{8.9}$$

Since $C'X = BX - C'B = BX - c/2$, if we can find BX and BG in terms of a, b and c, we will have all the relationships we need.

The point X is at the intersection of the common internal tangents, so X lies on the angle bisector of $\angle BCA$. From this, we know that

$$\frac{BX}{BC} = \frac{AX}{AC}. \tag{8.10}$$

Since $AX + BX = c$ and $AC = b$ we can solve equation 8.10 for BX:

$$\begin{aligned}
\frac{BX}{CX} &= \frac{c - BX}{b} \\
b \cdot BX &= a(c - BX) \\
BX(a + b) &= ac \\
BX &= \frac{ac}{a + b}.
\end{aligned}$$

To obtain a value for BG, notice from figure A.19 that $BG = (s - b) - (s - a) = a - b$. If we put this together with equation 8.9, $C'X = BX - c/2$ and the result above that $BX = ac/(a + b)$, we obtain:

$$
\begin{aligned}
C'B'' &= \frac{BG \cdot C'X}{BX} \\
&= \frac{(a - b)(BX - c/2)}{BX} \\
&= \frac{(a - b)c(a - b)/(2(a + b))}{ac/(a + b)} \\
&= \frac{(a - b)^2}{2a}.
\end{aligned}
$$

Now we can finish the proof:

$$
B'C' \cdot C'B'' = \frac{a}{2} \cdot \frac{(a - b)^2}{2a} = \frac{(a - b)^2}{4} = C'D \cdot C'D.
$$

8.17.3 Feuerbach Consequence

Let H be the orthocenter of $\triangle ABC$. Then the nine-point circle of $\triangle ABC$ is tangent to all twelve of the incircles and excircles of the triangles $\triangle ABH$, $\triangle BCH$, and $\triangle CAH$.

Using Feuerbach's theorem, this is pretty easy to show—see figure 8.30. The nine-point circle of the original triangle is the same as the nine-point circle of each of the smaller triangles, since the nine-point circle passes through the midpoints of the segments connecting the orthocenter with the vertices. Since it is still a nine-point circle, it is tangent to all the circles named. (To avoid too much clutter, the figure includes only one typical incircle and one typical excircle.)
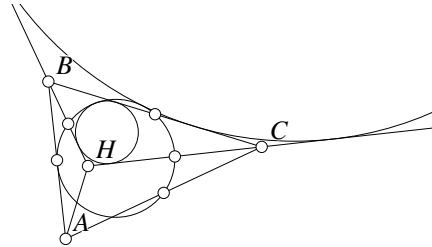
Figure 8.30: Twelve Tangencies
Inversion/Ninefeuer.T [M]
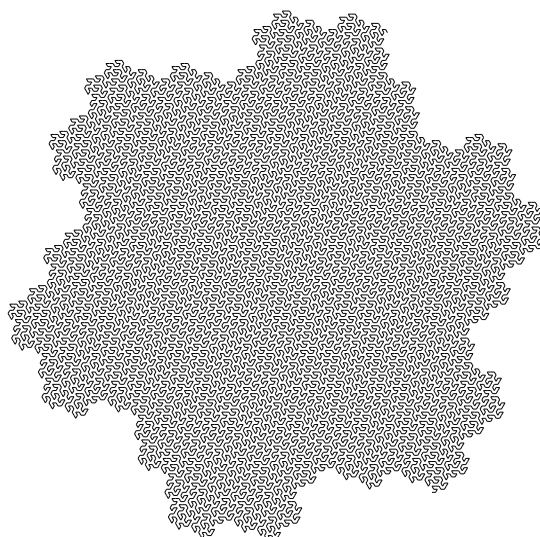
Chapter 9

♦ Projective Geometry

This chapter will not supply many new techniques for using **Geometer** (although it will provide a few). It does, however, provide some insight about how **Geometer** (and probably most other computer geometry programs) do many of their internal calculations in terms of projective geometry as opposed to Euclidean geometry.

We will begin by proving the Euclidean version of a purely projective theorem. A careful analysis of the theorem will then show why it is often useful to view a situation as if it were in projective space rather than Euclidean space.

9.1 Desargues' Theorem

Desargues' theorem, like most of the others in this chapter, is actually a theorem from projective geometry. We will begin by proving the Euclidean form of that theorem.

Theorem 9.1 (Theorem of Desargues) *Let* $\triangle ABC$ *be an arbitrary triangle and* P *be a point different from* A, B *and* C. *Choose points* A', B' *and* C' *on the lines* PA, PB *and* PC, *respectively. Find the intersections of the corresponding sides of triangles* $\triangle ABC$ *and* $\triangle A'B'C'$, *if they exist. In other words, find the points at the intersections of lines* AB *and* $A'B'$, *of* BC *and* $B'C'$ *and of* CA *and* $C'A'$. *The points obtained will lie on a straight line.*

Of course one or more pairs of corresponding sides may be parallel so the points of intersection do not exist.) If not all three intersections exist, the theorem is trivially true, so let us assume that all three points of intersection exist. (See figure 9.1.
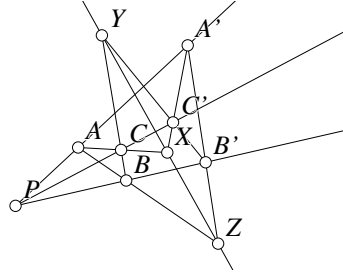
Figure 9.1: Desargues' Theorem
`Projective/Desargues.T [P]`

The proof invokes four applications of Menelaus' theorem. One of the problems with Menelaus' theorem is that it applies to *any* triangle and *any* line that does not go through one of the triangle vertices. In figure 9.1 there are a lot of triangles and a lot of lines.

To use Menelaus' theorem we need to discover length relationships between the points X, Y and Z in the figure with some triangle. In the diagram, these points lie on the lines that make up the sides of either $\triangle ABC$ or $\triangle A'B'C'$. Since the two triangles enter the figure symmetrically, we can use either, so we will arbitrarily choose to use $\triangle ABC$. The triangles are symmetric in this figure only in the sense that if we exchange the labels A with A', B with B' and C with C', the theorem would be no different.

If we are to apply Menelaus' theorem to $\triangle ABC$ and what will be shown to be line XYZ, we will need to find ratios that involve the lengths AX, CX, BY, CY, AZ and BZ. Other interesting lines that are available to us include $A'B'Z$, $A'C'X$ and $B'C'Y$. We also need to make use of the fact that the point P is a center of projection from $\triangle ABC$ to $\triangle A'B'C'$.

All of the conditions in the above paragraph are satisfied if we apply Menelaus' theorem to the following three pairs of triangles and sides: $\triangle ACP$ and line $A'C'X$, $\triangle CBP$ and line $B'C'Y$ and finally $\triangle PAB$ and line $A'B'Z$. If we do so, we obtain the following equations:

$$\frac{PA'}{AA'} \cdot \frac{AX}{CX} \cdot \frac{CC'}{PC'} = 1 \tag{9.1}$$

$$\frac{CY}{BY} \cdot \frac{BB'}{PB'} \cdot \frac{PC'}{CC'} = 1 \tag{9.2}$$

$$\frac{BZ}{AZ} \cdot \frac{PB'}{BB'} \cdot \frac{AA'}{PA'} = 1. \tag{9.3}$$

If we multiply together the three equations 9.1 through 9.3, there is a great deal of cancellation and we finally obtain:

$$\frac{AX}{CX} \cdot \frac{CY}{BY} \cdot \frac{BZ}{AZ} \quad = \quad 1,$$

which, together with Menelaus' theorem applied to $\triangle ABC$, implies that X, Y and Z are collinear. In this book, of course, things work out smoothly, but when you try to do this yourself the first time, it is almost certain that you will write down one of the products of ratios upside-down and you might obtain the following, for example, in place of equation 9.1:

$$\frac{AA'}{PA'} \cdot \frac{CX}{AX} \cdot \frac{PC'}{CC'} \quad = \quad 1.$$

Obviously it is equivalent, but you will find that you sometimes need to invert the ratios to make the terms cancel properly.

Note: There is an interesting arrangement of the code in the **Geometer** diagram for this theorem. The text reuses the various expressions representing the results of different applications of Menelaus' theorem and for this reason it is instructive to examine the **Geometer** code in a text editor.

9.2 Projective Geometry

In this section we take a short but very interesting detour from Euclidean to projective geometry that sheds light on the relationship between the two geometries. Later we will see how projective geometry simplifies writing computer graphics code such as that used in **Geometer**.

In the statement of Desargues' theorem we had to be careful to consider situations where some of the lines of interest were parallel since in those cases, certain of the points of intersection did not exist. It is interesting to experiment with the **Geometer** diagram to see exactly what happens when that occurs. If you move A' or B' to make the line $A'B'$ approximately parallel to AB (keeping, for the moment, the intersections of the pairs AC and $A'C'$ and of BC and $B'C'$), it is certainly true that the point Z moves farther and farther away, but when the lines are parallel and Z, in a sense, "gets to infinity", the infinity it reaches is in the directions of the lines AB, $A'B'$ *and* XY.

In other words, the closer $A'B'$ is to being parallel to AB, the closer XY is to being parallel to both of them. If we were to think of a "point at infinity" where the parallel lines AB and $A'B'$ meet, it would have to be the same "point at infinity" where AB and XY or $A'B'$ and XY would meet.

In Euclidean geometry, two lines can intersect at only one point and we would like this to be true in our new geometry. If we add a point at infinity where parallel lines meet, there would have to be a different point at infinity for every possible direction. If non-parallel lines shared any point at infinity, they would meet at two points—the one at infinity and their normal Euclidean intersection.

Another nice feature of Euclidean geometry is that any two points lie on a unique line. If we have added a point at infinity for every possible direction, on which line do pairs of those points at infinity lie? The easy solution is simply to add a "line at infinity" that includes all the points at infinity.

If we begin with plane Euclidean geometry and just add a point at infinity corresponding to every direction and then place all of those new points on a single line at infinity, we have essentially constructed a perfectly good model for two-dimensional projective geometry!

Note: There is only one point at infinity for any pair of parallel lines. In other words, you "arrive" at the same point at infinity if you follow a line in either direction. In a sense, the lines in projective geometry form loops—if you follow one "to infinity" and then "continue", you find yourself returning to the starting point from the opposite direction.

The main difference between Euclidean and projective geometry is that in Euclidean geometry we have the following two axioms:

- *Every two points determine a unique line.*

- *Every two lines determine a unique point, unless they are parallel.*

In projective geometry, the corresponding axioms are completely symmetric:

- *Every two points determine a unique line.*

- *Every two lines determine a unique point.*

It is worthwhile to verify that these axioms hold in the model of projective geometry that we just constructed by adding points at infinity to the Euclidean plane and the line at infinity:

- If the two points are Euclidean points, they lie on the normal Euclidean line that passes through both of them. If one is a Euclidean point and the other is a point at infinity, they lie on the Euclidean line passing through the Euclidean point and going in the direction indicated by the point at infinity. If both points are at infinity, they lie on the newly created line at infinity.

- If two lines are intersecting Euclidean lines, they meet at the normal Euclidean point. If they are parallel Euclidean lines, they meet at the appropriate point at infinity. If one is the line at infinity, the lines meet at the point at infinity in the direction of the Euclidean line. Since there is only one line at infinity, we need not consider the situation where both lines are at infinity.

In fact, Desargues' theorem is a pure projective geometry theorem. If we consider the points and lines mentioned in the theorem to be projective points and lines, the theorem holds in *every* case, whether the lines be parallel in the Euclidean sense or the points be at infinity. (Note that in projective geometry there is no such thing as parallel lines, since *every* pair of lines meet at a point.)

In the case we first considered in this section where the only problem was that lines AB and $A'B'$ were parallel, their intersection point Z would still exist in projective geometry and since XY in that case will be parallel to AB and $A'B'$, XY will also pass through the same point Z at infinity in that same direction.

If AB is parallel to $A'B'$ *and BC is parallel to $B'C'$*, then both Y and Z will be (different) points at infinity. This means that X must also be a point at infinity, so AC and $A'C'$ must also be parallel in this case. (This is not hard to prove.) Finally, and this is left as an exercise to the reader, the point P could be at infinity, or so can some of the vertices of these (projective) triangles. Think about a few of these and how each situation would correspond to a Euclidean configuration.

9.3 Monge's Theorem

At first glance, Monge's theorem seems completely different from Desargues' theorem but in fact, they are virtually equivalent.

Theorem 9.2 (Monge's Theorem) *Given three circles in the plane, none of which intersect, none of which lie one within the other, and none of which have the same radius, construct the common external tangents of each pair of circles. The three intersections of those common pairs lie on a straight line. See figure 9.2.*

The statement of the theorem can be loosened somewhat. Pairs of circles can intersect and it is still true. If we make the same restriction that we did in the statement of Desargues' theorem that if the external tangents do not meet (in other words they are parallel because the circles are the same size), then the intersections lie on a straight line (although in the case where there are circles with equal radii there can obviously be only two or zero of them).

In fact, if two circles have equal radii, their external tangents meet at a point at infinity, and you will find that the other two pairs of external tangents meet at points such that the line passing through them is parallel to the parallel pair of external tangents. If all three circles have the same radius, the three sets of parallel lines meet at three different points at infinity, and those all lie on the line at infinity, so there is a definite projective flavor to Monge's theorem[1].
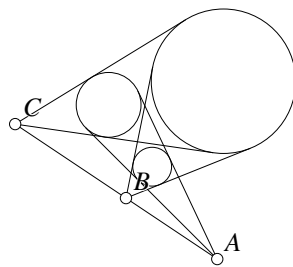

Figure 9.2: Monge's Theorem
`Projective/Monge.T [D]`

---

[1]Monge's theorem is *not* a theorem in projective geometry, since in projective geometry there is no such thing as a circle. A circle can be projected to any conic section, and Monge's theorem does not work if the circles are replaced by ellipses or hyperbolas.
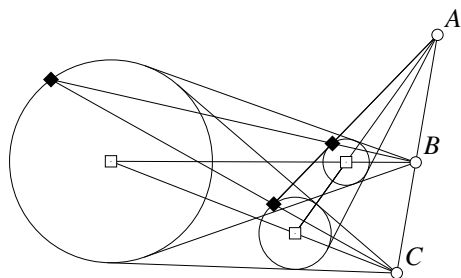
Figure 9.3: Monge's Theorem
`Projective/MongeDes.T [M]`

In figure 9.3 we see Monge's theorem again, but with some additional constructions. It is clear that the line connecting the centers of a pair of circles passes through the intersection of their common external tangents, and in fact, if we think of that point of intersection as a point of projection, we are really just projecting one circle onto another. If that is the case, then *any* corresponding points on the two circles will lie on a line passing through that center of projection.

In the figure, a point drawn as a solid diamond is drawn in the corresponding position (at the same angle from the center) of each circle, and the lines connecting those diamond points also pass through the common intersection of the external tangents and the line of centers of the circles.

Now look at what is basically the same figure with some lines and circles erased (see figure 9.4). In this figure it is obvious that the two triangles whose vertices are all diamonds or all squares are projected from the same point at infinity (the lines connecting the centers and corresponding points on the circles are parallel). Thus the projective version of Desargues' theorem (or a suitably extended Euclidean version) shows that Monge's theorem and Desargues' theorem are equivalent.

One final interesting feature of Monge's theorem is that there is a wonderful, obvious proof if we first make the situation a bit more complicated. Begin with the three circles in a plane, but now consider them not to be circles, but rather spheres that extend above and below the plane. The external tangents become cones that touch each pair of spheres all the way around, and it is clear that the three-dimensional model of cones and spheres, when cut by the plane

Figure 9.4: Monge's Theorem
`Projective/MongeDes1.T [M]`

through the common centers of the spheres, gives the two-dimensional model with which we began.

In three dimensions, consider two additional planes that touch all three spheres above and below. Clearly each of these planes will be tangent to the cones and will thus pass through all three tips. But two planes in three-dimensional space intersect in a line, so the three tips of the cones must lie on that line.

9.4 The Theorem of Pappus

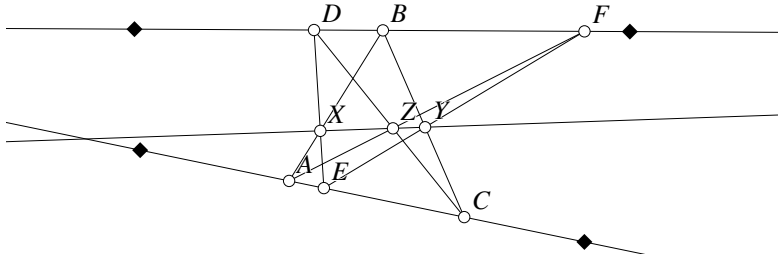Let us take a look at another truly projective theorem—the theorem of Pappus.

Figure 9.5: The Theorem of Pappus
Projective/Pappus.T [M]

Theorem 9.3 (Theorem of Pappus) *Let $ABCDEF$ be a hexagon with points A, C, and E on one line and points B, D, and F on a different one. Then the intersections of the opposite edges of the hexagon (the intersection of AB with DE, of BC with EF, and of CD with FA) lie on a straight line. See figure 9.5.*

The "hexagon" above is probably not what you normally think of as a hexagon, but in projective geometry a hexagon is just a set of six distinct points connected by lines in order. In such a general hexagon, the lines may or may not cross each other (and they surely will cross each other in this example), but to satisfy the conditions of Pappus' theorem, they will clearly have to cross.

We can see that just as in Desargues' theorem, the theorem of Pappus will usually make sense in the Euclidean plane, but there will certainly be situations where one or more pairs of lines are parallel so the points of intersection do not exist in the Euclidean sense. The theorem as stated above, however, is always true in the projective sense, even if, for example, the line containing the points A, C and E happens to be the line at infinity.

Play with the **Geometer** diagram a bit to study its properties. Modify the order of the hexagon's vertices on the lines, and adjust the positions of the lines on which they lie. Check to see that "reasonable" things happen (in the projective sense) when some of the points or lines are the projective points or lines at infinity.

The theorem can, of course, be restated in a Euclidean form (which is the one we will prove) as follows:

Theorem 9.4 (Theorem of Pappus in the Euclidean plane) *If $ABCDEF$ is a hexagon such that points A, C, and E lie one line and B, D, and F lie on a different line. Then the points that lie at the intersections of AB with DE, of BC with EF and of CD with FA, if they exist, all lie on a single line.*
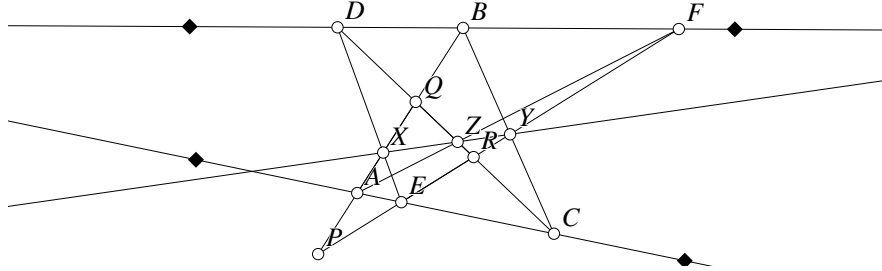
Figure 9.6: The Theorem of Pappus
Projective/Pappus.T [M]

Notice that the Euclidean form could be extended to explain the situations that could occur if various combinations of lines are parallel, but the complete statement of the theorem in this form would probably cover an entire page. Such a "complete" extension would have to include the following cases:

- None of the corresponding pairs of sides are parallel.

- One pair of corresponding sides is parallel.

- If two pairs of corresponding sides are parallel, so is the third.

- One of the points on a line is the point at infinity.

- One point on each of the lines is a point at infinity.

- One of the lines is the line at infinity.

The proof of the Euclidean version of Pappus' theorem where nothing is parallel basically involves an orgy of applications of Menelaus' theorem (see figure 9.6). In that figure, we need to show that the points X, Y, and Z are collinear, and we will do so by showing that the conditions for Menelaus' theorem hold for that line relative to the triangle formed by the sides AB, CD, and EF, which meet at points P, Q, and R in the figure. (We could equally well have used the triangle formed by the sides BC, DE and FA.)

You have to fiddle around a bit to get this to work out nicely with the appropriate things in the numerator and denominator, but the goal is to find the right number of expressions of Menelaus' theorem so that they can be multiplied or divided to give the result. Looking ahead, we want to show that the three points X, Y, and Z lie on a line, and Menelaus' theorem says they will be on a line if and only if

$$\frac{PY}{RY} \cdot \frac{QX}{PX} \cdot \frac{RZ}{QZ} = 1. \tag{9.4}$$

Looking at $\triangle PQR$, we can find a bunch of lines that cut all three of its sides: BYC, DXE, FZA, DBF, and AEC. For each of these lines Menelaus' theorem

will give a product of ratios that is equal to one. In the order above, those products of ratios are:

$$\frac{QB}{PB} \cdot \frac{RC}{QC} \cdot \frac{PY}{RY} \;=\; 1$$

$$\frac{QX}{PX} \cdot \frac{RD}{QD} \cdot \frac{PE}{RE} \;=\; 1$$

$$\frac{QA}{PA} \cdot \frac{RZ}{QZ} \cdot \frac{PF}{RF} \;=\; 1$$

$$\frac{PB}{QB} \cdot \frac{QD}{RD} \cdot \frac{RF}{PF} \;=\; 1$$

$$\frac{PA}{QA} \cdot \frac{QC}{RC} \cdot \frac{RE}{PE} \;=\; 1.$$

If we multiply all five equations above, everything cancels except for the desired result: equation 9.4. Remember that when you are looking for a proof of your own, you probably will not come up with exactly these five equations—some of yours may be inverted and you will find you have to multiply and divide your equations, but the net result will be the same.

Of course this is only a proof of the Euclidean version of the theorem, but the theorem does hold in the general projective sense.

9.5 Another View of Projective Geometry

Imagine that you are looking out a glass window at the three-dimensional world outside, and then, without moving your head at all, you take a paintbrush and at every point on the window, you paint on the glass exactly the color your eye sees there. If your paint colors were accurate enough, if you were a good enough technician, and if you were unable to move your head, you would be unable to tell if you were looking at the scene outside the window or at the painting you had just made on the window glass.

The image on the glass is a projection of the objects in the three-dimensional world on the plane of the window, and if you wish to study this sort of operation, projective geometry is just the thing. Three dimensional computer graphics software would not exist without it.

Although projective geometry exists in any number of dimensions (in the example above we were projecting the three-dimensional world to a two-dimensional painting), in this chapter, unless we state otherwise, the words "projective geometry" will refer to the two-dimensional version (two-dimensions projected onto two dimensions). We can learn quite a bit by looking at one-dimensional projective geometry, and much of that can be applied to higher-dimensional projective spaces as well. In fact, there are even finite projective geometries where the entire geometry consists of a finite number of lines and planes and theorems like those of Desargues and Pappus continue to be true in those finite geometries as well. See Section 9.11.

Go back to the painting example, but imagine that you are not looking at the world through the window, but at a geometric diagram drawn on a flat plane. You can do the same trick with the paintbrush and draw the image of the two-dimensional diagram on the two-dimensional window, and this is a perfectly good projection too. As you move your eye around behind the window, or as you move or tilt the glass window, the painting you would obtain on the window will clearly change.

9.5.1 Geometric Invariants

One way to think about Euclidean geometry is that it is the study of properties of figures that do not change if you move them around or rotate them (or even if you reflect them in a mirror). For example, if two triangles are congruent, they will continue to be congruent if you move them around without distorting them. In the same way, the angles between the lines in a figure remain unchanged if you move the figure rigidly. The areas of geometric objects are similarly unchanged if you move them elsewhere.

Other properties are also preserved, and in a sense they are almost too obvious to mention. For example, if you rigidly move a line or a circle, it remains a line or a circle. If you rigidly move two intersecting lines as a unit, they continue to intersect. There are lots of other properties that are preserved under rigid motions.

Of course there are properties that are not preserved under rigid motion—the distance to a fixed point, for example, or which of two objects is to the right or left of the other (since you can rotate the figure to reverse this).

What if the operations allowed include projections in addition to mirrorings, translations, and rotations? Now what properties are preserved?

Certainly not things like lengths, areas or angles. Projections of circles are very rarely circles. Projections of parallel lines are not parallel.[2] Even things like betweenness drop by the wayside—if point B is between points A and C on a line, the projection of B may not be between the projections of A and C. They will, however, remain on the same projected line.

But plenty of things are preserved (although we do need to make some very slight changes to what we mean by "points" and "lines"):

- If two lines intersect at a point, the projections of those lines intersect at the projection of the point.

- If two points lie on a line, the projections of those points lie on the projection of the line.

- Although circles are not preserved, conic sections are. In other words, if you project any conic section (a circle, ellipse, parabola, or hyperbola), its projection will also be a conic section, although any one of the types might be converted to any of the other types.

---

[2]In fact, in true projective geometry, there really is no such thing as a pair of parallel lines.

- Although betweenness is not preserved, something similar is—separation. If A, B, C, and D lie on a line, and A and C are "separated" by B and D, their projections satisfy the same separation condition. By "A and C are separated by B and D" we mean, roughly, one of B or D is between A and C and the other is not. Another way to think of it is to imagine that the line is bent into a circle, then the points A and C determine two arcs, and if B is on one of those, D is on the other.

Actually, many different geometries besides Euclidean and projective can be classified by the properties that are preserved under different types of transformation. Affine geometry, for example, studies the properties that are preserved under transformations that preserve parallel lines. In this book we will stick to Euclidean and projective geometry, and a bit of inversive geometry in the previous chapter.

9.6 Projective Duality

In Section 9.2 we stated a very interesting property of projective geometry theorems—the axioms are stated in such a way that points and lines are treated completely equivalently, and therefore if you take any projective theorem and swap the occurrences of the words "point" and "line" you will get another theorem that is true.

So what is the dual of Pappus' theorem? Try and figure it out before you read the following description. Even better, try to make a computer geometry drawing that illustrates it and play with it a bit.
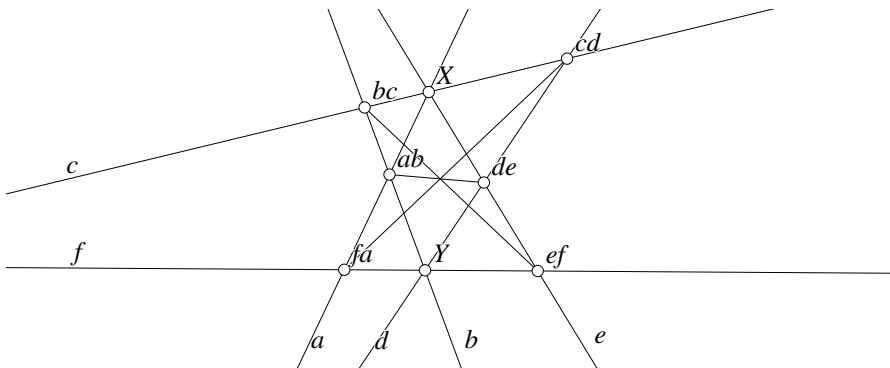
Figure 9.7: Dual of the Theorem of Pappus
`Projective/Pappusdual.T [M]`

Theorem 9.5 (Dual of the Theorem of Pappus) *Let lines a, c, and e pass through a common point, and let lines b, d, and f all pass through another point. Next construct three lines: the line connecting the intersection of lines a and b with the intersection of lines d and e, the line connecting the intersection of the lines b and c with the lines e and f, and the line connecting the intersection of the lines c and d with the intersection of lines f and a. All three of those lines meet at a point.*

In figure 9.7, most of the important intersections are labeled according to the pairs of lines that define them. The lines a, c, and e pass through point X and lines b, d, and f pass through point Y. The intersection of lines a and b is labeled ab; the intersection of lines b and c is labeled bc, and so on. As you can see in the figure, the lines connecting appropriate pairs of intersections all meet at a point.

Since the projective axioms make no distinction between points and lines, a purely projective proof of Pappus' theorem also implies the truth of this dual theorem. If this were only a Euclidean theorem, the proof could be carried out in the same way, but using Ceva's theorem a bunch of times. Ceva's theorem is much like a "dual" of Menelaus' theorem, but remember that since Ceva's theorem (and Menelaus' theorem) both involve lengths, they can not possibly be projective theorems.

9.7 Pascal's Theorem
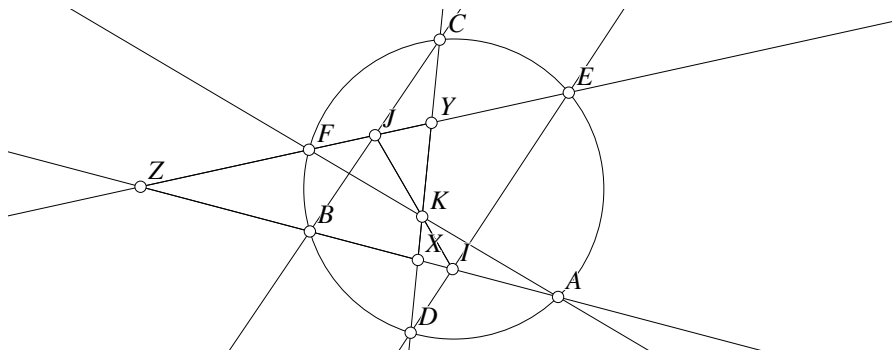
Figure 9.8: The Theorem of Pascal
`Projective/Pascal.T` [M]

Theorem 9.6 (Pascal's Theorem) *Given a hexagon all of whose vertices lie on a circle then the intersections of the opposite sides lie on a straight line.*

In other words, if the points A, B, C, D, E, and F lie on a circle, and I is the intersection of AB and DE, J is the intersection of BC and EF, and K is the intersection of CD and FA, then I, J, and K lie on a straight line.

As you can see from the figure, "hexagon" in this case simply means any six distinct points—the lines connecting them can cross or not. (The reason the lines cross in the figure is so that the intersections will be relatively close to the circle, but this is certainly not required—play with a computer geometry diagram to convince yourself that the ordering of $A, \ldots, F$ on the circle makes no difference.

Also note that the theorem as stated above is not true. There is no reason that the opposite pairs of sides cannot be parallel. In the same way that the Euclidean version of the theorem of Pappus can be written, so can Pascal's theorem. It is not hard to do,

and you should think of exactly how to state it. The statement begins with "If none of the opposite sides are parallel," and then has the statement above. Then work it out if the statement begins, "If one pair of opposite sides is parallel", and so on.

Finally, as stated above, the theorem is not really a theorem in projective geometry either, since there is no such thing as a circle in projective geometry. We will talk about that later. (In fact, it will turn out that the theorem of Pappus is just a special case of the full-blown projective version of Pascal's theorem that is true for hexagons on any conic section.)

As in the case for Desargues' theorem, we will just give the proof in the hard case—where all pairs of opposite sides are not parallel. If one or three pairs of sides are parallel, things are much easier, and you can solve them yourself for "extra credit", if you wish.

The proof is extremely similar to the proof of Pappus' theorem. The only thing we have got to work with is some triangles and some transversals, and the fact that the points lie on a circle, so after you come up with the key observation, the rest of the proof is pretty mechanical.

Exactly as we did for Pappus' theorem, consider $\triangle XYZ$ that is made up of sides AB, CD, and EF of the hexagon. In figure 9.8, X is the intersection of AB and CD, et cetera.

It seems like for these projective-like theorems, the best way to prove the points I, J, and K lie on a line is to use Menelaus' theorem. Here is what we would like to get (relative to $\triangle XYZ$):

$$\frac{ZI}{IZ} \cdot \frac{XK}{KY} \cdot \frac{YJ}{JZ} = 1. \tag{9.5}$$

Relative to $\triangle XYZ$, we can also apply Menelaus' theorem for three different transversal lines: AKF, BJC, and DIE, yielding:

$$\frac{ZA}{AX} \cdot \frac{XK}{KY} \cdot \frac{YF}{FZ} = 1$$

$$\frac{ZB}{BX} \cdot \frac{XC}{CY} \cdot \frac{YJ}{JZ} = 1$$

$$\frac{ZI}{IX} \cdot \frac{XD}{DY} \cdot \frac{YE}{EZ} = 1.$$

but we cannot get the two additional equations that we got for Pappus' theorem since the points do not lie on two lines, but rather on a circle.

We can multiply the three equations above (and rearrange it a bit) to get the following:

$$\frac{(ZI \cdot XK \cdot YJ)}{(IX \cdot KY \cdot JZ)} \cdot \frac{(ZA \cdot ZB) \cdot (XC \cdot XD) \cdot (YF \cdot YE)}{(FZ \cdot EZ) \cdot (BX \cdot AX) \cdot (DY \cdot CY)} = 1. \tag{9.6}$$

The grouping shows that we have exactly what we want on the left-most part of the equation, but there are a bunch of other factors that had better turn out to be one.

Luckily, the six points lie on a circle, so by looking at pairs of chords that intersect at X, Y, and Z we get the following ratios which are exactly what we need to cancel off all the unwanted pairs:

$$
\begin{aligned}
ZA \cdot ZB &= FZ \cdot EZ \\
XC \cdot XD &= BX \cdot AX \\
YF \cdot YE &= DY \cdot CY.
\end{aligned}
$$

These, combined with equation 9.6, yield equation 9.5 and we are done.

9.7.1 Projective Version of Pascal's Theorem

As we stated above, there is no such thing as a circle in projective geometry. Under projections, circles are just one form of a conic section, and conic sections include ellipses, hyperbolas, parabolas, and even degenerate forms, such as a pair of crossing lines. But lines, under projection, remain lines.

Pascal's theorem holds if the six points lie on *any* conic section. (If the conic section happens to be a pair of crossing lines, Pascal's theorem becomes the theorem of Pappus.)
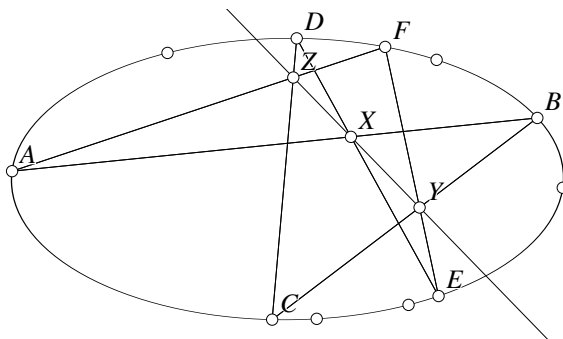
Figure 9.9: The Theorem of Pascal on an Ellipse
`Projective/Pascalconic.T` [M]

We will not prove the fully projective version here, but figure 9.9 shows an example where the conic section happens to be an ellipse and the vertices of the hexagon are in fairly general positions. (The unnamed points in the figure control the shape of the conic section.) The **Geometer** diagram can be modified to illustrate Pascal's theorem on other conic sections.

Figure 9.10 illustrates a proof of a special case of Pascal's theorem for a particular hyperbola ($y = 1/x$), a particular point (the point at infinity in the direction of the y-axis) and a particular line (the x-axis). In the figure, the point A is at infinity, and the hexagon is $ABCDEF$. Lines AB and DE intersect at point α, lines BC and EF at γ, and lines CD and FA at β, where α, β, and γ lie on the x-axis.
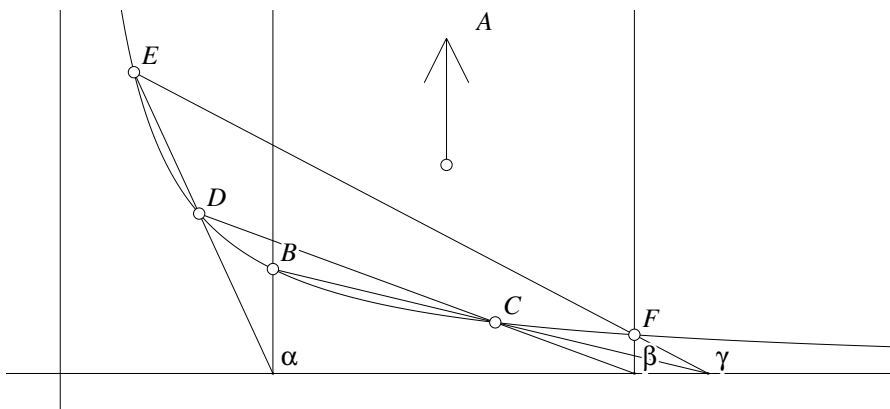
Figure 9.10: The Theorem of Pascal on $y = 1/x$
`Projective/Hyperbola.T` [M]

As you can see, the positions of points D and E completely determine the location of point B, and C and D determine the location of F, so since A is fixed at infinity in the y-axis direction only points C, D and E can be moved in the **Geometer** diagram.

With this setup it is amazingly easy to show that lines BC and EF meet at a point γ that is on the x-axis using just a little bit of analytic geometry.

For *any* two points $P_1 = (x_1, 1/x_1)$ and $P_2 = (x_2, 1/x_2)$ on the hyperbola, we will find where the line $P_1 P_2$ intersects the x-axis. The equation of the line through P_1 and P_2 is given by:

$$y - \frac{1}{x_1} = \left(\frac{\frac{1}{x_2} - \frac{1}{x_1}}{x_2 - x_1} \right)(x - x_1).$$

Set $y = 0$ to find the intersection of the line with the x-axis, and we obtain:

$$-\frac{1}{x_1} = -\left(\frac{1}{x_1 x_2} \right)(x - x_1),$$

or

$$x = x_1 + x_2.$$

In other words, if P_1 has x-coordinate x_1 and P_2 has x-coordinate x_2, then the line $P_1 P_2$ intersects the x-axis at the point having the sum of those coordinates.

Looking back at figure 9.10, let the x coordinates of the points B, C, D, E, and F be b, c, d, e, and f, respectively. We know that $e + d = b$ and $d + c = f$. To show that γ lies on the x-axis, we need to show that the lines BC and EF intersect the x-axis at the same x-coordinate. For BC, that point is at $b + c$, and for EF, it is at $e + f$. But since $e + d = b$ and $d + c = f$ it is clear that $b + c = e + f$ and we are done.

9.7.2 Brianchon's Theorem—Pascal's Dual

Since Pascal's theorem is a pure projective theorem, it must have a dual. But what is the dual of a conic section?

First, we will show a Euclidean version of Brianchon's theorem, then we will talk about the dual of a conic section, and then we will show an example of Brianchon's theorem on a conic section.
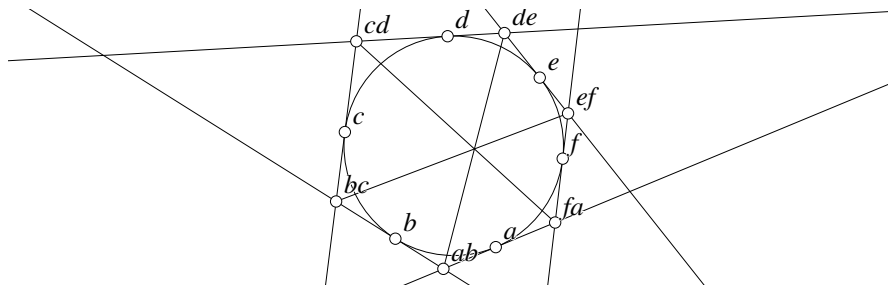
Figure 9.11: Brianchon's Theorem on a Circle
`Projective/Brianchon.T [M]`

Figure 9.11 illustrates Brianchon's theorem on a circle. a, b, ..., f are six lines tangent to a circle. The point labeled ab is where lines a and b intersect. The one labeled bc is where lines b and c intersect, and so on.

If we connect opposite pairs of intersections (ab with de, et cetera), we find that those three lines meet at a point.

The Brianchon projective dual version of Pascal's theorem looks the same—take six lines tangent to a conic, find the intersections of the adjacent ones, and connect opposite pairs of those with lines, and those lines will all meet at a point. This is true, but notice that it is not *exactly* stated as a dual theorem should be. Pascal's theorem begins "Take six points of a conic section...". Brianchon's theorem should then begin "Take six lines of a conic section..."? What is a line of a conic section?
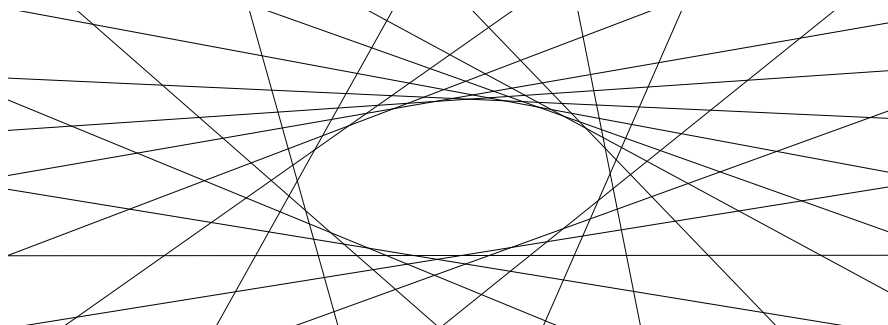
Figure 9.12: A Conic of Lines
`Projective/Coniclines.D [D]`

The answer is that in projective geometry, there can be conic sections composed of points and conic sections composed of lines. The ones composed of lines look to us like the collection of all the lines that are tangent to the conic. Figure 9.12 shows a few of the lines that compose an ellipse-like conic section. It all the lines were shown, of

course, the only white portion of the figure would be the interior of the ellipse. Some even better illustrations of this appear in figure 10.10.
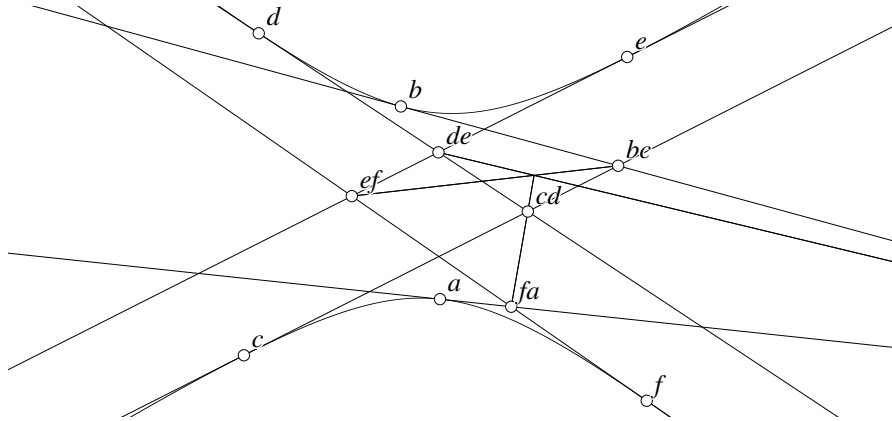
Figure 9.13: Brianchon's Theorem on a Hyperbola
Projective/Brianconic.D [D]

Finally, figure 9.13 illustrates Brianchon's theorem on a conic section (a hyperbola, in this case).
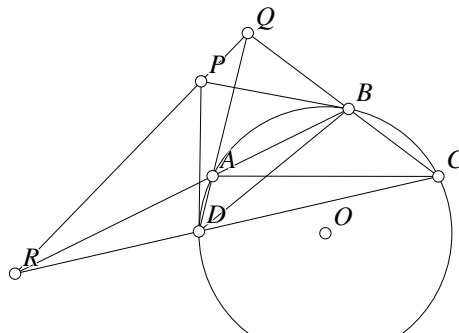
9.7.3 ♦ Degenerate Hexagons

Figure 9.14: Pascal's Theorem on a Degenerate Hexagon
Projective/Degenpascal.T [M]

There are a few interesting facts we can derive by applying Pascal's and Brianchon's theorems to degenerate hexagons. We will look at two examples. See figure 9.14. To do the proof, we actually use a bit of the theory of limits, which is why there is a "♦" in front of this section. It is a very straightforward use, however.

Let $ABCD$ be a cyclic quadrilateral with no pair of parallel sides. Then the points determined by the intersections of the lines AB and CD, of AD and BC, and of the

tangent lines to the circle at points B and D (or of the tangents at A and C) all lie on the same line.

Consider the hexagon $ABBCDD$, or, if you like, consider a hexagon $ABB'CDD'$ where B and B' are *very* close together and D and D' are similarly close. Then Pascal's theorem states that the intersections of AB and CD, of $B'C$ and $D'A$, and of BB' and CC' all lie on a line. As B approaches B' and D approaches D', clearly BB' and DD' approach the tangent lines, and $B'C$ and $D'A$ approach BC and DA.

By relabeling, it is similarly easy to show that the tangents at A and C likewise meet on the same line.
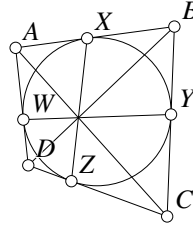
Figure 9.15: Brianchon's Theorem on a Degenerate Hexagon
Projective/Degenbrian.T [M]

As a second example, consider figure 9.15. Given any quadrilateral $ABCD$ circumscribed about a circle, the diagonals AC and BD, and the lines connecting the opposite points of tangency XZ and WY all meet at a point.

Consider the degenerate hexagons $AXBCZD$ and $BYCDWA$. Apply Brianchon's theorem to both, and conclude that all four lines meet at a point.

9.8 Homogeneous (Projective) Coordinates

Just as in Euclidean space you can assign coordinates to your points so that you can do mechanical calculations with them, the same thing can be done in projective space, and the coordinates in this case are called **homogeneous coordinates**.

We will consider two dimensions here, but just by adding or deleting coordinates the situation in more or fewer dimensions is very similar.

If you want to start with the Euclidean plane and add the points at infinity and the line at infinity and to have a uniform method of dealing with them in homogeneous coordinates, here is how to proceed.

- Every point has three coordinates: (x, y, w) and the only constraint is that at least one of the three has to be non-zero.[3]

---

[3]In three dimensions, this becomes: Every point has four coordinates: (x, y, z, w), and the only constraint is that at least one of the four has to be non-zero.

- To convert a standard Euclidean point (not a point at infinity) to homogeneous coordinates, take the two Cartesian coordinates and add a third coordinate of 1. In other words, the point 3 units to the right and 7 units up from the origin which was labeled $(3, 7)$ in cartesian coordinates becomes $(3, 7, 1)$ in homogeneous coordinates.

- Every point has many different representations in homogeneous coordinates. If $\alpha \neq 0$ is any real number, then (x, y, w) and $(\alpha x, \alpha y, \alpha w)$ represent exactly the same point. This is very much like what you do with fractions: $1/3 = 2/6 = 30/90$.

- Remember that every point at infinity can be thought of as a direction. From the origin, pick a point (x, y) on the Euclidean plane, so that the line from the origin to (x, y) goes in the direction you want. Then the point with homogeneous coordinates $(x, y, 0)$ represents that point at infinity. So for example, the point at infinity whose direction is the bisector of the angle between the positive x and y axes has homogeneous coordinates $(1, 1, 0)$ (or $(2, 2, 0)$, or $(\pi, \pi, 0)$—all are just multiples of each other).

- Lines will also have three coordinates, and to differentiate them from points here, we will enclose them in brackets, like this: $[a, b, c]$. Like points, any set of three numbers represents a valid line as long as at least one of the three is non-zero. Similarly, if $\alpha \neq 0$ is any real number, the lines $[a, b, c]$ and $[\alpha a, \alpha b, \alpha c]$ represent the same line.

- If the Euclidean line has the equation $ax + by + c = 0$, it will be represented by the homogeneous coordinates $[a, b, c]$. In this case, it is obvious why $[a, b, c]$ and $[\alpha a, \alpha b, \alpha c]$ represent the same line—one equation for the line is just a multiple of the other.

- The line at infinity can be represented using any of the following sets of homogeneous coordinates $[0, 0, 1]$, $[0, 0, 2]$, or $[0, 0, \sqrt{17}]$—they are all the same. This is clearly not a Euclidean line since it would represent the equation $0x + 0y + 1 = 0$, or $1 = 0$.

- A point (x, y, w) lies on the line $[a, b, c]$ if and only if $ax + by + cw = 0$. For our standard translation of Euclidean points to homogeneous coordinates this makes perfect sense. $w = 1$ and the x and y are the standard coordinates, so the equation becomes $ax + by + c \cdot 1 = 0$. Clearly only the points at infinity will lie on the line $[0, 0, 1]$—the two zeroes in the line's coordinates will eliminate anything the point might have in its first two slots, and if the point is not at infinity, it will have a non-zero number as its third coordinate making $ax + by + cw \neq 0$.

- If (x_0, y_0, w_0) and (x_1, y_1, w_1) are two distinct points, then the line connecting them is given by: $[y_0 w_1 - w_0 y_1, w_0 x_1 - x_0 w_1, x_0 y_1 - y_0 x_1]$. This is easy to check—just make sure that both of the points lie on the line.[4]

---

[4]Although this is not what is going on, if you know something about the three-dimensional vector cross-product, this equation is easy to remember, since the line connecting the points A and B is given by $A \times B$,

- If $[a_0, b_0, c_0]$ and $[a_1, b_1, c_1]$ are two lines, then the point where they meet is given by: $(b_0c_1 - b_1c_0, c_0a_2 - a_0c_2, a_0b_1 - a_1b_0)$. Notice (as must be the case) that this equation and the previous are completely symmetric.

- ♦ If A is any 3-by-3 non-singular matrix, and if v is any point in homogeneous coordinates, then vA represents a projective transformation (where v is considered to be a row vector). In other words, if you multiply every point in your space by the same matrix A, you will get another space where all the projective properties are preserved. If you transform all the points as above, you need to transform the lines l by considering them to be column matrices and calculating $A^{-1}l$. This clearly works, since if a point v is on a line l, then (in matrix notation) $v \cdot l = v(AA^{-1})l = (vA)(A^{-1}l)$. If $v \cdot l = 0$ (the original point was on the line), then the transformed point will be also, and conversely.

9.9 ♦Higher Dimensional Projective Geometry

Like Euclidean geometry, we can consider projective geometry in 1, 2, 3 (and even more) dimensions. To improve your intuition about geometry of any sort in higher dimensions, it is *very* useful to understand thoroughly what goes on in lower dimensional situations. Unless you understand rotation in the two-dimensional plane, you will be totally mystified if you try to understand it in three dimensions.

Many people are interested in the "fourth dimension", but being fundamentally three-dimensional creatures[5] it is difficult to imagine what four dimensions might be like.

Perhaps the best way to get a feeling for four dimensions is to look carefully at the situation in all the dimensions available to you and perhaps you can see patterns that might continue into the fourth dimension. Without too much trouble, you can usually investigate the situation for many kinds of geometry (Euclidean, projective, and others) in one, two, and three dimensions.

An example of this sort of approach to looking at four-dimensional Euclidean geometry can be found in Edwin Abbott's wonderful little book *Flatland: A Romance of Many Dimensions* (see [Abbott, 1952]). Abbott's book was originally written in the late nineteenth century, and has been "updated" by a book called *Sphereland* (see [Burger, 1965]) that extends Flatland's ideas to Einstein's curved and expanding space-time. Both books are worth reading. For those interested, a more advanced, but highly visually-oriented book is called *The Shape of Space* (see [Weeks, 1985]).

Even three-dimensional projective geometry is a bit tricky to visualize (just ask anyone who has tried to write code for a three-dimensional computer graphics system), so we will begin our investigation by looking at the one-dimensional case in detail.

---

where "$\times$" is the three-dimensional vector cross product. Of course the same thing is true in the dual case—the "cross product" of two lines gives the point of intersection.

[5]At least we are three-dimensional in space—treating time as the fourth dimension and going into Einstein's theory of relativity is a bit beyond the scope of this book. Or if the string-theory guys are right, who knows what our "true" dimension is.

Since we have already looked at the situation in two dimensions, we will then be in a good position to talk about what goes on in three (and more) dimensions.

9.9.1 Basic 1-Dimensional Projections

Two-dimensional projective geometry concerns projecting planes onto planes, so one-dimensional projective geometry concerns projecting lines onto lines. This is very nice to work with, since we can draw the entire situation on a plane—simply draw the two lines, the point of projection, and a set of projections, as in figure 9.16. In this figure, P is the center of projection, and the two one-dimensional lines that are projected are AG and $A'G'$. This particular projection takes A to A', B to B', and so on. Exactly the same illustration can be interpreted as a projection from the second line to the first, taking A' to

Figure 9.16: One-Dimensional Projections
`Projective/Onedim.D` [D]

A, B' to B, et cetera. In this particular case, the point P is in the Euclidean plane, but the projection could be from a point at infinity, so all the projection lines would be parallel, coming from the same direction.

Note that the projection maps *all* the points on one line—not just the half-dozen that are shown in the figure. *Every* line through P projects a point on one line to a point on the other. If we happen to choose a line through P that is parallel to one or both of the lines, it maps the "points at infinity" appropriately.
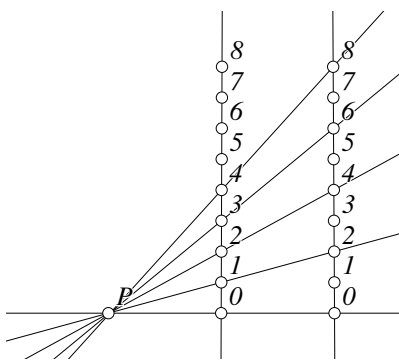
Figure 9.17: A Scaling Projection
`Projective/Scale.D` [D]

In general, the two lines can be anywhere in the plane, at any orientation. P can also be anywhere as long as it is not on one of the lines (otherwise a degenerate projection will occur that maps all the points to a single point), and in fact, P can even be interpreted as a "point at infinity" so that the lines of projection are parallel to each other.

Since our world is fairly nearly a Euclidean one (at least locally), it will be convenient to apply a coordinate system to the lines to be able to describe how the points on the lines are mapped by various sorts of projections.

Before we begin, however, we need to look at one other thing—are we going to need

any "points at infinity" or "lines at infinity"? The answer is yes, but it is not so compli-
cated as in the two-dimensional case. We need only add a single point at infinity to do
the job. You can think of the one-dimensional projective line as a loop—if you go out
forever in the positive direction, you come to the point at infinity, and if you continue in
that direction, you will find yourself coming in toward 0 through the negative numbers.
Since we do not have a whole bunch of points at infinity corresponding to different
directions, but rather a single point, we will just call it "∞" here.

No "line at infinity" is required in 1-
dimensional projective geomety.

In figure 9.17 we see two lines, each of
which is fitted with the usual Euclidean coor-
dinate system—evenly spaced points. These
two happen to be parallel to each other, both
vertical, and aligned so that the two zero-
points are on the same horizontal line. In this
case, the center of projection, P, is on that
same horizontal line of the zeroes, such that
its distance to the first line is the same as the
distance from the first line to the second. It
should be clear that this particular projection
maps every point to a new point with double
the coordinate of the original. (At least that
is if you project from the left line onto the

Figure 9.18: A Translating Projection
`Projective/Translate.D` [D]

right—if you think of the projection as mapping the right line onto the left, then this
projection cuts every coordinate value in half.)

In mathematical terms, the particular projection in figure 9.17 could be denoted by
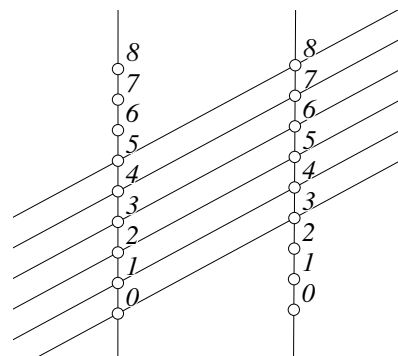$f(x) = 2x$—every value is mapped to twice its original value by the projection.

Figure 9.19: The $1/x$ Projection
`Projective/Invert.D` [D]

It is not hard to convince yourself that
by moving P back and forth along the line,
the projection can represent any multiple you
want. In general, with a setup like this, we can
get anything of the form $f(x) = \alpha x$, where
$\alpha > 0$.

How about negative α? Well, if the sec-
ond line were upside-down so that the nega-
tive coordinates were heading toward the top
of the page, we would have projections for
$f(x) = -\alpha x$, where $\alpha > 0$.

What other operations could be obtained?
Is there a projection that just adds a constant?
Can we make a projection function that does
this: $f(x) = x + 3$? Yes. See figure 9.18. The
two vertical lines are arranged the same way
as previously, but this time we just use a parallel projection that moves each point up

three units. It should be easy to see that by changing the slope of the parallel projection, an arbitrary amount can be added or subtracted, so in general, it is possible to have a projection described mathematically as $f(x) = x + \alpha$, for any α.

Let us consider one more example. See figure 9.19. In this case, the two lines are perpendicular, but they cross at their 1 coordinate, and the center of projection, P, is where the horizontal and vertical lines through the zeroes would meet. Again, without too much difficulty, you can show that $f(1) = 1$, $f(2) = 1/2$, $f(3) = 1/3$, $f(1/4) = 4$, et cetera. In fact, this projection represents $f(x) = 1/x$. What happens when $x = 0$? Just look at the diagram—where would it be mapped? The projection line is parallel to the line it needs to hit, so the intersection must be at the point at infinity, so effectively we have: $f(0) = \infty$ and $f(\infty) = 0$.
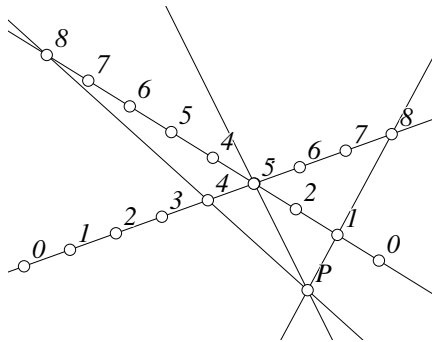
Figure 9.20: Three Point Projections
`Projective/Threepoint.D [D]`

9.9.2 Projection from 3 Points

How much freedom do we have in "designing" one-dimensional projections? The answer is that you can take any three distinct points (including possibly the point at infinity) and specify the three targets of those three points (which also must be distinct, but may include the point at infinity). A projection that achieves this is possible, but once that much is specified, the projections of all the other points are completely determined.

Figure 9.20 shows generally how to do it (the strategy has to be changed somewhat for points at infinity, but not much). In this particular example, suppose we are seeking a projection that makes the following mappings:

$$f(3) = 5 \quad f(1) = 8 \quad f(8) = 4.$$

Move the lines around so that the 3 on the first line lies exactly on top of the 5 on the other. Now no matter where the center of projection is located, the first condition in the list above will be satisfied.

Next, draw the lines connecting 1 on the first line with 8 on the second, and 8 on the first line with 4 on the second. Let those lines intersect at P (and if those lines are parallel, let P be the point at infinity). If P is the center of projection, it clearly satisfies all the conditions of the problem. Furthermore, once we know where P and the two lines are, the projection is completely determined.

Obviously, this does not prove that with another selection of orientations of the lines that a different projection is possible, but that is also true.

9.9.3 Combining 1-Dimensional Projections

What if we do one projection and follow it by another, as in figure 9.21? Here we have two projections. First, using P as the center of projection, the points A, B, …, E are projected to A', B', …, E'. Then those "primed" points are projected to a new line from a new center of projection, Q to make A'', B'', …, E''. This operation could be repeated as many times as desired.
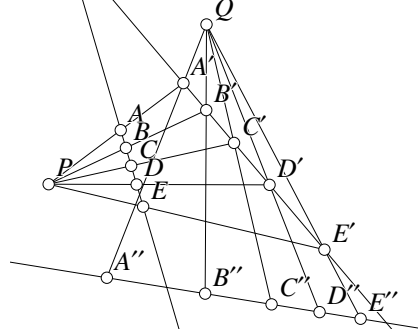
It should be clear that the operation of repeated projection is equivalent to the composition of the mathematical representations of the projections. For example, if the projection centered at point P from line l_1 to line l_2 is represented by $f(x)$ and the projection centered at point Q from line l_2 to line l_3 is represented by $g(x)$, then the combination—project from P, then from Q—will be represented by the composition of functions: $g(f(x))$—first we apply f, then we apply g to the result of applying f.

Figure 9.21: Combining 1-Dimensional Projections

`Projective/Concatenate.T [M]`

To take a concrete example, if the first projection is $f(x) = x + 5$ and the second projection is $g(x) = 1/x$ (both of which we know to be possible projections), the resulting projection is

$$g(f(x)) = g(x + 5) = 1/(x + 5).$$

We already know how to make projections for $f(x) = \alpha x$, $g(x) = x + \beta$, and $h(x) = 1/x$, for any real $\alpha \neq 0$, and for any real β. If we combine these in various ways, how complicated an expression can we get? The answer is that starting from only those three functions, we can make any function of the form:

$$f(x) = \frac{\alpha x + \beta}{\gamma x + \delta},$$

where $\alpha\delta - \beta\gamma \neq 0$, and we cannot get anything more complicated.

We will leave it as an (easy) exercise to show that the function above can be generated from the three primitive operations, but we will show that nothing more complicated can be generated. To see that we cannot get more complicated, let

$$f(x) = \frac{\alpha x + \beta}{\gamma x + \delta} \tag{9.7}$$

and

$$g(x) = \frac{\mu x + \nu}{\zeta x + \theta}, \tag{9.8}$$

where $\alpha\delta - \beta\gamma \neq 0$ and $\mu\theta - \nu\zeta \neq 0$.

Let us calculate $(g(f(x))$:

$$
\begin{aligned}
g(f(x)) &= \frac{\mu\left(\frac{\alpha x + \beta}{\gamma x + \delta}\right) + \nu}{\zeta\left(\frac{\alpha x + \beta}{\gamma x + \delta}\right) + \theta} \\
&= \frac{\mu(\alpha x + \beta) + \nu(\gamma x + \delta)}{\zeta(\alpha x + \beta) + \theta(\gamma x + \delta)} \\
&= \frac{(\mu\alpha + \nu\gamma)x + (\mu\beta + \nu\delta)}{(\zeta\alpha + \theta\gamma)x + (\zeta\beta + \theta\delta)},
\end{aligned}
$$

Which is of the same form, since all the sums of products in parentheses are just real numbers.

The only thing that should be shown is the condition about a non-zero condition like the original $\alpha\delta - \beta\gamma \neq 0$.

For the final expression, this amounts to showing that

$$(\mu\alpha + \nu\gamma)(\zeta\beta + \theta\delta) - (\mu\beta + \nu\delta)(\zeta\alpha + \theta\gamma) \neq 0,$$

which can be simplified as follows:

$$
\begin{aligned}
\mu\alpha\zeta\beta + \mu\alpha\theta\delta + \nu\gamma\zeta\beta + \nu\gamma\theta\delta - \mu\beta\zeta\alpha - \mu\beta\theta\gamma - \nu\delta\zeta\alpha - \nu\delta\theta\gamma &\neq 0, \\
\mu\alpha\theta\delta + \nu\gamma\zeta\beta - \mu\beta\theta\gamma - \nu\delta\zeta\alpha &\neq 0, \\
(\alpha\delta - \beta\gamma)(\mu\theta - \nu\zeta) &\neq 0.
\end{aligned}
$$

This last is clearly true since neither of the terms above can be zero since the definitions of f and g require that each of the terms be non-zero.

It turns out that every one-dimensional projection can be expressed in this form. In fact, the form is a little over-specified. If

$$f(x) = \frac{\alpha x + \beta}{\gamma x + \delta}$$

then multiplying α, β, γ, and δ all by the same constant non-zero number yields exactly the same projection. But since any one of the four numbers can be zero we cannot arbitrarily say that α, for example, is 1.

9.9.4 ♦ Matrix Form in One Dimension

If you know something about matrix algebra, the calculation above should be pretty familiar. Assuming that f and g are defined as in equations 9.7 and 9.8, take the

numbers in the definitions of f and g and put them in matrices as follows[6] and multiply them together as follows:

$$\begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} \begin{pmatrix} \mu & \zeta \\ \nu & \theta \end{pmatrix} = \begin{pmatrix} \mu\alpha + \nu\gamma & \zeta\alpha + \theta\gamma \\ \mu\beta + \nu\delta & \zeta\beta + \theta\delta \end{pmatrix}.$$

The matrix on the right above corresponds exactly to the composition of $g(f(x))$.

But there is even more to the matrix notation. Following what we did in two dimensions in Section 9.8, we can define one-dimensional homogeneous coordinates as follows. The Euclidean point x is represented by $(\alpha x, \alpha)$, where $\alpha \neq 0$, and the point at infinity will be represented by $(\alpha, 0)$, where $\alpha \neq 0$. In other words, if $\beta \neq 0$, then the point with homogeneous coordinates (α, β) corresponds to the Euclidean point α/β.

Now if we write the homogeneous coordinates of a point as a row vector, we can represent the operation of projection as a matrix multiplication. Let f be defined as in equation 9.7, and let it operate on a point with Euclidean coordinate x (which we will represent as the row vector $(x \quad 1)$. The matrix multiplication gives:

$$\begin{pmatrix} x & 1 \end{pmatrix} \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} = \begin{pmatrix} \alpha x + \beta & \gamma x + \delta \end{pmatrix},$$

but the point having the homogeneous coordinates in the vector on the right corresponds to the Euclidean point

$$\frac{\alpha x + \beta}{\gamma x + \delta},$$

as long as the denominator is non-zero, which is exactly what we wanted.

What all of the above shows is that by simply putting the coefficients of the projections into matrices, we can represent the operation of repeated projections as a product of matrices. Similarly, the actual projection can be interpreted as the multiplication of a vector by the projection matrix.

Just as both the homogeneous coordinates of a point may be multiplied by the same non-zero constant and the point will remain the same, all the elements in any projection matrix can be multiplied by the same non-zero constant and it will continue to represent the same projection. This will also hold in any number of dimensions—see Section 9.9.5.

9.9.5 ♦ Matrix Form in n Dimensions

The other nice thing about the matrix representation is that it extends trivially to higher dimensional projective spaces. In two dimensions, for example, any projection or combination of projections can be represented as a 3×3 matrix such as:

$$\begin{pmatrix} \alpha & \beta & \gamma \\ \delta & \epsilon & \theta \\ \eta & \zeta & \iota \end{pmatrix}, \tag{9.9}$$

---

[6]Notice that the rows and columns are swapped—in other words, although the numerator (the "first row") of the expression for f contains $\alpha x + \beta$, the α and β are written in the first column of the matrix. They could have been written as rows, but then the notation would be incompatible with the way that **Geometer** represents transformations.

where the matrix above is non-singular. (The condition $\alpha\delta - \beta\gamma \neq 0$ was sufficient to assure a non-singular matrix in the 2×2 case, but it is more complicated in higher dimensions[7].) The composition of two such projections corresponds to the product of their matrix representations.

If (xw, yw, w) is the homogeneous representation of a two-dimensional projective point, then the operation of projection by the projection represented in 9.9 is achieved by multiplication:

$$\left(\begin{array}{ccc} xw & yw & w \end{array} \right) \left(\begin{array}{ccc} \alpha & \beta & \gamma \\ \delta & \epsilon & \theta \\ \eta & \zeta & \iota \end{array} \right).$$

The pattern continues—in three dimensions, we need a 4×4 matrix, and in n dimensions, an $(n+1) \times (n+1)$ matrix.

This also shows why the images of three points are required to completely determine a one-dimensional projection, and why four points are required in two dimensions, five in three, and so on. The $(n+1) \times (n+1)$ matrix has $(n+1)^2 = n^2+2n+1$ entries, but since any constant multiple of the matrix is equivalent, there are basically only $(n+1)^2 - 1 = n^2 + 2n = n(n+2)$ values that can be chosen freely. In n-dimensional projective space, each image of a point adds n constraints that must be satisfied, or equivalently, n additional equations that must be satisfied. Since there are $n(n+2)$ variables, the images of $n+2$ points are required to completely determine the projection. If $n = 1$ (one dimension), then $n + 2 = 3$ points are required, and so on.

One other nice feature of the matrix representation is that if you want to invert a projection, you simply need to invert the matrix. In other words, if you have the matrix to project from line 1 to line 2, and you want the matrix that undoes that operation and does the reverse projection from line 2 to line 1, simply take the inverse of the matrix. It is clear this will work, since a matrix times its inverse is the identity matrix and multiplying the identity matrix by anything has no effect. That means if you apply the first projection followed by the second, you will come exactly back to where you started which is exactly what is meant by undoing the projection.

9.9.6 Points and Lines in Geometer

Recall in Section 9.8 that the coordinates for a general (two-dimensional) projective point were (a, b, c) and for a line, we wrote them as $[A, B, C]$ where at least one of the numbers in both cases was non-zero. In fact, instead of being written $[A, B, C]$, we should more properly have written a line as a column vector since we have designated points to be row vectors. It takes too much space to write column vectors as a stack of three numbers in text, so it is often written as $(A \ \ B \ \ C)^T$, where the "T" exponent stands for "transpose".

With lines written as column vectors, the transformation of lines by a matrix is again a matrix multiplication. If $(A \ \ B \ \ C)^T$ represents a line and $(A' \ \ B' \ \ C')^T$ is the

---

[7]Well, not *that* complicated—the condition is simply that the determinant be non-zero.

result of transforming it by the matrix above, then:

$$\left(\begin{array}{ccc} \alpha & \beta & \gamma \\ \delta & \epsilon & \theta \\ \eta & \zeta & \iota \end{array} \right) \left(\begin{array}{c} A \\ B \\ C \end{array} \right) = \left(\begin{array}{c} A' \\ B' \\ C' \end{array} \right).$$

The point $(x \ \ y \ \ 1)$ lies on the line $(A \ \ B \ \ C)$ if and only if the product of the row matrix and column matrix is zero:

$$\left(\begin{array}{ccc} x & y & 1 \end{array} \right) \left(\begin{array}{c} A \\ B \\ C \end{array} \right) = (Ax + By + C) = (0).$$

If calculations are done this way in homogeneous coordinates, there is no need ever to check to see if lines are parallel or if points lie at infinity. The right thing happens in every case. If **Geometer** needs to find the intersection of two lines, it blindly performs the cross product operation described in Section 9.8 which works whether the lines be parallel or not. If they were parallel, the result is simply the correct point at infinity that can be used in further calculations.

9.9.7 Transformations Used by Geometer

In **Geometer**'s transformation commands, and even when operations like the rotation of all the primitives are performed, all the transformations are achieved by multiplying the coordinates by a suitable 3×3 matrix after having converted them to homogeneous coordinates with the addition of a "1" as the third coordinate. Here are the basic matrices used by **Geometer** to transform points. In the equations below, $P = (x \ \ y \ \ 1)$ is the input vector:

$$P \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ t_x & t_y & 1 \end{array} \right) = (x + t_x \ \ y + t_y \ \ 1) \tag{9.10}$$

$$P \left(\begin{array}{ccc} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{array} \right) = (s_x x \ \ s_y y \ \ 1) \tag{9.11}$$

$$P \left(\begin{array}{ccc} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{array} \right) = (x \cos \theta - y \sin \theta \ \ x \sin \theta + y \cos \theta \ \ 1) \tag{9.12}$$

Transformation 9.10 translates a point by t_x in the x-direction and by t_y in the y direction. Transformation 9.11 scales a point by a factor of s_x in the x-direction and by s_y in the y-direction. Finally, transformation 9.12 rotates a point counter-clockwise by an angle of θ about the origin.

Combinations of these can be combined simply by multiplying the appropriate matrices together. For example, to rotate about the point (x_0, y_0) one could simply translate by $(-x_0, -y_0)$, bringing the point (x_0, y_0) to the origin. Then a rotation is applied, and finally a translation by $(+x_0, +y_0)$ to return the origin back to its original location.

9.9.8 ◆◆ Constructing a Projection Matrix

In Section 9.9.2 we stated that if you know the images of any three distinct points in the one-dimensional case, the projection is completely determined. So (using homogeneous coordinates), suppose we would like to find the matrix form of a projection f that does the following:

$$
\begin{aligned}
f(x_1, w_1) &= (X_1, W_1) \\
f(x_2, w_2) &= (X_2, W_2) \\
f(x_3, w_3) &= (X_3, W_3)
\end{aligned}
$$

We want to find a matrix

$$
\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}
$$

satisfying these three equations:

$$
\begin{aligned}
(x_1 \quad w_1)\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} &= (X_1 \quad W_1), \\
(x_2 \quad w_2)\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} &= (X_2 \quad W_2), \\
(x_3 \quad w_3)\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} &= (X_3 \quad W_3).
\end{aligned}
$$

Well, not quite. If we could solve the equations above, that would be great, but sometimes we cannot. The reason is, to take the first equation as an example, that (X_1, W_1) is equivalent to $(k_1 X_1, k_1 W_1)$ for any non-zero k_1. Variables k_2 and k_3 could similarly be inserted in the other equations.

Adding the k_i terms, and re-writing the three equations above as a single matrix equation, this is basically the system we are trying to solve:

$$
\begin{pmatrix} x_1 & w_1 \\ x_2 & w_2 \\ x_3 & w_3 \end{pmatrix}\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} k_1 X_1 & k_1 W_1 \\ k_2 X_2 & k_2 W_2 \\ k_3 X_3 & k_3 W_3 \end{pmatrix},
$$

where the twelve known values are the x_i, w_i, X_i, and W_i, and the unknowns are α, β, γ, δ, and the three k_i.

If you multiply the whole mess out, you get six equations and 7 unknowns which is usually bad news, but in our case, we can lock down one of the k_i to be 1 since that would be equivalent to multiplying the whole system by a constant, and then we can solve the six equations and six unknowns in the usual (ugly) way.

A nicer way to do it is the following. It is not too hard to come up with the matrix if the input values are simple. A great set of simple input values are these: $(x_1, w_1) = (1, 0)$, $(x_2, w_2) = (0, 1)$, and $(x_3, w_3) = (1, 1)$. All the ones and zeros make the calculation quite a bit easier, and it is not hard to get the desired matrix.

If you can do this, you can find the matrix form for a projection that takes the points 0, 1, and ∞ to three arbitrary points. But the original problem required a matrix representing the projection from any three points to any other three points. Well, we can find the two transformations that take 0, 1, and ∞ to either the original or final set of points, invert one of them, and multiply the results get the matrix we want.

In other words, suppose we want to map P_1 to Q_1, P_2 to Q_2, and P_3 to Q_3 where the P_i and Q_i are any points represented in homogeneous coordinates. Suppose we find a matrix M such that M takes 0 to P_1, 1 to P_2, and ∞ to P_3. Suppose further that N is a different matrix that takes 0 to Q_1, 1 to Q_2, and ∞ to Q_3. M^{-1} (M inverse) takes P_1 to 0, P_2 to 1, and P_3 to ∞ (it undoes the operation of M).

If we apply M^{-1} first, then N, we take P_1 first to 0, then to Q_1. Similarly, P_2 goes to 1, then to Q_2, and P_3 goes to ∞, and then to Q_3. The matrices M^{-1} and N are combined with a simple matrix multiplication.

This may seem ugly, and it is probably overkill for the one-dimensional case, but in 2, 3, or more dimensions, it is by far the best approach.

In 3 dimensions, for example, a transformation is completely determined by the images of 5 points, so there are basically 20 equations and 20 unknowns (16 unknowns in the 4-by-4 matrix plus the 5 k_i values). But the problem becomes tractable if we find the matrices for a particularly simple set of points: $(1, 0, 0, 0)$, $(0, 1, 0, 0)$, $(0, 0, 1, 0)$, $(0, 0, 0, 1)$, and $(1, 1, 1, 1)$. Find two such transformations and combine one with the inverse of the other to get the desired general transformation.

If you just look at the matrices, it is easy to see why this approach works well. We will show it here for two-dimensional projective space which should make the general approach obvious.

Given 4 points (x_1, y_1, w_1), ..., (x_4, y_4, w_4), we need to find the 3-by-3 matrix $[a_{ij}]$ and constants k_1, k_2, k_3, and k_4 such that:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} k_1 x_1 & k_1 y_1 & k_1 w_1 \\ k_2 x_2 & k_2 y_2 & k_2 w_2 \\ k_3 x_3 & k_3 y_3 & k_3 w_3 \\ k_4 x_4 & k_4 y_4 & k_4 w_4 \end{pmatrix}.$$

If we just multiply this out, we obtain the following 12 equations:

$$\begin{array}{lll} a_{11} = k_1 x_1 & a_{12} = k_1 y_1 & a_{13} = k_1 w_1 \\ a_{21} = k_2 x_2 & a_{22} = k_2 y_2 & a_{23} = k_2 w_2 \\ a_{31} = k_3 x_3 & a_{32} = k_3 y_3 & a_{33} = k_3 w_3 \\ a_{11} + a_{21} + a_{31} = k_4 x_4 & a_{12} + a_{22} + a_{32} = k_4 y_4 & a_{13} + a_{23} + a_{33} = k_4 w_4 \end{array}$$

The first three rows of equations tell us that the a_{ij} are just multiples of the x_i, y_i, and w_i, so the only real unknowns are the k_i. We can let $k_4 = 1$, so all we need to do is solve this system:

$$\begin{aligned}
k_1 x_1 + k_2 x_2 + k_3 x_3 &= x_4 \\
k_1 y_1 + k_2 y_2 + k_3 y_3 &= y_4 \\
k_1 w_1 + k_2 w_2 + k_3 w_3 &= w_4,
\end{aligned}$$

which is just a straightforward system of three equations and three unknowns.

9.10 The Equation of a Conic

Pascal's theorem provides a computationally efficient method to compute the classical equation of a conic given any five distinct points that determine a conic. In classical Euclidean geometry, the equation of a conic is given by:

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0. \qquad (9.13)$$

Obviously, any multiple of this equation yields the same conic, but we cannot, for example, divide through by A, since A might be zero. We are similarly prevented from dividing through by any of the other coefficients.

Thus although there appear to be six degrees of freedom, there are only five. If we know 5 distinct points on it, we should be able to determine all the points on a conic. (We have to be careful of degeneracies—if any three of the points lie on a line, we will get a degenerate conic.)

Suppose you are given the coordinates of five points on a conic: (x_0, y_0), (x_1, y_1), ..., (x_4, y_4). How can you determine A, B, C, D, E, and F in equation 9.13, above?

One way is that since you know the points lie on a conic, simply plug the values (x_i, y_i) into equation 9.13 and get five equations that look like:

$$Ax_i^2 + Bx_i y_i + Cy_i^2 + Dx_i + Ey_i + F = 0.$$

Assume that $F \neq 0$ and divide through by F so the constant term is 1. Then there are five equations and five unknowns, and these can be solved using standard techniques from linear algebra. (Remember that the only variables are A, B, et cetera. The terms with x_i and y_i are constants which you know.)

If this succeeds, great. But what if $F = 0$? What if three of the points lie in a straight line? Well, on a failure, we can then assume $F = 0$ and divide through by E and see if we have better luck. If that fails, we can assume $E = F = 0$ and divide through by D, et cetera. If all of those fail, then we know the set of points must contain three in a line.

But if you are doing this on a computer, it is extremely unlikely that your solution will fail—because of round-off errors and the finite precision of the computer, you will get an answer as if F were something like 10^{-15}—a completely bogus result.

Pascal's theorem provides a much better way.

Given five points (x_i, y_i), convert them to homogeneous coordinates: $(x_i, y_i, 1)$. Add a sixth unknown point that lies on the conic with coordinates $(x, y, 1)$. Now write down the equations that the six points must satisfy according to Pascal's theorem. This is easy to do—we know exactly how to find the equations of the lines that pass through pairs of points, the equations of the points that lie on the intersections of those lines, et cetera. For all the points but one, we have actual numbers. After you grind out the conditions that must be satisfied by the coordinates of your sixth point, it will be in the form of equation 9.13, and you can just read off A through F. If the solution fails for some reason, it is not because of an equation with a zero coefficient; it is because three or more of your given points lie on a straight line.

9.11 Finite Projective Planes

Although the typical computer geometry program will not help much with this subject, it is interesting to note that the axioms of projective geometry nowhere say that there must be an infinite number of points (or an infinite number of lines). There exist, in fact, some finite geometries, one of which has as few as seven points and seven lines.

The smallest finite projective plane contains seven points and is called the Fano Plane..

The seven points are A, B, C, D, E, F, and G, and the seven lines each contain three points. Here are the lines: ADB, BEC, CFA, AGE, BGF, CGD, and DEF. You can easily check that every two points determine a line, and every two lines intersect in exactly one point.

Clearly this is nothing like Euclidean geometry—you cannot draw it with straight lines on a Euclidean plane.

For certain $n \geq 2$ there exists a finite projective plane of with $n^2 + n + 1$ points and the same number of lines. Each line contains $n + 1$ points, and there are $n + 1$ lines passing through each point. You may find it interesting to try to construct examples for $n = 3$ (containing 13 points and lines) or even for $n = 4$ (with 21 points and lines).

There are still open questions about exactly which values of $n \geq 2$ admit the existence of a finite projective plane. It is known that if n is a power of a prime number, then a finite projective plane with $n^2 + n + 1$ points and lines exists, but for other values of n nobody knows. No finite projective planes for such n have ever been found, and for certain values of n the existence of such a projective plane has been proved impossible.

Most small n are powers of primes: $2 = 2^1$, $3 = 3^1$, $4 = 2^2$, $5 = 5^1$, $7 = 7^1$, $8 = 2^3$, $9 = 3^2$, and $11 = 11^1$. So among the first ten possibilities, only 6 and 10 are not powers of prime numbers. It has been proved that a projective plane of order $6^2 + 6 + 1 = 43$ is impossible, but nobody knows whether or not a projective plane corresponding to $n = 10$ (having $10^2 + 10 + 1 = 111$ points and lines with 11 points on each line and 11 lines going through each point) can exist. For more information, see [Ryser, 1963].

9.11.1 The Finite Projective Plane of Order 13

Here is a list of lines for such a plane, where the points are simply A, B, C, ..., M, and the lines are denoted by l_1, l_2, l_3, ..., l_{13}:

$$
\begin{array}{rclrclrcl}
l_1 & = & ABCD & l_2 & = & EFGA & l_3 & = & EHKB \\
l_4 & = & ELJC & l_5 & = & EIMD & l_6 & = & AHIJ \\
l_7 & = & BFIL & l_8 & = & CFHM & l_9 & = & DFJK \\
l_{10} & = & AKLM & l_{11} & = & BGJM & l_{12} & = & CGIK \\
& & & l_{13} & = & DGHL & & &
\end{array}
$$

You can check a few examples to see that every two points have one line in common and every two lines have one point in common.

10
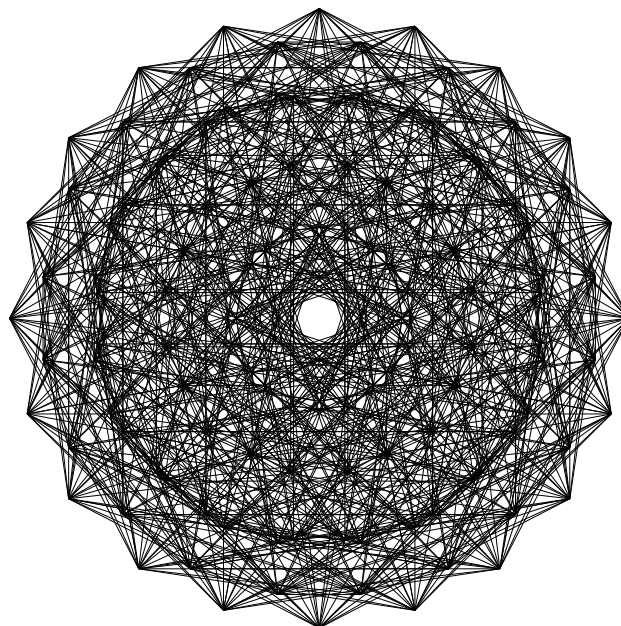
♦ Harmonic Point Sets

Harmonic point sets arise in a surprisingly large number of situations. The condition of harmonicity is fundamentally a concept from projective geometry (see Chapter 9). However, it is often easy to use in Euclidean situations, and it is also related to inversion in a circle (see Chapter 8) and even to musical harmony!

10.1 What is an Harmonic Set?

Let us begin with a simple example. Consider the diagram in figure 10.1. Given collinear points A, B and C, construct a circle with AB as its diameter. Choose a point X arbitrarily on the circle and copy the angle $\angle AXC$ to the other side of the ray XA. (This is most easily done in **Geometer** by reflecting the point C across the line XA.) Let the line XC thus reflected across XA intersect the line AB at a point D. This works whether C is between A and B or not, and only fails if C happens to be the midpoint of the segment AB.

Notice that if you move the point X around on the circle, the location of D does not change—it seems to depend only on the locations of points A, B and C. In fact, if we move the pont X such that the line DX is tangent to the circle, it certainly appears that the line XC is perpendicular to AB, so it *looks* like points D and C are inverses of each other relative to the circle.
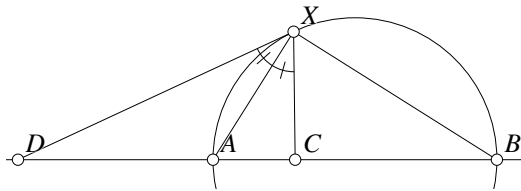
Figure 10.1: Harmonic Division
`Harmonic/BisectHarm.T` [M]

10.1.1 Geometer Analysis of the Situation

If we test the diagram with **Geometer**, moving the points X and C during the test, we obtain a list of relationships that are preserved, and in addition to the obvious ones there are a few interesting ones as well:

```
Points [A B, C D] form an harmonic set.
Points C and D are inverse relative to circle c1.
Triangle area ratios:
    [A C X] x [B D X] = [A D X] x [B C X]
Ratios:
    [B C] x [X D] = [B D] x [X C]
    [A C] x [X D] = [A D] x [X C]
    [A C] x [B D] = [A D] x [B C]
```

The triangle area ratio is the same as the ratio of its bases where the bases are all taken along the line DB with X at the vertex of all of them. **Geometer** also reports that C and D are inverse relative to the circle which we suspected.

But the statement that the points $[AB, CD]$ "form an harmonic set" is new. What does it mean?

10.1.2 Definition of an Harmonic Set of Points

If the points A, B, C and D lie on a line, then the points are said to form an harmonic set (written $[AB, CD]$) if:

$$AC \times BD = AD \times BC. \tag{10.1}$$

When this occurs, one of C or D will be interior to the segment AB and the other will be outside as they are in figure 10.1.

If we rewrite equation 10.1 as ratios, it looks like this:

$$\frac{AC}{BC} = \frac{AD}{BD}. \tag{10.2}$$

If we think of AB as a fixed segment, and if point C lies between A and B, then C divides the segment into a certain ratio. What is interesting is that once C is chosen, there exists another point D that also produces the same ratio, except that now one of the lengths (AD in the case of figure 10.1) will be larger than the original segment.

Another way of saying this is that the points C and D divide the segment AB internally and externally in the same ratio. Once the location of C or D is fixed, the location of the other is determined. An interesting feature of this subdivision is that if C and D divide the segment AB internally and externally in the same ratio, then A and B divide the segment CD internally and externally in the same ratio, so the relationship is symmetric. This is easy to see if we begin with the same equation 10.1 and divide by AD and BD instead to obtain:

$$\frac{AC}{AD} = \frac{BC}{BD}, \tag{10.3}$$

which is exactly what is required to say that A and B divide the segment CD internally and externally in the same ratio. That is why the notation for an harmonic set is symmetric. In fact, we have just proved that if $[AB, CD]$ then $[CD, AB]$.

10.1.3 Back to the Original Problem

The results we guessed after studying figure 10.1 are not hard to prove. An exercise in Chapter 6 whose solution appears in the file `Locus/Locus5.T` proves that the locus of points that lie in a fixed ratio of distances from two given points is a circle, so if the circle passes through A and B, the distances from any point X on the circle to C and D will be in the same ratio as are the distances from A and from B.

If we wish to show that C and D are inverses of each other relative to the circle, that is equivalent to showing that when $XC \perp AB$ the line DX is tangent to the circle at X. The easiest way to proceed is to construct the line OX, where O is the center of the circle.

If it is still not clear how to proceed, construct the **Geometer** diagram similar to the figure (containing the segment OX, where O is the center of the circle), but with a right angle at C and test the diagram. The large set of equal and supplementary angles obtained should provide more than enough hints to arrive at a proof.

10.1.4 Harmonic Point Sets in Projective Geometry

Of course you can see that if C is *very* near the midpoint of the segment AB, then in order for $[AB, CD]$ to hold, the point D must be *very* far away. In fact, once C is the midpoint, there is no Euclidean point D satisfying $[AB, CD]$, but if we choose to include the point at infinity on the line AB, the segment is effectively divided internally and externally in a ratio of $1 : 1$.

General projections do not, of course, preserve lengths, but (as we shall see later) they do preserve sets of harmonic points. In other words, if 4 points on a line are in
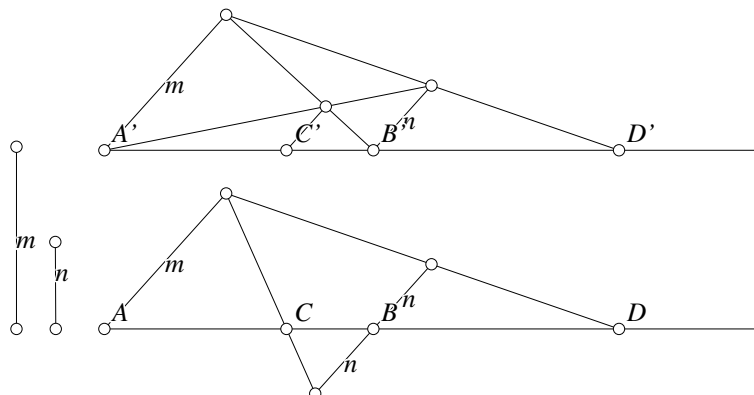
Figure 10.2: Constructing Harmonic Points
`Harmonic/HarmConstruct.T [M]`

an harmonic relationship and they are projected to another line, the projected points will remain in the same harmonic relationship. We will even discover a numerical measurement called the "cross-ratio" that measures for any 4 points how close those points are to being in an harmonic relationship. Even if the points are not in an harmonic relationship, the numerical value of this cross-ratio will also be preserved by projections.

Since harmonicity is a projective concept, then because of duality, there must be a way to assign a meaning to an harmonic set of lines. In fact, four lines are said to be in an harmonic relation if they all pass through a single point and any line not passing through that point intersects the four lines in an harmonic set of points. The term **pencil of lines** means a set of lines passing through a single point.

If you believe that harmonic sets of points are preserved under projection, it is clear that if one line intersects the harmonic pencil of lines, then all other lines will do the same: the intersections are projected from one line to the other from the common point of the four harmonic lines.

10.2   **Some Constructions of Harmonic Sets**

The first example in Section 10.1 provides a method to construct an harmonic set of four points given three of them. In this section we will display a few other approaches. The first two are very straightforward.

Figure 10.2 displays two related constructions of points C and D given A, B, and the ratio m/n into which the segment is to be divided internally and externally. In the two parts of the diagram the segments labeled "m" and "n" are parallel, as is the line passing though C' in the upper diagram. It is a fairly simple exercise in similar triangles
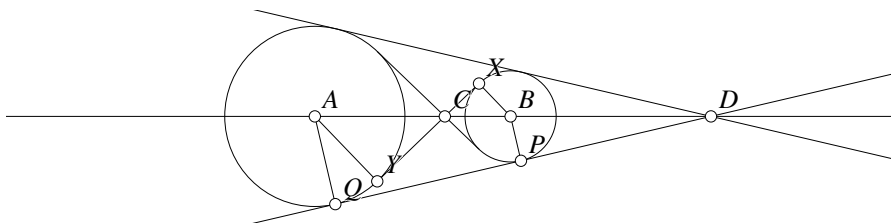
Figure 10.3: Interior and Exterior Tangents
Harmonic/InExt.T [M]

to prove that in both cases the appropriate harmonic division has been constructed:

$$\frac{AC}{BC} = \frac{AD}{BD} = \frac{m}{n} \qquad \text{and} \qquad \frac{A'C'}{B'C'} = \frac{A'D'}{B'D'} = \frac{m}{n}.$$

Note: it may be interesting to examine the **Geometer** file for figure 10.2 in a text editor to see how two constructions can be "synchronized" as they are in this diagram.

Figure 10.3 shows another situation in which harmonic points arise. If A and B are the centers of two circles that do not intersect nor lie one inside the other, the intersections of their common interior and exterior tangent lines (C and D in the figure) form an harmonic set with A and B. This is quite easy to prove if we simply label the points of tangency as in the figure. By similar triangles, we obtain: $AC/R = BC/r$ and $AD/R = BD/r$, where R and r are the radii of the larger and smaller circles, respectively, in the diagram. From this it is trivial to show that $[AB, CD]$.

Although the figures and constructions shown above look simple, an actual construction with straightedge and compass would require quite a few auxiliary lines to construct all the necessary parallel lines. In addition, this method definitely requires a compass since some lengths are copied. The next construction can be done entirely with a straightedge.

Figure 10.4 shows what is perhaps the fastest way to find the point D given A, B and C so that $[AB, CD]$. In the figure the method is applied twice (the second time with primed point names) to illustrate that it does yield the same result no matter how you select the free points.
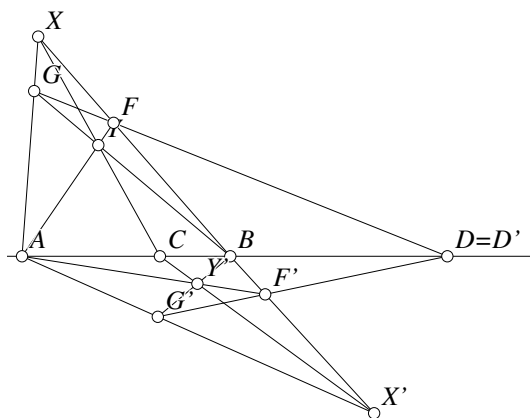
Figure 10.4: Harmonic Division
Harmonic/FreeConstruct.T [M]

names) to illustrate that it does yield the same result no matter how you select the free points.

Given three points A, B and C on a line, select an arbitrary point X not on the line and connect it with straight lines to A, B and C. Now select another arbitrary point Y
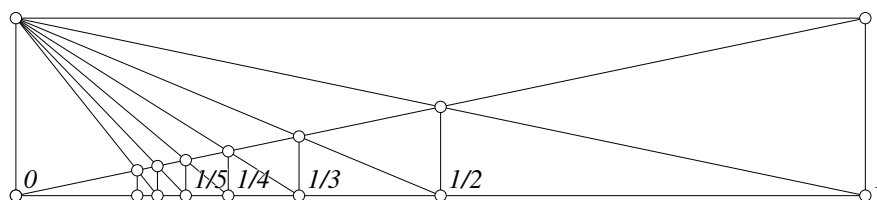
Figure 10.5: Harmonic Construction
`Harmonic/Music.T [M]`
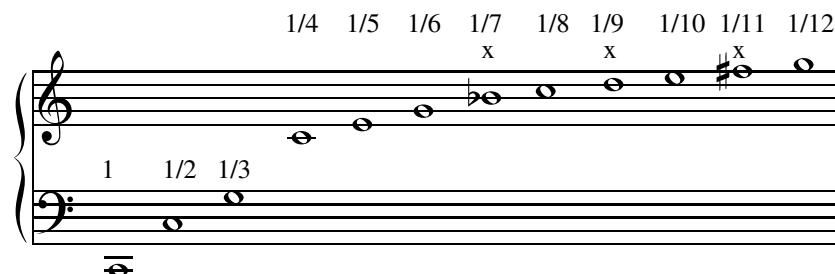
Figure 10.6: Musical Notes and Harmonic Numbers

on the line XC that is different from X and from C. AY and BY intersect BX and AX at points F and G, respectively. The required point D lies at the intersection of GF and AB.

We will delay the proof that this construction does, in fact, yield an harmonic set of points until Section 10.4.

10.3 Harmonic Points and Music

Figure 10.5 is a construction of the points with x-coordinates $1/2, 1/3, 1/4, 1/5, \ldots$ starting from the interval $[0, 1]$. Draw a rectangle whose base is $[01]$ and its diagonals meet over the point $1/2$. Draw the line connecting the upper left corner with the new point $1/2$ and it intersects the long diagonal from 0 over the point labeled $1/3$. Continue as in the diagram to obtain $1/n$ for any integer n.

Labeling the points as in figure 10.5, we obtain all of the following harmonic relationsips: $[0(1/2), (1/3)1], [0(1/3), (1/4)1], \ldots [0(1/n), (1/(n + 1))1]$, with the ratios into which the segments are being divided internally and externally being $1 : 2$, $1 : 3$, $1 : 4$, et cetera. Simple arithmetic shows that this is true, but can you show that the construction does in fact yield those ratios? Do not be surprised if **Geometer** is a little sluggish if you ask it to test the diagram `Harmonic/Music.T`; the list it generates contains 63,129 lines of text specifying relationships that are satisfied!

You may have heard of the **harmonic series** which is simply the diverging sum: $1 + 1/2 + 1/3 + 1/4 + \ldots$. The relationships illustrated by the diagram show why the name may be appropriate.

Figure 10.6 illustrates the relationship between the harmonic numbers and musical notes. If a taut string, when plucked, sounds the low C note (labeled with a "1" above the note in the figure), then if the same string has a node at its center, the note produced will be the C one octave higher (labeled "$1/2$" in the figure). Similarly, if the node is $1/3$, $1/4$, et cetera of the distance from the end, the notes produced are those indicated in the figure by $1/3$, $1/4$, and so on. The notes with an "x" above them are not exact—the frequency at which the string will vibrate is not exactly the defined frequency for the note.

10.4 The Complete Quadrilateral

The word "quadrilateral" means "four sides", and a **complete quadrilateral** begins with four lines shown in figure 10.7 (not including the dashed lines AH, IB and IE). The quadrilateral is made up of the solid lines AB, BC, CD and DA. The only constraint on the lines is that no three of them pass through the same point. This includes points at infinity, in the projective sense—in other words, in addition to no three of the lines passing though a point, no set of three of the lines may be parallel, either[1].

These four lines will meet at six points (A, B, C, D, E and F in the figure). Three pairs of those points are "opposite" each other: A is opposite C, B is opposite D and E is opposite F. The lines connecting these pairs of points (the dashed lines in the diagram) are called **diagonal** lines. If we draw all three diagonals, there will be three additional points where the diagonals intersect, and in the figure those are labeled G, H and I.
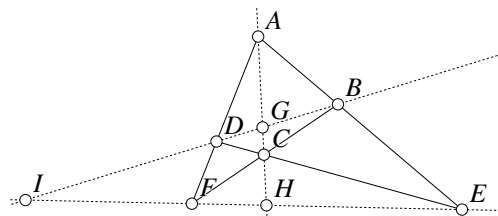
Figure 10.7: The Complete Quadrilateral
`Harmonic/CompleteQuad.T` [M]

From the point of view of harmonic point sets, complete quadrilaterals are very interesting. If you test the diagram in figure 10.7, the following results are obtained:

```
Points [E F, H I] form an harmonic set.
Points [B D, G I] form an harmonic set.
Points [A C, G H] form an harmonic set.
```

Notice also that the right thing happens if a pair of the lines in the quadrangle are parallel. Suppose point C is moved[2] so that $AB \parallel CD$. It is an interesting (but not

---

[1]Since the concept of a complete quadrilateral is a projective as well as a Euclidean notion, there should be a dual concept and there is. The **complete quadrangle** is the figure formed from four arbitrary points and the six lines that join them.

[2]To avoid many other extraneous relationships in the diagram, the points that control the locations of the lines in figure 10.7 are invisible, so to move them you will need to click on the button *Show Invis* to see and manipulate them.

too difficult) exercise in Euclidean geometry to show that when $AB \parallel CD$, the line $IEFH$ is parallel to both of them and $IE = EH$, so $[IH, E\infty]$ (where "∞" is the point at infinity in the direction of that line). In other words, to maintain harmonicity, the point F must move to infinity in the direction of the parallel lines AB and CD.

There is a great deal more harmonicity hidden in figure 10.7. If we simply connect I with C (intersecting AF and AE at N and M) and connect E with G (intersecting FB at K and FA at L), then testing the resulting diagram yields:

```
Points [E G, K L] form an harmonic set.
Points [E F, H I] form an harmonic set.
Points [C I, M N] form an harmonic set.
Points [B D, G I] form an harmonic set.
Points [B C, F K] form an harmonic set.
Points [A N, F D] form an harmonic set.
Points [A D, F L] form an harmonic set.
Points [A C, G H] form an harmonic set.
Points [A M, E B] form an harmonic set.
```

If we are going to prove, for example, that $[EF, HI]$ in figure 10.7 we are going to need a bit more than what **Geometer**'s *Test Diagram* tells us. One nice way to proceed is to look at the relationships that hold for a single point that divides a segment.

In figure 10.8 we have a segment AB that is divided (in this case internally) by a point C. In the complete quadrilateral, all three points are connected to another one (called P in this example).

We would like to find an alternative expression for AC/BC involving the other sides of the triangle or the angles or something. In the figure, the other three sides have lengths a, b and x, and the angles are named α and β. If we apply the law of sines to the triangles $\triangle APC$ and $\triangle BPC$ we obtain:
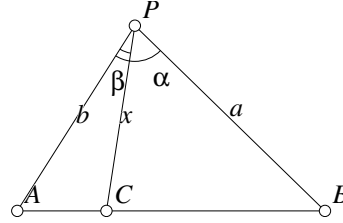
Figure 10.8: Dividing a Segment
`Harmonic/SegDivide.D [D]`

$$\frac{AC}{\sin \beta} = \frac{x}{\sin A} \qquad \text{and} \qquad \frac{BC}{\sin \alpha} = \frac{x}{\sin B}. \tag{10.4}$$

If we combine the two equations in 10.4 in such a way that we can eliminate the variable x, we obtain:

$$\frac{AC}{BC} = \frac{\sin \beta \sin B}{\sin \alpha \sin A} \tag{10.5}$$

Note that this relationship holds whether C is internal or external to the segment AB.

If there is another point D dividing AB, we can obtain a similar expression by replacing C by D, α by γ and β by δ in equation 10.5, where γ and δ are the angles made by the line PD with the sides PB and PA, respectively. Here is the result:

$$\frac{AD}{BD} = \frac{\sin \delta \sin B}{\sin \gamma \sin A} \tag{10.6}$$

If $[AB, CD]$ then $(AC/BC = AD/BD$, or $(AC \cdot BD)/(AD \cdot BC) = 1$, so let us calculate the ratio $(AC \cdot BD)/(AD \cdot BC)$ from equations 10.5 and 10.6:

$$\frac{AC \cdot BD}{AD \cdot BC} = \frac{\sin \beta \sin \gamma}{\sin \alpha \sin \delta}. \tag{10.7}$$

Notice that the expression on the right side of equation 10.7 depends only on the angles made by the sides PC and PD relative to the sides PA and PB. This means that once the four lines that meet at P are determined, *any* line not passing through P will have the same value of the expression on the right of equation 10.7 as any other. In other words, the value of that expression is preserved by projection onto another line. If it is 1, meaning that the four points are in a harmonic configuration, they will remain harmonic after a projection, and if they were not, they can never be projected into an harmonic configuration.

The quantity $(AC \cdot BD)/(AD \cdot BC)$ is very important since it is not only preserved by projection, but its value is 1 exactly when the four points are in an harmonic configuration. It is called the **cross ratio** of the four points[3]. The cross ratio is usually written: $\{AB, CD\} = (AC \cdot BD)/(AD \cdot BC)$.

If we return to figure 10.1 which is the first time we encountered points in an harmonic ratio, it is easy to prove that the construction there generates an harmonic set. In that figure, let $\theta = \angle DXA = \angle CXA$ and let $\psi = \angle BXC$. We know that $\theta + \psi = 90°$. If we substitute these angles into the expression on the right of equation 10.7 we obtain:

$$\frac{AC \cdot BD}{AD \cdot BC} = \frac{\sin \theta \sin(90° + \theta)}{\sin \psi \sin(-\theta)} = \frac{\sin \theta \sin \theta}{\sin(90° - \theta)(-\sin \theta)} = \frac{\sin \theta \sin \theta}{(-\sin \theta)(-\sin \theta)} = 1.$$

Since the cross ratio is preserved under projection, we only need to prove that the various harmonic ratios observed in the complete quadrilateral hold in a single configuration. This is because any configuration can be projected to any other (four points can be mapped to four other arbitrary points, assuming the the initial and final configurations are non-degenerate—see Section 9.9.8). We can choose that single configuration any way we want. Clearly we could choose a nice Euclidean configuration and calculate all the distances, so the result can be proved completely in Euclidean geometry, but if we are willing to look at the result in a projective space, the result becomes almost trivial.

In figure 10.7, for example, if $ABCD$ is a square, then clearly G will be at the center, so AG will meet EF at a point H that is the midpoint of FE. The segment DB will be parallel to FE, so the point I will become the point at infinity. Thus the points I, F, H and E satisfy $[EF, HI]$ and so they will for any configuration of A, B, C and D. The harmonicity of $[BD, GI]$ and $[AC, GH]$ can be proved similarly.

A quick comparison of figures 10.7 and 10.4 should make it obvious why the straightedge-only construction works—it just constructs a complete quadrilateral with A, B and C suitably placed.

---

[3]Sometimes directed segments are used in the definition of the cross ratio, and if so, the statement that the four points lie in an harmonic configuration replaces the value 1 by -1

One final note about the cross ratio is this: If we know the cross ratio of four points in a particular order, the cross ratios of all other possible arrangements of those points are determined. A little algebra shows, for example, that if $\{AB, CD\} = k$ then $\{BA, CD\} = 1/k$, $\{AC, BD\} = 1 + k$, and so on.

10.5 Poles and Polars

We noted earlier that there seems to be a relationship between inversion in a circle and an harmonic set of points. In fact, in figure 10.1 the points D and C are inverses relative to the circle with diameter AB. Inversion is very powerful in Euclidean geometry so it is reasonable to search for a similar operation in projective geometry where this new operation will, of course, have a dual operation where the roles of "points" and "lines" are interchanged.

This is accomplished with poles and polars but instead of mapping a point to its inverse point, a point is mapped to its polar which is a line that is **reciprocal** to that point. The reciprocal of a line is its pole which is the point whose polar is the line. In other words, the projective version of inversion will associate each point with a line and vice-versa in a manner that is similar to inversion in Euclidean geometry.

Of course in projective geometry there is no such thing as a circle, and a full-blown projective definition of poles and polars must apply to a general conic section, but let us begin by looking at poles and polars in the Euclidean plane where this new sort of inversion will be relative to a circle.

Figure 10.9: Pole and Polar
`Harmonic/PolePolar.D` [M]

Figure 10.9 shows the construction. Given a circle and a point A, let A' be the inverse (point) of A relative to that circle. Now construct a line through A' perpendicular to the line AA'. This line is the **polar** of A, and A is the **pole** of this line. Thus there is a polar for every point except the center of the circle (and in projective geometry, of course, the polar of the center is the line at infinity). Similarly, every line has a pole except for the lines that pass through the center of the circle (and in projective geometry those lines have poles at infinity in the direction perpendicular to each line).

At first it seems that there is very little difference between the concept of pole and polar versus the Euclidean inverse in a circle. Figure 10.10 illustrates the result of finding the polars of a set of evenly-spaced points around a circle relative to a fixed circle in the center of each of the four examples. From the figure it is obvious that the polars of the points in a circle do not form another circle (or straight line) as they do when those same points are subjected to the standard Euclidean inversion. Figure 10.10 illustrates only a small number of examples. The **Geometer** diagram allows you to
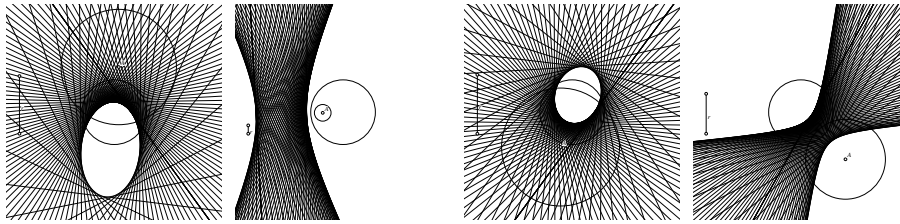
Figure 10.10: Polars of Circles
`Harmonic/PolarInverse.T [S]`

Figure 10.11: Poles, Polars, and Conjugates
`Harmonic/PolePolar1.D [M]`

adjust the center of the circle of points whose polars are to be drawn and the length of the line segment on the left controls the radius of that circle. When the center and radius are set, the *Run Script* command draws a set of polars relative to 90 evenly-spaced points around the circle. The **Geometer** script itself perhaps merits an examination in a text editor to see exactly how it works.

Here are a few basic properties of Euclidean (as opposed to projective) poles and polars as defined relative to a circle (instead of relative to a general conic section). None of them is particularly hard to prove, and most depend on the properties of inversion in a circle.

1. If A lies on the circle, $A = A'$ and the polar of A will pass through A and be tangent to the circle at that point. See the left diagram in figure 10.11.

2. If B is a point on the polar of point A, then A lies on the polar of B. In this situation, A and B are called **conjugate points** and their polars are **conjugate lines**. See the right diagram in figure 10.11.

3. If a secant through a circle is drawn through a point P and tangents are drawn at the points where the secant intersects the circle, then the intersection of those tangent lines lies on the polar of P. If this process is repeated, the two points found on the polar can be used to construct the polar. See figure 10.12 where the intersections of the tangents at Q' and Q'' determine the polar of the point P.

4. If the line joining two conjugate points intersects the circle, the two intersections of that line with the circle and the two conjugate points form an harmonic set.

5. See figure 10.13. If two secants are drawn from a point P that is exterior to the circle and the complete quadrangle is formed from the four intersections (A', A'',
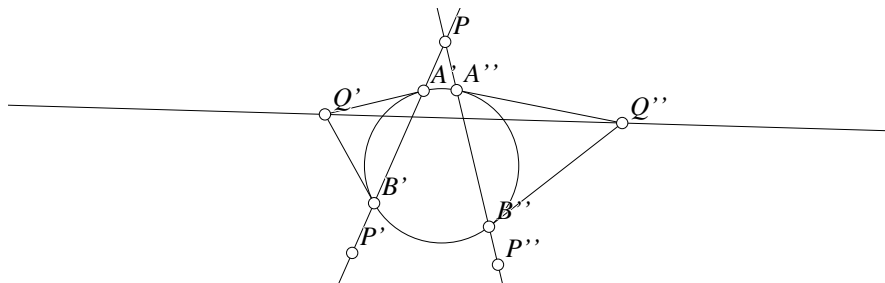
Figure 10.12: Constructing a Polar
Harmonic/PolePolar2.D [M]
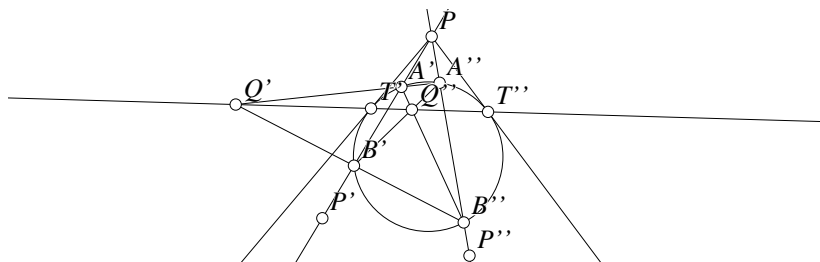
Figure 10.13: Polar and Tangent Construction
Harmonic/PolePolar3.D [M]
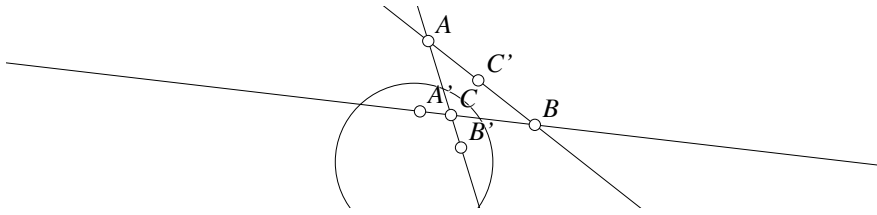
Figure 10.14: The Polar Circle
`Harmonic/PolarCircle.T` [M]

B' and B'') of the secants with the circle, then the two additional intersections in the complete quadrangle (Q' and Q'') lie on the polar of the point P. This is a straightedge-only construction of the polar. Furthermore, the intersections of the polar with the circle when connected to P (T' and T'') are the two tangents to the circle passing through P, so tangents to a circle can also be constructed with a straightedge only.

6. The angle between the polars of A and B is the same as the angle $\angle AOB$, where O is the center of the circle.

7. The distances from the center of a circle to any two points is proportional to the distances of each point from the polar of the other.

8. If two circles are orthogonal, then the (extension of a) diameter of one will cut the other in two points that are harmonic relative to the endpoints of the diameter in the first, and conversely.

9. A triangle is called **self conjugate** relative to a circle if its sides are the polars of its vertices. A self conjugate triangle can be constructed by taking an arbitrary point other than the center of the circle as one vertex, then choosing another point on its polar as the second vertex, and the third vertex at the intersection of the polars of the first two. Such a triangle must contain an obtuse angle. In figure 10.14, the triangle $\triangle ABC$ is self conjugate since the polars of points A, B and C are the lines BC, CA and AB, respectively.

10. If a triangle is self conjugate relative to a circle, that circle is called the **polar circle** of the triangle. Any obtuse triangle has a polar circle. In figure 10.14 the circle is the polar circle of $\triangle ABC$.

11. The nine-point circle and the circumcircle are inverses of each other relative to the polar circle. This is "obvious" since the circumcircle passes through the vertices of a triangle $\triangle ABC$ and the nine-point circle passes through the feet of the altitudes, which are the inverses of the vertices A, B and C. Figure 10.14 makes it clear exactly what is going on.

12. If a complete quadrangle $ABCD$ is inscribed in a circle, the intersections of the lines AB with CD, AC with BD and AD with BC form a self-conjugate triangle. See figure 10.15.
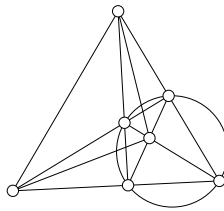
Figure 10.15: Constructing a Self Conjugate Triangle
`Harmonic/SelfConj.T [M]`
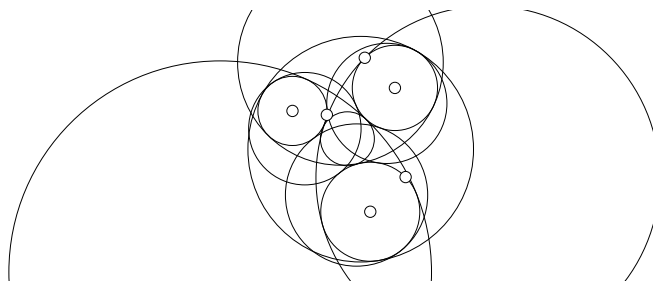
Figure 10.16: Apollonius' CCC problem
`Harmonic/App10.T [M]`

13. The altitudes of a self-conjugate triangle pass through the center of the circle.

14. The polar of a point A relative to a circle with center O is the radical axis of that circle with the circle having diameter OA.

15. The reciprocal of any rectangle with the same center as the circle is a rhombus.

10.6 The Problem of Apollonius (Again)

In Section 8.15 we presented a solution, using inversion in a circle, to the problem of constructing all the circles that are simultaneously tangent to three given circles which we labeled there as CCC. The solution there required that certain decisions be made during the construction. For example, the first step was to decide which of the three circles had the smallest radius so that that length could be subtracted from the other two to reduce the problem to one of finding the circles tangent to two given circles and passing through a point.

The method works fine, but it is difficult to produce a **Geometer** diagram that works for arbitrary circles using this method. In fact, in figure 8.25 the centers and radii of the three given circles can only be moved a little bit before the **Geometer** diagram fails to work.

Using poles of a circle, there is a mechanical construction of the tangent circles due to Gergonne which we present here without proof.

For each pair of circles, find the two homothetic centers. The homothetic centers are the points from which one circle is a projection of the other. If the circles are disjoint, for example, the homothetic centers are the points where the internal tangents intersect and where the external tangents intersect.

These six points lie on four lines. For each of those lines, find the poles of that line relative to each of the three circles, and if they all exist, connect them to the radical center of the three circles. The lines connecting the poles to the radical center will intersect each circle twice, and those six points are the points of tangency of two of the solutions to the CCC Apollonius' problem. Since there are four lines, there are up to eight solutions. The **Geometer** file `Harmonic/App10.T` contains that construction, and is illustrated in figure 10.16.

10.7 Poles and Polars Relative to Conics

We cannot construct poles and polars for a general conic using the Euclidean concept of inversion, but a construction like that used in example 5 in the previous section will work for any conic section—not just a circle.

Figure 10.17: Projective Pole and Polar
`Harmonic/ProjPolar.T` [M]

It is surprisingly difficult to construct a **Geometer** diagram that will produce the polar for any pole and any conic section as illustrated by the example in figure 10.17. In that figure (and in the **Geometer** diagram) the points indicated with solid diamonds control the shape of the conic section. They can be moved as can the point A. The line is the polar of A.

With a fixed point and a fixed conic, it is easy—begin by drawing two lines that pass through the conic in two different places and then continue exactly as was done with the circle. The problem, of course, is how does one draw two lines that are guaranteed to be different and to intersect the conic without knowing ahead of time what the conic looks like and where the point is relative to it? It is a very good exercise to try to make a diagram like that in figure 10.17 before reading the following material.

We can easily obtain one point (say P) on the conic using **Geometer**'s *P on Conic* command. The line connecting A and P will intersect the conic in another point, but even that is hard to pick consistently. If you simply choose the other point of intersection Q of the line AP with the conic using **Geometer**'s *LCon=>P* command it will work fine for the initial construction, but as you change the shape of the conic and the location of A you may discover that suddenly P and Q are the same point so the line passing through the pair of them will behave erratically, to say the least!

What happens is that upon selecting the line and conic when you use the *LCon=>P* command, **Geometer** finds the point nearest the last location of the mouse when you clicked on the conic. **Geometer** solves the line-conic intersection problem by solving a quadratic equation that has two roots, depending upon whether the sign of the

discriminant is positive or negative. If it is positive for for the point of intersection, **Geometer** *always* uses the positive value of the discriminant for that solution and vice-versa. The new location of the point P on the conic, however, is constantly reprojected to the conic as A moves or as the shape of the conic changes. The location of P may suddenly become the same as Q because of that.

The best solution is that once you have the line AP that intersects the conic twice to change P to the invisible color, and then create two new points at the two intersections of the line with the conic. **Geometer** will find that it uses a positive discriminant for one and a negative discriminant for the other, so even if P swaps to a different solution, the two intersections will remain different.

Thus it is easy to obtain *one* line and its two different intersections with the conic. How do you obtain another line and pair of points?

If you simply choose another point P' constrained to lie on the conic, as you move things around and P and P' are constantly reprojected, it is overwhelmingly likely that they will eventually converge to the same point. Try making a conic section through 5 points, constrain two points to lie on that conic and then change the shape of the conic by moving the control points. You will find that almost always the two points will converge to the same location on the conic. They may not be identical due to computer round-off error, but the distance between them will be so tiny that the line that joins them is very unstable.

But a perpendicular through P is guaranteed to intersect the conic in a new point, and that new point is generally not near either of the two intersections of the line AP. Of course this point has the same problem with respect to its definition and if we are not careful, it may collapse to be the same as P. If you try the same trick, however, that may not work, since P and this new point may swap, and your line through P and A will, essentially, not pass through P.

Here is the solution: Pick P on the conic and draw the line AP. Through P, construct a perpendicular to AP. Make P invisible and find the two intersections of the perpendicular line with the conic. When these different points are connected to A, they are guaranteed to be different lines[4].

Draw the two lines from the intersections of the line perpendicular to AP at P. Next, color the intersection points invisible, and make two new points on each line that intersect the conic in different places. These four points can then be used to construct the polar in the same manner that it was constructed for a circle.

---

[4]Well, almost guaranteed: with incredibly bad luck, P may lie at a point where AP is tangent to the conic and hence only intersects it in one place, but this is *very, very* unlikely.
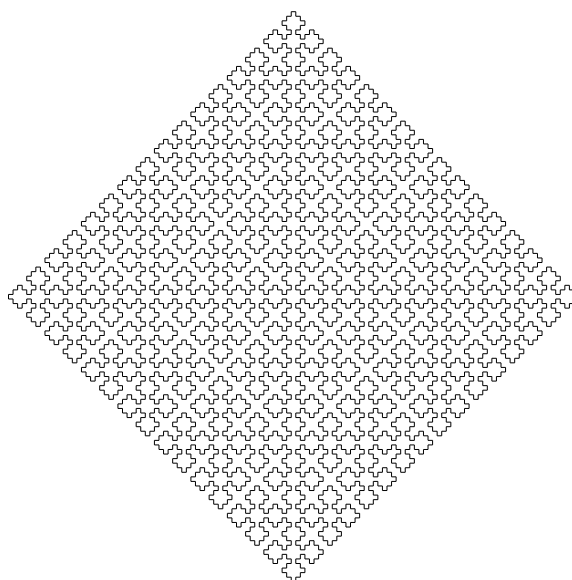
Chapter 11

Geometric Presentations

If you are trying to solve a problem for yourself then as long as you understand what you are doing, the beauty and neatness of the solution is not too important. But as soon as you try to present your results to someone else it is important that the presentation be neat and clear. If you are a geometry teacher using **Geometer** in front of a class, the presentation will be easier for your students to understand if you make some effort to polish each diagram before you present it.

In this chapter we will investigate methods to make diagrams easier to understand. Teachers of geometry need to do this every day, but many mathematicians need to do the same thing from time to time. In addition to dynamic presentations, **Geometer** can also be quite useful for helping generate illustrations for publication, either electronically or printed on paper. Every illustration in this book (and every one in the reference manual except for the screen shots) was prepared using **Geometer**.

The chapter is much more about techniques for using **Geometer** (or other computer geometry programs) than it is about specific geometric facts, although the examples

have been chosen to include some interesting results. Since many of the features and methods described here are specific to **Geometer** you may wish to have the reference guide available as you read this. Other computer geometry programs have most of these features, but probably not all, and **Geometer** is missing some features found in those programs.

The reference guide that comes with the **Geometer** release contains a tutorial for teachers that is a useful addendum to the material in this chapter.

11.1   General Considerations

There are some obvious things to keep in mind, no matter whether your presentation is electronic, on an overhead projector, written in chalk on a blackboard or printed in a book or handout. Here are a few:

- Use standard notation if possible. If that is impossible, at least use consistent notation. In other words, use "I" for the incenter of a triangle, "O" for the circumcenter and "H" for the orthocenter. If your triangle is $\triangle ABC$, let a, b and c represent the lengths of the sides opposite the vertices A, B and C, respectively. If you need to name angles, use Greek letters, and use α, β and γ for the angles at vertices A, B and C, respectively. Use "s" for the semiperimeter, et cetera.

- Make certain it is easy to read labels, especially the important ones. The nice thing about **Geometer** diagrams is that you can usually adjust the diagram in such a way that the labels do not fall on top of each other and so that they are not obscured by other lines or circles in the diagram.

- Use appropriate fonts and point sizes for the illustration. In a classroom the text has to be big enough to be read from the back of the class. In a book, smaller fonts may help remove the clutter from a diagram.

- Show a general configuration if possible. In other words, if the theorem you wish to demonstrate applies to *all* triangles, do not display it with a triangle that is close to being equilateral or isosceles.

- Eliminate extraneous information. Hide your construction lines by making them invisible. Eliminate names from points that are not important to the presentation. Do not color things differently unless you have a good reason to do so.

- Do use colors, line stippling, angle markings, point shapes, et cetera when there is a good reason. For example, if you have a diagram where you construct a similar object based on each side of a triangle and it might be confusing what corresponds to which side, draw the three sides in different colors, and draw the objects associated with each side in the same color as the side. Similarly, if you are using inversion in a circle, you might draw objects in the original configuration in one color and in the inverted configuration in another.

- In locus-of-points illustrations, try to make the representative points regularly spaced.

11.2 Improving the Appearance of a Diagram

Many of **Geometer**'s primitives can be drawn in different styles. All can be drawn in different colors, but the other styles are primitive-dependent.

11.2.1 Colors

In addition to the standard colors that are available in the color selection box in **Geometer**'s control area there are 24 additional available colors that you can customize and use from **Geometer**'s text editor. The additional colors have names .c8, .c9, ..., .c31. (Colors .c0 through .c7 are the reserved colors: black, red, green, yellow, et cetera.)

In a **Geometer** file, the definition of the colors .c8 through .c31 can be set to a particular combination of red, green and blue components that vary from 0.0 to 1.0. To set the color .c8 to be pink and then to use it as the color of point p1, insert a line defining the color components for .c8 and edit the line for p1 so that it looks something like this:

```
.c8 = (1.000000, 0.400000, 0.400000);
p1 = .free(0, 0, .c8, "A");
```

The definition of color .c8 above will remain in effect only as long as that particular **Geometer** file is loaded.

11.2.2 Names

Every primitive can have a user-assigned name, but by default, names are only assigned to points as they are created, and this feature can be turned on and off in the *Edit Preferences* dialog. The name can be as simple as a single character or it can be a whole string. To create or change the name of a primitive, select it and then use the *Edit Name* command.

The names of points, lines and angles are drawn on the screen slightly above and to the right of a point, or at the center of a line segment, either slightly above or slightly below the segment depending upon whether the segment slopes up or down. The name of an angle is drawn inside the angle. It is a bit difficult to draw angle names in exactly the right place, so for high-quality drawings you will sometimes need to resort to other methods to place angle names properly.

Another primitive whose name is used by **Geometer** is the floating-point number (the .flt) that you can only have entered with the text editor. If a .flt has a name, the value of that number is drawn in the upper left of the display window, together with the name. For example, if the .flt named pi has the value 3.14159, it will be presented as the string: "pi = 3.142" in the upper left of the screen. All values are presented only to three decimal places, since **Geometer** cannot guarantee accuracy much better than that. If there is more than one named .flt, all the values are listed in a column.

If you wish to display the length of a line segment, the area of a polygon, or the measure of an angle, you can select that item and use the *Display Value* command. In that case the primitive's length, area, or measure is drawn in the same way in the upper left portion of the display window. If the line segment named b has length 1.234 it will be presented as "b = 1.234". In the current version of **Geometer**, the coordinate system runs from -1.0 to 1.0 in the shorter dimension of the viewing window so all lengths and areas will be presented based on these dimensions.

The *Display Value* command is easy to use—you simply need to select a primitive and toggle it on or off—but it is somewhat restricted by **Geometer**'s coordinate system. For high-quality presentations you may need to resort to working directly with the .flts. See Section 11.5.

Finally, the text you use for a primitive name can have certain special characters, superscripts, subscripts, and so on. See Section 11.2.4.

11.2.3   Primitive-Specific Styles

In addition to colors and names, points, lines, polygons and angles can be displayed in various modes. Points can be drawn as dots, small circles, diamonds, or as nothing at all. Lines can be drawn with various stippling options, hash marks (to indicate congruence to other segments), or as a segment, ray or full line. Different sorts of angle markings are available, different polygon fill types are available, and the line width of any line, circle, conic section or Bézier curve can be set from 0.5 to 6 pixels in width. On the computer screen a width of 0.5 pixels cannot be shown (so it is displayed as one pixel wide), but when printed it will produce a lighter line.

One of the main uses of a point with nothing displayed is for use as a label in a high-quality illustration. Although no mark is made for the point, its name is displayed, so it is basically a piece of text that can be moved to any position on the screen. If you would like to have it locked at a certain location, use a .pinned point instead of a .free point.

All these display styles for primitives are available in the *Style* pulldown menu or in the selection boxes in the control area. Rather than give a complete description of each here, it is easier just to draw some primitives with geometer, select them, and change the styles of each to look at the various options.

In classroom presentations, sometimes the thin width of the lines in **Geometer** makes them difficult to see across a classroom, especially if the image is projected on a television screen instead of a computer screen projector. You can make the lines in your diagram wider, of course, but then they will be inappropriate for close-up viewing. The command *Wide Lines* (available in the *Styles* pulldown) doubles the width of all the lines. The keyboard shortcut is **Ctrl-L** or **Ctrl-l** (lower-case "L") which toggles the wide lines on and off.

The only reason to use the polygon primitive in **Geometer** is to obtain its area automatically or to have an easy way to stipple or fill regions for display purposes.

11.2.4 Descriptive Text

The text that appears at the bottom of the display window is the same as the text that names primitives with the one exception that a primitive name must consist of a single line of text—no newline characters.

In addition to the standard characters, **Geometer** is capable of drawing a number of special characters, subscripts, superscripts, and even lines or arcs over text. The special characters are all introduced by a "\" character. If you would like to talk about $\triangle ABC$, enter the following text: "`\triangleABC`". The leading "\" tells **Geometer** to try to find a command name after it (and it finds "triangle" in this case). If it does find a match, the matching text is replaced by the corresponding symbol. See the reference manual for a complete set of special characters supported in this way.

Two other commonly-used text commands produce superscripts and subscripts. The idea is similar: To produce "A_1", type: "`A\sub{1}`"; to produce "B^3", type: "`B\sup{3}`". There are a few other special text commands that are rarely used, but are documented in the reference manual.

Finally, **Geometer** has a somewhat unusual method for indicating newlines within a piece of text in the textual description. If you type in your text and edit it using the dialog box available under the *Descriptive Text* command in the *Primitives* pulldown this is all handled automatically. If you are typing in text by hand using the editor, here is what it looks like.

Suppose you would like your textual description to read like this with two lines of text:

> Congruent Triangles:
> $\triangle ABC$, $\triangle DEF$, $\triangle GHI$

This is exactly what you would type into the **Geometer** text file to obtain that result:

```
.text("Congruent Triangles:
\triangleABC, \triangleDEF, \triangleGHI");
```

Note that the opening and closing quotation marks are on different lines which is very non-standard in computer languages. It is easier to type, however.

11.3 Building Geometer Proofs and Constructions

It takes a fair amount of work to build a beautiful proof or construction that is ideally suited for presentation in a classroom. Once it is done, however, it can be used over and over again. In this section we will build a few examples in a step by step manner.

In this section we assume that you know how to prove a result. You are simply trying to present that proof in the clearest possible way.

Perhaps the best way to use this section is to read the description of how a particular demonstration is built and then try to build it yourself using **Geometer**. If you have problems, read the section again, presumably with much more interest.

11.3.1 The Incenter

As a first example, let us build a proof of the fact that the three angle bisectors of a triangle meet at a single point that is the center of the inscribed circle (the point is also called the incenter) of the triangle.

Although we may have to use the text editor to put in the final touches, we would like to do as much as possible using **Geometer**'s standard graphical user interface. Usually the best way to construct a proof is to draw the whole thing at once, including all construction lines that you want to be visible during the proof. Then return to the individual steps of the proof, making lines, points, angles, et cetera, invisible if they are not needed in that step and possibly emphasizing them if they are important.

Figure 11.1 displays all the visible lines and points that we draw before beginning to construct the step by step proof. Refer to this diagram in the discussion below.

In the case of the incenter, we begin by drawing an arbitrary triangle $\triangle ABC$ where A, B and C are free points. Next, construct two angle bisectors at angles A and B using the *PPP=>P Bis* command to find a point that bisects an angle defined by three points. Since we know that the three bisectors meet at a point (remember, we are only trying to present a proof, not to find it), there is no need to construct the third bisecting point.
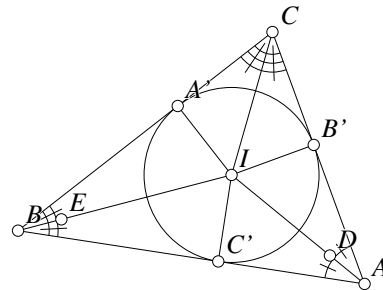
Figure 11.1: Three Incenters Proof
`Presenting/ThreeIns.T` [P]

Any two angle bisectors could have been used, but it is conventional in cases like that to use the angles with the names closest the beginning of the alphabet. In this simple case it is hard to make a mistake, but if you had somehow made a mistake and drawn an incorrect point and later erased it, your triangle might be $\triangle ABD$. Since it is again conventional to use a "logical" set of names, if this occurs, simply change the name of the point mislabeled "D" to be "C".

In complex diagrams, it may be easiest just to turn off the automatic point naming of **Geometer** and to fill in the names that you need later or as you go. In this case, for example, the points that bisect the angles will automatically be assigned the names D and E and we will have to remove those names later or make the points invisible. The command to toggle on and off the automatic assignment of names is *Point Names* under the *Edit* pulldown menu.

Draw the angle bisectors from A to D and from B to E and find their intersection which we label as "I" since we know it will be the incenter.

At this point, since we are surely not going to need to see points D and E in the final proof or the short segments AD and BE, it might be a good idea to "erase" them by changing them to have the invisible color[1].

---

[1]It is a minor point, but it is a good idea to make segments AD and BE invisible even though the lines

WE complete the drawing of the bisectors by drawing the segments AI, BI and CI.

Since we are going to want to indicate to the viewers that the lines AI, BI and CI are angle bisectors, we should also draw in the six angles $\angle CAI$, $\angle IAB$, $\angle ABI$, $\angle IBC$, $\angle BCI$ and $\angle ICA$ using the *PPP=>A* command. To indicate that there are three pairs of equal angles, use the different angle styles to make them different. The best way to do this is probably to use the angle styles with the slashes, since otherwise it will not be obvious whether you are talking about the single angle $\angle CAB$ or the two angles $\angle CAI$ and $\angle IAB$. See the figure.

To complete the drawing we will need to drop perpendicular line segments to the sides of $\triangle ABC$. Use the command *PL=>L Perp* to draw the three perpendiculars and then use the *LL=>P* three times to create points at the feet of these perpendiculars. We need points here since we will need to name those points in the textual presentation of the proof. It might be nice to give those points logical names, so we use the *Edit Name* command in the *Edit* pulldown menu to change the names to A', B' and C' as in the figure. The final touch is to draw the incircle centered at I and passing through A' (or B' or C'), or just draw the circle passing through the three points A', B' and C'.

Now that everything is drawn, let us convert our diagram to a proof. At this point, every primitive is either visible on all the layers or invisible on all the layers. The ones that are invisible at this point were made that way because they will never be shown, so we do not need to worry about them. We will, however, need to make the visible primitives change color and/or become invisible as we move from layer to layer in the presentation of the proof.

A **Geometer** proof works as follows. When you press the *Start* button, only items on layer zero are visible. Each time you press the *Next* button the visible layer advances by one so after one press you will view items on layer one, et cetera. If you want geometer to begin with only layer zero visible, you can add the line `.10;` to the file. This is commonly done on prepared proofs or constructions.

By default, the layer information is hidden from users, so if it is not visible you probably want to turn it on with the *Layer Display* command in the *Edit* pulldown menu. If you are going to produce a lot of proof or geometric construction demonstrations you can use the *Edit Preferences* entry in the *Edit* pulldown to have the layers displayed every time you start the **Geometer** program.

An easy way to build the proof at this point is to press the *Start* button (so that only layer zero can be viewed and edited) and then to make the diagram on that layer look exactly as you want it to look. When that step of the proof is complete, press the *Next* button to advance to the next step of the proof, and continue that way to the end. **Geometer** has 32 layers, which should be sufficient for almost any demonstration.

In this section we will describe in detail the production of the first two steps of the proof, but to complete it, you will have to perform the same general process on each

---

AI and BI will be drawn on top of them. This is because due to the fact that the lines are drawn as discrete pixels, the part of the line AI that is supposed to be identical to AD will almost certainly differ in a few pixels, so if AD is not made invisible, the line from A to I will be marginally wider during the part from A to D. Also, if the triangle is made small enough, points D or E could be pushed outside of it.

step. To view the final result, load the file `Presenting/ThreeIns.T` and press the *Next* button to step through the entire proof.

In a computer geometry program it is probably best in the first step of the proof to display the final result and explain it so that users can manipulate the figure to see exactly what is happening. A clear statement of the theorem should also be presented in that first step (on layer zero).

Save your diagram so far (as a sort of backup in case you make a terrible error) and then click on the *Start* button. Under *Layer Control* only box zero should be checked after doing this.

While viewing layer zero, select the points A', B' and C' and set them to the invisible color. Similarly, set the line segments connecting I to those three points also to be invisible. All the rest of the diagram should be visible since it demonstrates the theorem we are going to prove.

Next, enter the description that is to appear on layer zero. It should probably say something like this:

```
In \triangleABC the three internal angle bisectors
meet at a point I which is the center of the inscribed
circle--the circle inside the triangle that is
simultaneously tangent to all three sides.  Move
points A, B and C.
```

You can enter this description by selecting the *Descriptive Text* entry under the *Text* entry in the *Primitives* pulldown menu. If you prefer, you can enter it using the *Edit Geometery* command also in the *Edit* pulldown menu, but in this case you will have to include with it the "wrappings" that make it a text file command:

```
.text("In \triangleABC the three internal angle bisectors
meet at a point I which is the center of the inscribed
circle--the circle inside the triangle that is
simultaneously tangent to all three sides.  Move
points A, B and C.", .l0);
```

It is nice to include in the description an indication to the user of what parts of the diagram can be manipulated. In this case, for example, the user cannot move point I—only points A, B and C are free.

When this first step is completed, press *Next* to view and edit layer 1. On that layer, many of the constructions are invisible, so select each one that is supposed to be invisible and then select the invisible color. The keyboard shortcut that sets the selected primitive to be invisible is **Ctrl-i**. Since the first step of the proof concerns two angle bisectors, select those and set them to the blinking color, and since they are angle bisectors, it might also be a good idea to select the angles themselves and make them blink as well.

To complete this step, add text to describe this first step of the proof.

Continue in this way until all the steps are finished. In this presentation, it required 9 steps, ending with a view of layer 8. (Remember that the first step is on layer 0.)

It is now a good idea to go to the beginning (using the *Start* button) and step forward and backward through the demonstration, checking to see that all the primitives that were supposed to be invisible are, in fact, invisible. Correct any errors.

Next, edit the **Geometer** file with the *Edit Geometery* command in the *Edit* pull-down menu and add the line:

```
.l0;
```

at the beginning of the file so that when the diagram is loaded the viewer will see only the contents of layer 0 (as if the *Start* button were pressed automatically upon opening the file).

Finally, it is nice to add a line at the end of the file like this:

```
.text("Press 'Next' to continue ...", [8 .red, .in]);
```

This line appears after everything, but appears in red on the first 8 layers. Thus it will be displayed underneath the text you entered earlier on every step except the final one. It helps the user to know when the end of the proof is reached.

Now, load the diagram `Presenting/ThreeIns.T` and step through it, going forward and backward with the *Next* and *Prev* keys to see what a reasonably good proof looks like. You might also find it educational to view the code with the *Edit Geometry* command. Finally, it would be very useful to try to reproduce this file, or something similar, starting from scratch. You can then refer to `Presenting/ThreeIns.T` if you get stuck.

11.3.2 Squaring a Rectangle

We will finish this section with one more example, the presentation of a geometric construction: Given an arbitrary rectangle, construct a square that has the same area.

The same techniques and cautions that were mentioned in Section 11.3.1 apply here.

We first draw a **Geometer** diagram that does what we want and test it. Then we will begin with the entire diagram, complete with all the construction lines, and convert it to a step by step construction.

In this case, the solution is not too hard. If the dimensions of the rectangle are a and b, then the side of the desired square is $\sqrt{ab}$. If a circle has diameter $a + b$, then the distance to the circle along a line perpendicular to the diameter that splits it into lengths a and b is $\sqrt{ab}$.

The complete construction is illustrated in figure 11.2. The user can change the shape of rectangle $OPRQ$ by moving points Q or P. Those points are constrained to be on vertical and horizontal lines emanating from O. Since OP extends to the right, we can make a segment of length $OP + OQ$ by copying OQ down to the line OP with a circle centered at O passing through Q and S. Clearly $SP = OP + OQ$.

In the figure, T is the midpoint of SP, so a circle centered at T will have SP as diameter. A perpendicular to SP at O will intersect the circle with diameter SP at U.
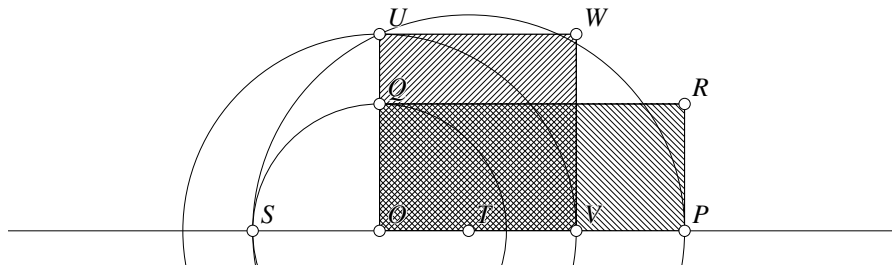
Figure 11.2: Rectangle to Square
`Presenting/RectSquare.T` [P]

The required square has a side equal to AU, and a circle centered at A passing through V will determine the two sides of the square. Finally, polygons with different colors and stippling patterns are draw around the rectangle and the square so the student can see their relative sizes and shapes.

Now we do much the same thing we did with the proof in Section 11.3.1. We press the *Start* button and modify the layers so that each presents the next step. As before, it is probably a good idea to show the adjustable rectangle and square in the initial layer so the user can experiment and see what the proposed construction is supposed to accomplish.

Probably the best way to display this first step is with only points P and Q visible, as well as the two rectangles. Everything else is invisible.

On the next step (layer 1), eliminate everything but the magenta rectangle with points O, P and Q visible. Also include the line OP and the circle centered at O of radius OQ intersecting the line at S.

In the next step we introduce the midpoint of SP and notice that it was named T in the original diagram. It would better perhaps to change the name to M to make it easier to remember that it is the midpoint. Again, add a textual description. Since viewers are convinced already that $OQ = OS$ we can probably make the circle centered at O passing through Q invisible.

Continue in this way, and when you are finished, as before, add commands with the text editor to display only layer 0 when the file is loaded and another text entry at the end that indicates to the user that pressing the *Next* key will present additional information.

As before, load the file `Presenting/RectSquare.T` and see how it looks and works. View the constructions with the text editor to see in what order the primitives were added. Make sure you understand how you could build it yourself.

11.4 Some Final Touches

Sometimes **Geometer** diagrams can be improved, either by making them less cluttered or more robust. The next sections show a few examples where this is done.

11.4.1 Avoiding Jumping Points

Sometimes a **Geometer** diagram can be made more robust by using one of the commands *LCP=>P Other* or *CCP=>P Other*, both available under the *Point* submenu of the *Primitives* pulldown menu.

A line and a circle or two circles generally intersect in two points or zero points. Thus, when you use the **Geometer** command to construct a point that is at one of those sorts of intersections, it solves the equations and finds two solutions. One of them is arbitrarily called solution 1 and the other, 2 (depending of the sign of the square root taken in the solution of a quadratic equation, if you know what that means). When the graphical user interface is used to identify an intersection, **Geometer** uses the solution closest to the final click of the mouse when the final circle is identified.

Unfortunately, this is only *guaranteed* to be correct at the time of creation. As you move free points later, the intersection may jump suddenly to the other possibility. Somewhat surprisingly, this occurs rarely, or when it does occur, it is masked because both intersections are used. But the two commands above can be useful for making this occurrence even more rare.

Quite often you have a point that is at one of the intersections and you would like to obtain the other intersection. The two commands above work exactly like

Figure 11.3: Angle Bisector Illustration
`Presenting/Bisect.T` [P]

LC=>P and *CC=>P* except that they require a third parameter—the intersection point you do *not* want to be chosen. When the more complex commands are executed, **Geometer** finds both intersections and chooses the one furthest from the point that is not to be chosen.

11.4.2 Using Arcs

Especially in constructions, the entire circle is not needed, and if entire circles are drawn, the figure can be quite cluttered. As an example, consider the construction of an angle bisector.

The construction is illustrated in figure 11.3. Given $\angle ABC$, construct a circle centered at B and find its intersection at D with ray BC. Draw the two circles centered at A and D that pass through B. They will have another intersection at E. (This is an ideal place to use the *CCP=>P Other* discussed in Section 11.4.1.) The line BE is the angle bisector of $\angle ABC$.

The figure is certainly correct, but it is a bit cluttered. Compare it with figure 11.4 where small arcs of circles are used instead of the entire circles.

An arc is specified somewhat like an angle. If three points A, B and C are given, point A is the beginning of the arc, point B is the center of the circle upon which the arc lies, and BC is the ray that marks the end of the arc. The arc goes counterclockwise from A to the ray BC.
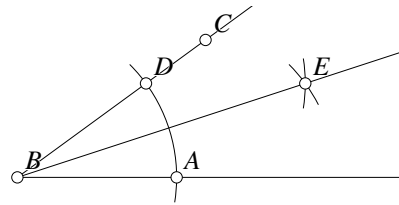
Figure 11.4: Better Angle Bisector Illustration

```
Presenting/Bisect.T [P]
```

Thus to specify an arc we need beginning and ending points for it that are invisible in figure 11.4. The way those were constructed is that two additional points were placed in the diagram a short distance apart (which could be adjusted later). Using these two points as a radius, small circles were drawn centered at points A, D and E. The intersections of these small circles with the larger circles in figure 11.3 were determined, and those points were used as the endpoints of the arcs.

If you cannot visualize this, look at the **Geometer** code for `Presenting/Bisect.T` or load the diagram and press the *Next* key twice. The first press shows the intermediate construction and the second shows the final state with the non-essential primitives set to the invisible color.

11.5   Calculation in Geometer

Geometer has a built-in calculator that works with floating point numbers. The source of the numbers can be from the distance between points, the measure of an angle, the coordinates of points, the area of a polygon, a current value of a `.script` command (see Section 11.7), or simply constants typed into a **Geometer** program.

These calculations can be useful directly for constructions, or for writing scripts that allow **Geometer** to display an entire animation at the press of a button. See Section 11.7.

Perhaps the most useful way to use floating point numbers is simply to display them on the screen. See Section 11.5.3.

11.5.1   Obtaining and Using Floating Point Numbers

Here is a list of commands to obtain floating point numbers for calculations:

- You can simply enter them as "12" or "-3.14159".

- `.f.vv`: Generate a number that is the distance between two points.

- `.f.area`: Find the (non-negative) area of a polygon and assign it to a number.

- `.f.vvvratio`: Find the number that represents the ratio of the distance between the first and second points to the distance between the first and third.

- `.f.vxcoord`: Assign the x-coordinate of a point to a number.

- `.f.vycoord`: Assign the y-coordinate of a point to a number.

- You can also use angles directly as floating point numbers. The angles are measured in degrees or radians, depending on the mode of **Geometer**. By default, **Geometer** is in degree mode.

Once you have some numbers to work with, you can do arbitrary calculations with the command `.f.rpn` (discussed below).

After you have done your calculations, you can use the results of your calculations as follows:

- `.v.ff`: Create a point whose coordinates are given by two floating point numbers.

- `.c.vf`: Create a circle with a point as center and number as radius.

- `.a.f`: Create an angle whose value is the floating point number. The angle's size is set depending on whether **Geometer** is in degree or radian mode.

- Floating point numbers can be used in almost all of the transformation commands. See Section 11.6.

11.5.2 Calculating with Floating Point Numbers

Finally, you need to know how to calculate. All calculations are done with a single function, `.f.rpn`. The "rpn" stands for "reverse Polish notation" which is a stack-based system of calculation that eliminates the need for parentheses and complex parsing.

In the case of **Geometer**, the stack is simply an initially empty pile of numbers, and each new number is conceptually piled on top of the others. You can only access the stack from the top—sort of like a pile of trays in a cafeteria. You can put trays on top or take them off the top, but it is difficult to access the trays below without first removing the ones on top.

The `.f.rpn` command is a list of numbers and operations that are evaluated from left to right. If it is a number or a variable that contains a number, the value of that number is simply put on the top of the stack. If it is an operation, that operation is applied to the top few elements of the stack and the result is usually put back on top of the stack. At the end of the evaluation, whatever number is left on top of the stack is the result of the calculation.

For example, the operation `.add` adds the top two stack elements, removing them from the stack, and the sum is put back on the stack. The following command assigns the value of 5 to `r`:

```
r = .f.rpn(2, 3, .add);
```

The first two items are numbers, so each is put on top of the stack. When it is time to evaluate the `.add`, the stack consists of a 3 on top of a 2. The `.add` command removes the top two numbers (all of them in this case), adds them to obtain 5, and then places the 5 on top of the stack. Since the `.add` is the final operation in this example, the value of 5 is assigned to r.

An application can be as simple as this:

```
p = .f.rpn(0.95);
```

which simply assigns the value 0.95 to the variable `r`.

It takes a bit of practice, but it is possible to evaluate extremely complex functions using this simple notation. For example, suppose we would like to evaluate:

$$r = \sin(x^2 - 3x)\cos(1/x).$$

Here is the statement that will accomplish it:

```
r = .f.rpn(x, x, .mul, 3, x, .mul,
    .sub, .sin, 1, x, .div, .cos, .mul);
```

The function puts on two copies of x, multiplies them, and returns the value of x^2 to the stack. Then a 3 and an x are added on top of the x^2. Those top two numbers are multiplied and replaced on the stack leaving an x^2 under a $3x$. Those two values are subtracted and the result is returned to the stack, leaving $x^2 - 3x$ on the stack. Next, the sine function is applied to the top stack element and the result is returned to the stack leaving a $\sin(x^2 - 3x)$. The process continues in this way until the entire expression is evaluated. **Geometer** understands about 30 operations and all are documented in the reference manual.

In the sections that follow, a few additional examples appear, but as one final example, here is some code to calculate the area of a triangle given the lengths of the sides using Heron's formula. If a, b, and c are the lengths of the sides of a triangle, then if $s = (a + b + c)/2$ is the semiperimeter, Heron's formula tells us that the area of the triangle is given by:

$$A = \sqrt{s(s-a)(s-b)(s-c)}.$$

Here is **Geometer** code to do the calculation, supposing that the points are internally named `v1`, `v2` and `v3`:

```
a = .f.vv(v1, v2);
b = .f.vv(v2, v3);
c = .f.vv(v3, v1);
s = .f.rpn(a, b, c, .add, .add, 2, .div);
A = .f.rpn(s, s, a, .sub, .mul, s, b, .sub,
    .mul, s, c, .sub, .mul, .sqrt);
```

11.5.3 Displaying Floating Point Numbers

If a floating point number has a name, it, together with its name, is displayed on the **Geometer** screen.

In the calculation in the previous section we did not display any of the information on the screen. If the purpose of this demonstration were to present a triangle that the user could manipulate and have the lengths of the sides as well as the area displayed, the following **Geometer** code would do the trick. This code can be found in the **Geometer** diagram `Presenting/Triarea.T`.

```
.geometry "version 0.60";
v1 = .free(-0.323353, 0.00898204, "A");
v2 = .free(0.161677, 0.461078, "B");
v3 = .free(0.338323, -0.161677, "C");
l1 = .l.vv(v1, v2);
l2 = .l.vv(v2, v3);
l3 = .l.vv(v3, v1);
c = .f.vv(v1, v2, "AB");
a = .f.vv(v2, v3, "BC");
b = .f.vv(v3, v1, "CA");
s = .f.rpn(a, b, c, .add, .add,
        2.000000, .div);
A = .f.rpn(s, s, a, .sub, .mul,
        s, b, .sub, .mul, s,
        c, .sub, .mul, .sqrt, "A(\triangleABC)");
```

This code would generate a continuously updated display in the **Geometer** window that might look something like this:

$AB = 0.882$
$BC = 0.547$
$CA = 0.683$
$A(\triangle ABC) = 0.187$

As we recall, the **Geometer** coordinate system goes roughly from -1.0 to 1.0 in both dimensions. Students can probably think more clearly about numbers that are in a more "friendly" range—say from 1 to 10. The diagram above can easily be modified so that the sizes of the legs of the triangle are roughly in this range by multiplying all the lengths of the sides by a constant before applying Heron's formula. If we want the sides to have maximum lengths on the order of 10 to 15, then a multiplication by 5 will do the trick. This code can be found in the **Geometer** diagram `Presenting/Triarea5.T`.

```
.geometry "version 0.60";
v1 = .free(-0.269461, 0.0299401, "A");
v2 = .free(-0.0898204, 0.601796, "B");
v3 = .free(0.682635, 0.39521, "C");
l1 = .l.vv(v1, v2);
l2 = .l.vv(v2, v3);
l3 = .l.vv(v3, v1);
cg = .f.vv(v1, v2);
ag = .f.vv(v2, v3);
bg = .f.vv(v3, v1);
```

```
c = .f.rpn(cg, 5.000000, .mul, "AB");
a = .f.rpn(ag, 5.000000, .mul, "BC");
b = .f.rpn(bg, 5.000000, .mul, "CA");
s = .f.rpn(a, b, c, .add, .add,
        2.000000, .div);
A = .f.rpn(s, s, a, .sub, .mul,
        s, b, .sub, .mul, s,
        c, .sub, .mul, .sqrt, "A(\triangleABC)");
```

In this example it might even be better to display a coordinate system to help the
students visualize the size of the triangle's legs. The coordinate system could be drawn
with .pinned points so the user could not move it. Of course if **Geometer**'s coordinates
are all multiplied by 5 as in the last example, so would they be in the coordinate system.
Here is some code that does just that found in `Presenting/Coordsys.T`.

```
.geometry "version 0.60";
v0 = .pinned(-0.6, 0, .plus, "-3");
v1 = .pinned(-0.4, 0, .plus, "-2");
v2 = .pinned(-0.2, 0, .plus, "-1");
v3 = .pinned(0, 0, .plus, "0");
v4 = .pinned(0.2, 0, .plus, "1");
v5 = .pinned(0.4, 0, .plus, "2");
v6 = .pinned(0.6, 0, .plus, "3");
v7 = .pinned(0, -0.6, .plus, "-3");
v8 = .pinned(0, -0.4, .plus, "-2");
v9 = .pinned(0, -0.2, .plus, "-1");
v10 = .pinned(0, 0, .plus, "0");
v11 = .pinned(0, 0.2, .plus, "1");
v12 = .pinned(0, 0.4, .plus, "2");
v13 = .pinned(0, 0.6, .plus, "3");
l1 = .l.vv(v0, v6, .longline);
l2 = .l.vv(v13, v7, .longline);
```

For a triangle area demonstration, it would be nice to combine the following code
with that in the example above. You cannot, however, just put the code together in one
file since there are name conflicts. There is a `v1`, `v2`, `v3`, `l1` and `l2` in both files. If you
have something like the coordinate system above that you would like to use in multiple
examples you have two options. First, use very unusual primitive names in the file that
you will insert into the other files, or second, use the **Geometer** *Insert* command. If
you load the file with the triangle area demonstration into **Geometer** and then use the
Insert command to add the `Presenting/Coordsys.T` file you will find that **Geometer**
modified all the names in the inserted file.

The **Geometer** commands look something like this after the insertion:

```
v0__b38f9 = .pinned(-0.6, 0, .plus, "-3");
v1__b38f9 = .pinned(-0.4, 0, .plus, "-2");
v2__b38f9 = .pinned(-0.2, 0, .plus, "-1");
...
l1__b38f9 = .l.vv(v0__b38f9, v6__b38f9, .longline);
l2__b38f9 = .l.vv(v13__b38f9, v7__b38f9, .longline);
```

11.6 | **Using Geometric Transformations**

Although they can certainly be used alone, the transformation primitives are useful in **Geometer** primarily for constructing scripts. If you want to have a point move across the screen or in a circle, or in a more complicated path, often the easiest way to do this is with transformations.

In this section we will examine transformations used alone, but once you understand how to do this and how simple scripts work, it is not too difficult to use the transformations in **Geometer** scripts. See Section 11.7.3 for a non-trivial application of transformations, and look at the source code for the circle demonstration in figure A.26 for a vastly more complicated example—so complicated, in fact, that a computer program was used to generate it.

A **Geometer** transformation is a rule that tells how to move a point. It is possible to use any linear transformation within **Geometer**, but the most common ones are translation, rotation, scaling and combinations of those.

Every transformation is simply a modification of another transformation with the exception of the identity transformation that does nothing. To build a **Geometer** transformation, you begin with the identity transformation and modify it until it does what you want.

We begin with a very simple example. It will contain one free point A, and two translations of A. See `Presenting/Trans1.T`. Here is the entire code:

```
v1 = .free(-0.227545, 0.110778, "A");
id = .x.identity();
t1 = .x.translate(id, 0.400000, 0.000000);
t2 = .x.translate(id, -0.200000, 0.300000);
v2 = .v.vx(v1, t1, "A+(.4, 0)");
v3 = .v.vx(v1, t2, "A+(-.2, .3)");
```

Two different translations are built from the identity. The transformation `t1` translates points by 0.4 units to the right and `t2` translates them 0.2 units to the left (which is the negative of right) and 0.3 units up. Point `v2` is the result of applying `t1` to `v1` and `v3` is the result of applying `t2` to `v1`. Load the diagram and move point A. The other two stay at those fixed distances from A.

As a second example, the file `Presenting/Trans2.T` rotates a triangle by $75°$ about the origin. Here is the code:

```
v0 = .pinned(0, 0, "O");
v1 = .free(0.311377, 0.11976, "A");
v2 = .free(0.58982, 0.41018, "B");
v3 = .free(0.610778, 0.0718563, "C");
l1 = .l.vv(v1, v2);
l2 = .l.vv(v2, v3);
l3 = .l.vv(v3, v1);
id = .x.identity();
rot = .x.rotate(id, 75.000000);
w1 = .v.vx(v1, rot, .green, "A'");
```

```
w2 = .v.vx(v2, rot, .green, "B'");
w3 = .v.vx(v3, rot, .green, "C'");
l4 = .l.vv(w1, w2, .green);
l5 = .l.vv(w2, w3, .green);
l6 = .l.vv(w3, w1, .green);
```

The file begins with a point O pinned at the origin $(0, 0)$. All rotation occurs about the origin, so the original triangle $\triangle ABC$ will be rotated about that point by $75°$. The next six lines of code that displays $\triangle ABC$ were drawn with **Geometer**'s graphical user interface. The next two lines build a $75°$ rotation beginning with the identity and combining a rotation with it. Finally, the points w1, w2 and w3 are rotated versions of the vertices of the original triangle. Load the diagram and modify points A, B and C to see how it works.

Finally, let us modify the example above to do a rotation about an arbitrary point P. This will be done by translating P to the origin, doing a rotation about the origin, and finally moving P back out to its original location. Some arithmetic must be done to achieve this.

Here is the complete code for this more complex transformation:

```
vp = .free(-0.0958084, -0.239521, "P");
v1 = .free(0.248503, 0.00598802, "A");
v2 = .free(0.667665, 0.293413, "B");
v3 = .free(0.610778, 0.0718563, "C");
l1 = .l.vv(v1, v2);
l2 = .l.vv(v2, v3);
l3 = .l.vv(v3, v1);
dx = .f.vxcoord(vp);
dy = .f.vycoord(vp);
dxneg = .f.rpn(dx, -1.000000, .mul);
dyneg = .f.rpn(dy, -1.000000, .mul);
id = .x.identity();
xform1 = .x.translate(id, dxneg, dyneg);
xform2 = .x.rotate(xform1, 75.000000);
rot = .x.translate(xform2, dx, dy);
w1 = .v.vx(v1, rot, .green, "A'");
w2 = .v.vx(v2, rot, .green, "B'");
w3 = .v.vx(v3, rot, .green, "C'");
l4 = .l.vv(w1, w2, .green);
l5 = .l.vv(w2, w3, .green);
l6 = .l.vv(w3, w1, .green);
```

This time the point P is the center of rotation and it is free so can be moved anywhere in the window. the variables dx and dy are the distances of the point P from the origin. Calculations are done to produce the negatives of these values: dxneg and dyneg. Next, after the identity transform is produced, it is successively modified to translate it to the origin, to do a $75°$ rotation, and then the translate back to the original point. The **Geometer** diagram is found in Presenting/Trans3.T. It seems similar to the previous example except that in this case, the center of rotation P can be moved.

11.7 Constructing Geometer Scripts

A script is based on a very simple idea. A floating point variable is successively set to different values at a uniform rate of speed. For example, a script could simply set the value of the variable t to values between 0.0 and 1.0 in steps of 0.01. This script would redraw the geometer diagram 101 times with t taking on the values $0.00, 0.01, 0.02, 0.03, \ldots, 0.99, 1.00$.

Assuming that the value of t is used in the construction of the diagram, each step in the script will draw a different version.

As a very simple example, here is a **Geometer** diagram that will move a point along the x-axis in 101 steps from $x = -1.0$ to $x = 1.0$. This example can be found in **Geometer** diagram `Presenting/Script1.T`.

```
t = .script(-1.000000, 1.000000, 0.020000);
v0 = .v.ff(t, 0.000000);
```

The first line defines the script variable t to begin at -1.0 and to end at 1.0 taking steps of 0.02 between each rendering of the scene. If you run this script by clicking on the *Run Script* in the command area, a point moves across the center of the drawing area in uniform steps.

An easy modification to the file is to change the second line to:

```
v0 = .v.ff(t, 0.000000, .smear, .dot);
```

This changes the style of the point to be drawn as a single dot, and to change the color to the smearing color. When this script is run, the point steps uniformly across the screen, but only dots are drawn, and they are not erased after each rendering, so a line of dots is drawn. This modification is in the file `Presenting/Script2.T`.

Another small modification (available in file `Presenting/Script3.T`) changes the diagram into something that will plot the curve $y = x^2$ for $-1.0 \le x \le 1.0$:

```
t = .script(-1.000000, 1.000000, 0.020000);
t2 = .f.rpn(t, t, .mul);
v0 = .v.ff(t, t2, .smear, .dot);
```

The second line calculates the value of t^2 by multiplying t by itself. The plotted point has coordinates (t, t^2) and is drawn in the smearing color.

That is all there is to it! A **Geometer** diagram can contain only a single script variable, and everything that changes has to be based on different values of that variable.

Obviously things are going to be much more interesting if we can somehow control some geometric shapes and positions with a script variable. In addition to the floating point calculator commands, the commands that tend to be immediately valuable with the script variable are:

- .v.vvf: Produces a point between two others a given ratio of the distance be-
  tween them. If the floating point value is zero, the new point lies on the first one.
  If it is one, it lies on the second. A value of 0.5 puts it midway between them.
  Values less than zero or greater than one also make sense.

- .v.ff: Produce a point with coordinates given by two floating point values.

- .c.vf: Produce a circle centered at a point with a radius given by the floating
  point value.

- .a.f: Produce an angle equal to the floating point value. If the diagram is in
  degree mode (the default), a value of 90 makes a right angle. If the diagram is in
  radian mode, a right angle corresponds to a floating point value of $\pi/2$.

- Any of the transformation commands. See Section 11.6.

11.7.1 The Witch of Agnesi

Perhaps the most common use of scripts in **Geometer** is to illustrate the solution to a
locus problem, of which the witch of Agnesi is one.
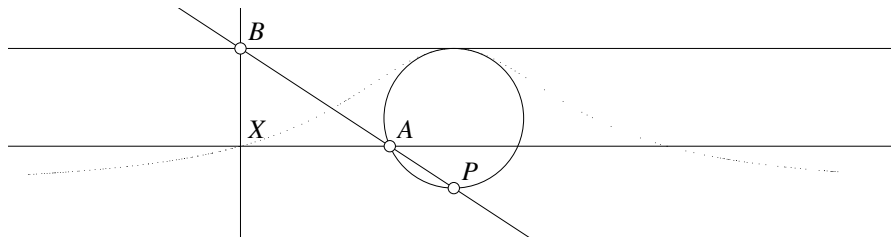
Figure 11.5: The Witch of Agnesi
`Presenting/Witch.T` [M]

Given a circle and a horizontal line l tangent to that circle, let P be the point on the
circle diametrically opposite the point of tangency with the line. Any secant though P
passes through the circle at a point A and through the line l at a point B. The witch of
Agnesi is the locus of points X that lie on the intersection of the vertical line through
B and the horizontal line through A. Figure 11.5 shows a portion of the curve.

Figure 11.5 is certainly enough to understand roughly what is going on, but if we
wish to make a nicer version with more evenly spaced points, we can do so with a
script.

Here is the code for `Presenting/Witch.T`:

```
.geometry "version 0.60";
v1 = .pinned(-1.11814, 0.649789, .in);
v2 = .pinned(1.09705, 0.649789, .in);
l1 = .l.vv(v1, v2, .longline);
v3 = .vonl(l1, 0, 0.649789, .in);
l2 = .l.vlperp(v3, l1, .in, .longline);
```

```
v4 = .vonl(l2, 0, 0.0442478, .in);
c1 = .c.vv(v4, v3);
v5 = .v.lcvother(l2, c1, v3, "P");
v6 = .vonl(l1, -1.84971, 0.649789, "B");
l3 = .l.vv(v5, v6, .longline);
v7 = .v.lcvother(l3, c1, v5, "A");
l4 = .l.vlperp(v6, l1, .longline);
l5 = .l.vlpar(v7, l1, .longline);
v8 = .v.ll(l4, l5, .smear, .dot, "X");
```

The point B is dragged along the line, and appropriate intersections are found that eventually result in the point X which moves along the witch.
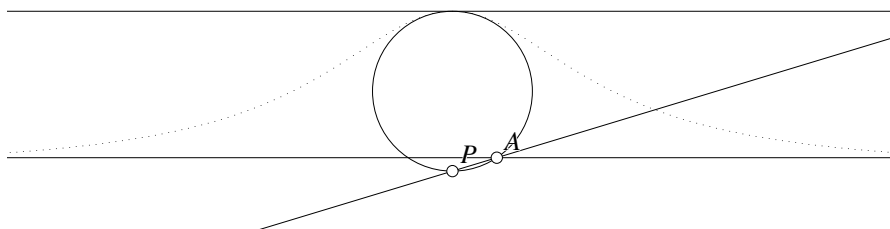
Figure 11.6: The Witch of Agnesi
`Presenting/ScriptWitch.T [S]`

To convert this to a script, all we need to do is make B move along the horizontal line uniformly. This could be accomplished by replacing the line:

```
v6 = .vonl(l1, -1.84971, 0.649789, "B");
```

with this:

```
t = .script(-4.0, 4.0, .05);
v6 = .v.ff(t, 0.649789, "B");
```

but this has the slight disadvantage that if the figure is scaled, only the y-coordinates of points will change; not the hard-coded value 0.649789 in the proposed solution.

Here is a better way: replace that same line defining point B with the following:

```
t = .script(-4.000000, 4.000000, 0.050000);
vx = .v.ff(t, 0.000000);
lx = .l.vlperp(vx, l1);
v6 = .v.ll(lx, l1, "B");
```

The difference is that a point `vx` is created that moves horizontally in a smooth manner. A perpendicular from `vx` is dropped to the line `l1` and the intersection of `l1` and that perpendicular line `lx` identify the point B. You can test this version which can be found in the file `Presenting/ScriptWitch.T`. The result is illustrated in figure 11.6.

11.7.2   Poncelet's Theorem Demonstration

In Section 3.8 we constructed a diagram to illustrate Poncelet's theorem for a quadri-lateral simultaneously tangent to an inner circle and inscribed in an outer one. In the demonstration in that chapter the user can move the point O, the center of the inner circle, and then leaving the point O in place move the point A around the outer circle and verify that for any starting position the quadrilateral remains tangent to the inner and inscribed in the outer circle.

A trivial modification changes that demonstration to a script that allows the user to adjust the position of O as before, but then the *Run Script* command drives the point A around the circle making two complete loops before it stops. The new code (in file `Presenting/Script4.T`) looks like this:

```
R = .f.rpn(0.950000);
O = .v.ff(0.000000, 0.000000, .in);
C = .c.vf(O, R);
V = .free(-0.125749, 0.191617, "O");
d = .f.vv(O, V);
r = .f.rpn(1.000000, R, d, .add, .dup,
        .mul, .div, 1.000000, R, d,
        .sub, .dup, .mul, .div, .add,
        1.000000, .exch, .div, .sqrt);
c = .c.vf(V, r);
//v1 = .vonc(C, 0.884919, 0.34557, "A");
t = .script(0.000000, 720.000000, 2.000000);
x = .f.rpn(t, .cos, R, .mul);
y = .f.rpn(t, .sin, R, .mul);
v1 = .v.ff(x, y, "A");
l1 = .l.vc(v1, c, 2);
v2 = .v.lcvother(l1, C, v1, "B");
l2 = .l.vc(v2, c, 2);
v3 = .v.lcvother(l2, C, v2, "C");
l3 = .l.vc(v3, c, 2);
v4 = .v.lcvother(l3, C, v3, "D");
l4 = .l.vv(v4, v1);
l5 = .l.vv(v1, v2);
l6 = .l.vv(v2, v3);
l7 = .l.vv(v3, v4);
```

The only modification from the old code is that the original line that defined `v1` as a point constrained to lie on a circle has been commented out, and replaced by the four lines that follow it. (The "//" at the beginning of the line containing the old definition of `v1` indicates that the rest of the line is commented out and is invisible to **Geometer**.)

The first of those lines defines a script that causes the value of `t` to vary from 0 to 720 (degrees). The following two lines calculate the location of the point on the circle by taking the sine and cosine of `t` and multiplying the results by the radius R. Finally, a new version of `v1` is created with `x` and `y` as its coordinates. All the rest of the code, before and after, is identical.
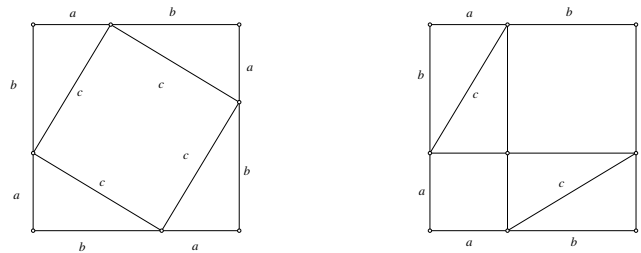
Figure 11.7: Proof 2 of the pythagorean theorem
`Review/Pyproof2.D, Review/Pyproof3.D [M]`

11.7.3 Script Plus Transformations

It is possible to construct animations with **Geometer** using the transformation commands discussed in Section 11.6. In this section we will construct a dynamic version of the pythagorean theorem demonstrated by the illustration shown in Section A.6.5, reproduced here as figure 11.7.

We would like to have the four triangles in the figure begin in one position and the *Run Script* button is pressed, to have them move to form the other configuration. A quick look at figure 11.7 shows us there is a problem: if we use the triangles as they appear in that figure, two of them will need to be flipped over during the transformation. In other words, in the diagram on the left, all the triangles are "right-handed" and in the diagram on the right, two are "right-handed" and two are "left-handed". If you cannot see this, try cutting out the triangles from small pieces of paper and see what happens if you try to slide them around. Eventually you will need to invert two of them.
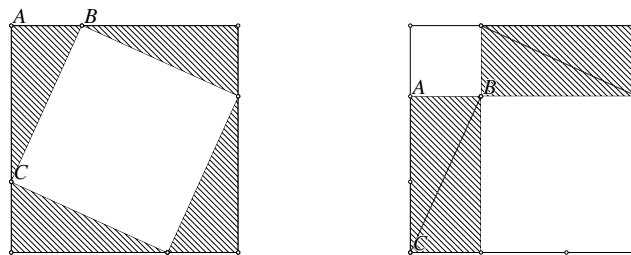
Figure 11.8: Dynamic Pythagorean Theorem
`Presenting/Script5.T [S]`

A better pair of initial and final configurations are shown in figure 11.8. In this section we will describe how to build a **Geometer** script that moves smoothly from the configuration on the left to that on the right. Notice that all four triangles in this diagram have the same handedness.

We would like to allow the user to adjust the shape of the triangle before running the script, but we know that the four triangles will initially have to be arranged around a square as in the diagram on the left of figure 11.8. The easiest way to do that is to fix the

square and to allow the point B to move along the edge. Because of the arrangement of the triangles, the location of B will determine the length AC as well, and all possible right triangle shapes can be illustrated with this single degree of freedom.

In the initial diagram it is important to analyze the motions that the four triangles must make. If we number the initial ones 1, 2, 3 and 4 beginning with the one labeled $\triangle ABC$ in the upper left and continuing clockwise, triangle 1 needs to be translated straight down by a distance equal to AB. Triangle 2 does not move, triangle 3 rotates counterclockwise for $90°$ about its uppermost tip, and triangle 4 rotates $90°$ clockwise as its uppermost point slides down to the lower left corner of the surrounding square.

The easiest way to build the script is one transformation at a time. Write each of the three transformations and test it before moving on to the next. Remember that one of the triangles does not need to move at all so there are three, rather than four transformations.

Here is the entire code for `Presenting/Script5.T`:

```
.geometry "version 0.60";
t = .script(0.000000, 1.000000, 0.010000);
v0 = .pinned(-0.85, 0.85);
v1 = .pinned(-0.85, -0.85);
v2 = .pinned(0.85, 0.85);
l1 = .l.vv(v0, v2);
l2 = .l.vv(v0, v1);
v3 = .vonl(l1, -0.320359, 0.85);
c3 = .c.ctrvv(v0, v2, v3, .in);
v5 = .v.lc(l2, c3, 2);
l4 = .l.vlpar(v1, l1, .in);
l5 = .l.vlpar(v2, l2, .in);
v6 = .v.ll(l5, l4);
l6 = .l.vv(v2, v6);
l7 = .l.vv(v6, v1);
c4 = .c.ctrvv(v1, v0, v5, .in);
c5 = .c.ctrvv(v6, v0, v5, .in);
v7 = .v.lc(l6, c5, 1);
v8 = .v.lc(l7, c4, 1);
ab = .f.vv(v0, v3);
xid = .x.identity();
dy = .f.rpn(t, ab, .mul, -1.000000, .mul);
xabc = .x.translate(xid, 0.000000, dy);
a1 = .v.vx(v0, xabc, "A");
a2 = .v.vx(v3, xabc, "B");
a3 = .v.vx(v5, xabc, "C");
poly1 = .polygon(3, a1, a2, a3, .red, .hashpoly);
ex = .f.vxcoord(v7);
ey = .f.vycoord(v7);
nex = .f.rpn(ex, -1.000000, .mul);
ney = .f.rpn(ey, -1.000000, .mul);
xe1 = .x.translate(xid, nex, ney);
xang = .f.rpn(t, -90.000000, .mul);
xe2 = .x.rotate(xe1, xang);
xe3 = .x.translate(xe2, ex, ey);
b1 = .v.vx(v7, xe3);
b2 = .v.vx(v6, xe3);
b3 = .v.vx(v8, xe3);
```

```
poly2 = .polygon(3, b1, b2, b3, .green, .hashpoly);
poly3 = .polygon(3, v2, v3, v7, .yellow, .hashpoly);
cx = .f.vxcoord(v5);
cy = .f.vycoord(v5);
ncx = .f.rpn(cx, -1.000000, .mul);
ncy = .f.rpn(cy, -1.000000, .mul);
xc1 = .x.translate(xid, ncx, ncy);
yang = .f.rpn(t, 90.000000, .mul);
xc2 = .x.rotate(xc1, yang);
xc3 = .x.translate(xc2, 0.000000, dy);
xc4 = .x.translate(xc3, cx, cy);
d1 = .v.vx(v1, xc4);
d2 = .v.vx(v5, xc4);
d3 = .v.vx(v8, xc4);
poly4 = .polygon(3, d1, d2, d3, .cyan, .hashpoly);
```

The first line defines a script variable `t` that varies from 0.0 to 1.0 in steps of 0.01. The 0.01 is not too important—we can adjust it later if the script seems to run too slow or too fast. The range 0.0 to 1.0 is arbitrary—we will have to transform it to move angles from $0°$ to $90°$ and to translation amounts that depend on the lengths of the sides of $\triangle ABC$.

The next few lines draw the surrounding square and define a point B that is constrained to move on the upper line of the square. Notice that the point constrained to be on the line does not have the name "B". This is because the real point B will be a transformation of that point. Initially the transformation will be zero, but although the point `v3` remains on the top edge of the enclosing square, its transformed image called B will move down as the script proceeds.

In fact, in the construction of the diagram, all the vertices defining the four triangles were visible and had readable names to make it easier to find the appropriate points to transform. When everything was working, the names were eliminated and certain constructions were made invisible.

The next lines are standard **Geometer** constructions of the rest of the vertices on the square whose positions vary based on the location of the controlling point `v3` (which looks like it is named B, but is not).

The script work begins with the line:

```
ab = .f.vv(v0, v3);
```

This line assigns the length of the segment AB to a variable called `ab`. Next we begin building a transformation that will slowly slide the triangle in the upper left downward. As all transformations must begin with an identity transform, that is what is done in the next line. The amount of movement `dy` is calculated in the next line, and it depends on the script variable `t`. As `t` varies from 0.0 to 1.0 `dy` varies from 0.0 to $-AB$. Next, `xabc` is produced by a translate command that moves objects by a distance `dy` in the y-direction and does not move them at all in the x-direction. Finally, transformation `xabc` is applied to the three initial vertices of the triangle in the upper left corner to produce the version with vertices labeled A, B and C that slides down the window.

A polygon is drawn connecting these three vertices so that the solid effect will be obvious.

Next, a similar procedure is written to rotate the triangle in the lower right for $90°$ about its upper point. We cannot simply use a rotation command, since the rotation transformation rotates about the origin. What we must do is translate the point to the origin, rotate it there, and then translate back to the original position. That is why the transformation xe3 is built in a series of steps: start with the identity, translate to the origin, do the rotation, and finally, translate back. When it is ready, the xe3 transformation is applied to the three vertices of the triangle in the lower right, and those transformed points b1, b2 and b3 are connected to form a polygon.

The final transformation is almost the same, but it is a combination of a rotation and a translation. Thus the point about which the rotation is to occur is translated to the origin, it is rotated and translated, and finally translated back to its original position, and the grand transformation that accomplishes all of this is called xc4. The original vertices are transformed by xc4 and are also connected to form a polygon.

11.7.4   Scripts with Conditionals

The example in this section is not very interesting, but it illustrates a technique that can be used in very interesting situations. Here we will construct a **Geometer** diagram which will move the center of a fixed-size circle around a square whose vertices are $(\pm 1/2, \pm 1/2)$. For an illustration of a much more interesting example that was constructed using similar techniques, look at Presenting/PythagProofFinal.T.

What we will do is to move the center first from $(-1/2, -1/2)$ uniformly to the point $(1/2, -1/2)$. Then we will move it from there to $(1/2, 1/2)$, then to $(-1/2, 1/2)$ and finally back to the start. If t varies from 0 to 1, a movement along the first path can be achieved by taking a vertex having coordinates $(-1/2 + t, 1/2)$. Unfortunately, when t reaches 1, a new formula is required that leaves the x-coordinate fixed and moves along the y direction, and so on.

A technique that will work well is to have a single script where the variable t runs from 0 to 4, and each time it advances by 1, a new segment of the square is traced.

Following is the code to achieve the desired result, and it can be found in the file Presenting/SquareMove.T:

```
.geometry "version 0.62";
t = .script(0.000000, 4.000000, 0.040000);
in1 = .f.rpn(t, 1.000000, .lt);
in2 = .f.rpn(t, 1.000000, .ge, t, 2.000000,
        .lt, .mul);
in3 = .f.rpn(t, 2.000000, .ge, t, 3.000000,
        .lt, .mul);
in4 = .f.rpn(t, 3.000000, .ge);
t1 = .f.rpn(t);
t2 = .f.rpn(t, 1.000000, .sub);
t3 = .f.rpn(t, 2.000000, .sub);
t4 = .f.rpn(t, 3.000000, .sub);
```

```
tx = .f.rpn(-0.500000, t1, .add, in1, .mul,
         0.500000, in2, .mul, .add, 0.500000,
         t3, .sub, in3, .mul, .add,
         -0.500000, in4, .mul, .add);
ty = .f.rpn(-0.500000, in1, .mul, -0.500000, t2,
         .add, in2, .mul, .add, 0.500000,
         in3, .mul, .add, 0.500000, t4,
         .sub, in4, .mul, .add);
v0 = .v.ff(tx, ty);
c0 = .c.vf(v0, 0.200000, .smear);
```
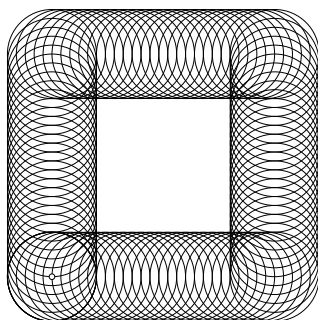

Figure 11.9: Circle Moving in a Square

The first line is the standard script running `t` from 0 to 4. The next four lines are essentially boolean values: 1 (true) if `t` is in the range and 0 (false) otherwise. Note that multiplication of zeroes and ones is the same as the boolan "and" operation. To see if `t` is between 1 and 2 we just multiply "$1 < t$" by "$t \le 2$", and so on.

The next four steps produce for convenience values of `t` that vary from zero to one in each of the ranges. Notice that we do not care about bad values outside the ranges, since we will be multiplying the bad values by zero anyway.

The somewhat complex calculations for the centers (t_x, t_x) of the circle effectively multiply the position of the point along each side of the square by a boolean value that is one only during the proper part of the path. Mathematically, it looks something like this, if p_i is only true (or equal to 1) on the correct part of the path. Here is a formula for the x-coordinate. t_i varies from 0 to 1 on the ith part of the path:

$$(-1/2 + t_1)p_1 + (1/2)p_2 + (1/2 - t_3)p_3 - (1/2)p_4.$$

A slightly different, but similar, calculation is done to find the y-coordinate, and the two are combined in the next-to-last step to identify the center of the circle that will follow a path along the square.

Figure 11.9 shows the result.

11.8 Scripts and Layer Conditionals

A nice touch in a diagram is to have the circles and lines be drawn not all at once, but rather as if they were being drawn by hand: an arc that expands to become a circle, or as a line segment that starts as a point and gradually lengthens to its final length.

Figure 11.10 shows the final result of a demonstration of the method to bisect a line segment that is not too interesting. What is interesting is to watch it being drawn by **Geometer**. As the script runs, both circles are drawn smoothly as is the line connecting points C and D, and the points C, D and the midpoint M appear one at a time.
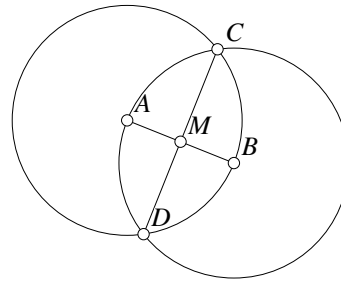
Figure 11.10: Smooth Drawing
`Presenting/SmoothArc.T` [S]

The script uses the `.layercondition` command which turns on a set of layers that depends on the value of the script variable, `t` in this case. In the example below, when `t` is between 0.0 and 1.0, layer 1 is displayed and so on. Using this trick it is simply a matter of drawing the correct figure or animation in each of the layers. Each animation is drawn in one layer, and then after that animation is complete, the finished figure appears in the following layers.

In this example, the original problem is presented on layer 0, the animation of the first circle appears on layer 1, and a finished version of the circle appears on layers 2 and beyond. There is no animation on some of the layers so the person viewing the script can think for a second or two about what was just done.

Also note that a diagram such as this can be viewed manually simply by repeatedly pressing the *Next* button.

Here is the code:

```
.geometry "version 0.64";
.l0;
t = .script(0.000000, 30.000000, 0.010000);
.layercondition(0.000000, t, 1.000000, 0x1);
.layercondition(1.000000, t, 2.000000, 0x2);
.layercondition(2.000000, t, 3.000000, 0x4);
.layercondition(3.000000, t, 4.000000, 0x8);
.layercondition(4.000000, t, 5.000000, 0x10);
.layercondition(5.000000, t, 6.000000, 0x20);
.layercondition(6.000000, t, 7.000000, 0x40);
.layercondition(7.000000, t, 8.000000, 0x80);
.layercondition(8.000000, t, 9.000000, 0x100);
.layercondition(9.000000, t, 10.000000, 0x200);
.layercondition(10.000000, t, 11.000000, 0x400);
v1 = .free(-0.11976, 0.161677, "A");
v2 = .free(0.194611, 0.0359281, "B");
l1 = .l.vv(v1, v2);
v1x = .f.vxcoord(v1);
v1y = .f.vycoord(v1);
```

```
v2x = .f.vxcoord(v2);
v2y = .f.vycoord(v2);
r12 = .f.vv(v1, v2);
theta1 = .f.rpn(v2x, v1x, .sub, v2y, v1y,
        .sub, .exch, .atan2);
x = .f.rpn(t, t, .floor, .sub, 360.000000,
        .mul, theta1, .add, .cos, r12,
        .mul, v1x, .add);
y = .f.rpn(t, t, .floor, .sub, 360.000000,
        .mul, theta1, .add, .sin, r12,
        .mul, v1y, .add);
v3 = .v.ff(x, y, .in);
arc1 = .arc.vvv(v2, v1, v3, .l1);
circ1 = .c.vv(v1, v2, .l2on);
theta2 = .f.rpn(v1x, v2x, .sub, v1y, v2y,
        .sub, .exch, .atan2);
x1 = .f.rpn(t, t, .floor, .sub, 360.000000,
        .mul, theta2, .add, .cos, r12,
        .mul, v2x, .add);
y1 = .f.rpn(t, t, .floor, .sub, 360.000000,
        .mul, theta2, .add, .sin, r12,
        .mul, v2y, .add);
v4 = .v.ff(x1, y1, .in);
arc2 = .arc.vvv(v1, v2, v4, .l3);
circ2 = .c.vv(v2, v1, .l4on);
p1 = .v.cc(circ1, circ2, 2, .l5on, "C");
p2 = .v.cc(circ1, circ2, 1, .l6on, "D");
frac = .f.rpn(t, t, .floor, .sub);
v5 = .v.vvf(p1, p2, frac, .in);
ltemp = .l.vv(p1, v5, .l7);
ll = .l.vv(p1, p2, .l8on);
mid = .v.vvmid(v1, v2, .l9on, "M");
.text("Construct the midpoint of the segment AB");
.text("Adjust the positions of A and B
and then press the 'Run Script' button.", .red, .l0);
```

11.9 Computer-Generated Diagrams

Sometimes complex diagrams are required with hundreds or even thousands of lines. When that occurs, it is inefficient to enter thousands of primitives using the **Geometer** program. **Geometer** diagrams, however, are simply text files that can be generated by any means, including computer programs.

11.9.1 The Koch Snowflake Curve

The first example generates the first few approximations of the Koch snowflake curve— one of the simplest examples of a fractal. The "curve" is illustrated in figure 11.11. The first iteration is simply a triangle. The succeeding steps can be considered to be successive modifications of the previous version.

To traverse the triangle, imagine following three straight-line paths. In each successive iteration, replace the movement along a path by a four-step path that follows

the original path for $1/3$ of its length, then turns to the right by $60°$ and continues for the same distance. Next, turn $120°$ to the left, move another $1/3$, and finally turn $60°$ to the right and follow the original curve to its end.

After each such modification, the curve is composed of straight-line segments; just four times as many of them. The first six stages are illustrated in figure 11.11. When the **Geometer** diagram is loaded, the triangle appears, but each time the *Next* button is pressed, the next stage of the curve is presented.

The **Geometer** diagram simply consists of computer-generated sets of short line segments that compose each stage, and those on each stage are presented on a new layer.
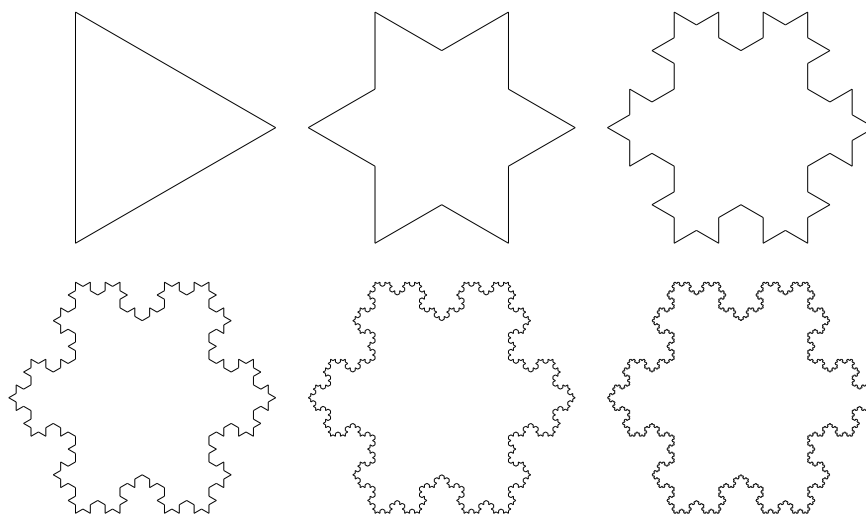
Figure 11.11: The Koch Snowflake Curve
`Presenting/Script5.T [S]`

The computer code (written in the C language) appears below and on the CD as `koch.c`:

```
#include <math.h>

#define PI 3.1415926535897

int vnum;
int glevel;

plot(double x, double y, double angle, double dist, int n)
{
  double x1, y1;
  if (n == 0) {
    x1 = x + dist*cos(angle);
    y1 = y + dist*sin(angle);
    printf("v%dx%d = .v.ff(%f, %f, .nomark, .l%d);\n", glevel,
```

```
            vnum, x1, y1, glevel);
    printf("l%dx%d = .l.vv(v%dx%d, v%dx%d, .l%d);\n", glevel, vnum, glevel,
            vnum-1, glevel, vnum, glevel);
    vnum++;
    return;
  }
  plot(x, y, angle, dist/3, n-1);
  x = x + dist*cos(angle)/3.0;
  y = y + dist*sin(angle)/3.0;
  angle -= PI/3;
  plot(x, y, angle, dist/3, n-1);
  x = x + dist*cos(angle)/3.0;
  y = y + dist*sin(angle)/3.0;
  angle += 2*PI/3;
  plot(x, y, angle, dist/3, n-1);
  x = x + dist*cos(angle)/3.0;
  y = y + dist*sin(angle)/3.0;
  angle -= PI/3;
  plot(x, y, angle, dist/3, n-1);
  return;
}

koch(int lev)
{
  printf("v%dx0 = .v.ff(1.0, 0.0, .nomark, .l%d);\n", glevel, glevel);
  plot(1.0, 0.0, 5*PI/6, 1.73205080757, lev);
  plot(-0.5, .866025403785, 3*PI/2, 1.73205080757, lev);
  plot(-0.5, -.866025403785, PI/6, 1.73205080757, lev);
}

main()
{
  int i;
  printf(".l0;\n");
  printf(".text(\"The Koch (snowflake) curve\n\");");
  printf(".text(\"Press 'n' to continue ...\", .red, .tol3);\n");

  for (i = 0; i < 5; i++) {
    vnum = 1;
    glevel = i;
    koch(i);
  }
}
```

Most of the work is done in the `plot` subroutine which is recursive and calls itself for the correct number of levels, depending on the approximation to the Koch curve it draws. If the level is zero, it simply draws a straight line of the given length in the given direction and returns. If the level is not zero, it calls itself four times to make the four subpaths, but with a level that is one less. Note that when each segment of the curve is drawn, it is drawn on the same layer controlled by the variable `glevel`.

To draw the Koch curve, the `plot` routine is called three times by the `koch` routine that draws paths that connect the three vertices of the original triangle.

Note: The actual Koch curve is none of the iterations above—it is the limiting case as the number of iterations increases without bound. It is interesting in that at each

stage after the first one, the area increases by $2/9$ of the amount added previously, but the perimeter increases by a factor of $4/3$. Thus the final, limiting Koch curve has finite area (if the original triangle has area 1, then the area enclosed by the limiting curve is $10/7 = 1.42857\ldots$) and infinite perimeter.

11.9.2 Apollonian Circle Packing

The next example is more complicated and it generates figure 11.12.
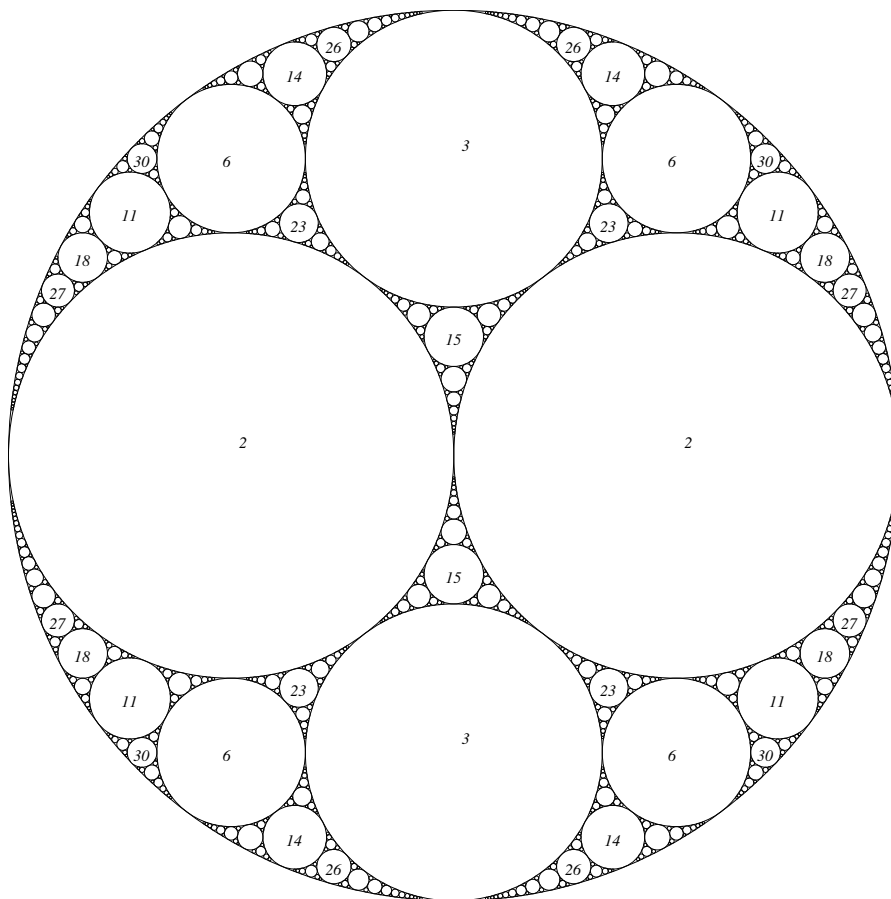
Figure 11.12: Apollonian Packing
`Presenting/soddy.T` [S]

The figure consists of a circle with two half-circles inside it. Beginning with those three mutually-tangent circles, each set of three mutually-tangent circles will have circles tangent to those, and so on. If the "holes" are filled in forever, the result is called the **Apollonian gasket**.

The radii of successive circles can be calculated using the Descartes circle theorem (see Section 2.14) and almost miraculously, if the radius of the outer circle is 1, then all the radii of the inner circles are of the form $1/n$, where n is some integer. In figure 11.12 some of the larger circles have numbers inside them and those numbers are the inverses of their radii. (For example, the circle with an 11 inside has radius $1/11 = .0909\ldots$.) What is more, every circle with radius $1/n$ is centered at a point with rational coordinates whose denominator is a divisor of n. With this information, it is not too hard to write a computer program that generates the figure.

Here is the listing of such a program (called `soddy.c` on the CD) written in the C language:

```c
#include <math.h>
#include <stdlib.h>

#define RMAX 400
#define RPRINTMAX 30

// for point: center = (x1/x2, y1/y2); radius = 1/R

int vertnum = 5;

struct point {
  int x1, x2, y1, y2, R;
};

int rightdist(struct point *p, int i, int j, int Rn)
{
  int x1 = p->x1, x2 = p->x2, y1 = p->y1, y2 = p->y2, R = p->R;

  if (R!=1 && (y2*y2*R*R*(x1*Rn-i*x2)*(x1*Rn-i*x2) +
      x2*x2*R*R*(y1*Rn-j*y2)*(y1*Rn-j*y2) ==
               x2*x2*y2*y2*(Rn+R)*(Rn+R))) return 1;
  if (R==1 && (i*i+j*j == (Rn-1)*(Rn-1))) return 1;
  return 0;
}

newcirc(struct point *p1, struct point *p2, struct point *p3, int depth)
{
  int R1, R2, R3, Rnew;
  int B, C, D, R;
  int i, j;
  struct point *p;
  double Rad, Cx, Cy;
  int imin, imax, jmin, jmax;

  if (depth == 0) return;
  R1 = p1->R;
  R2 = p2->R;
  R3 = p3->R;
  if (R1 > R2) { R = R1; R1 = R2; R2 = R; p = p1; p1 = p2; p2 = p;}
  if (R2 > R3) { R = R2; R2 = R3; R3 = R; p = p2; p2 = p3; p3 = p;}
  if (R1 > R2) { R = R1; R1 = R2; R2 = R; p = p1; p1 = p2; p2 = p;}

  if (R1 == 1) R1 = -1;
```

```
  B = -2*(R1 + R2 + R3);
  C = R1*R1 + R2*R2 + R3*R3 - 2*(R1*R2 + R2*R3 + R3*R1);

  // now solve x^2 + Bx + C = 0;

  D = B*B - 4*C;
  R = (int)sqrt(D + .01);
  Rnew = (R - B)/2;
  if (Rnew < 0) Rnew = (-R- B)/2;
  //printf("%d\n", Rnew);
  Rad = 1.0/Rnew + 1.0/R3;
  Cx = (double)(p3->x1)/p3->x2;
  Cy = (double)(p3->y1)/p3->y2;
  imin = (int)(Rnew*(Cx-Rad)-1);
  imax = (int)(Rnew*(Cx+Rad)+1);
  jmin = (int)(Rnew*(Cy-Rad)-1);
  jmax = (int)(Rnew*(Cy+Rad)+1);

  p = 0;
  for (i = imin; i <= imax; i++)
    for (j = jmin; j <= jmax; j++) {
      if (rightdist(p1, i, j, Rnew) &&
          rightdist(p2, i, j, Rnew) &&
          rightdist(p3, i, j, Rnew)) {
        p = (struct point *)malloc(sizeof(struct point));
        p->x1 = i; p->x2 = Rnew;
        p->y1 = j; p->y2 = Rnew;
        p->R = Rnew;
        if (Rnew < RMAX) {
          newcirc(p1, p2, p, depth+1);
          newcirc(p1, p, p3, depth+1);
          newcirc(p, p2, p3, depth+1);
        }
      }
    }
  if (Rnew <= RPRINTMAX) {
    printf("v%d = .pinned(%f, %f, .in);\n",
           vertnum, (double)p->x1/Rnew, (double)p->y1/Rnew);
    printf("vn%d = .v.vtranslate(v%d, -.037, -.037,"
           ".nomark, \"%d\", .l%don);\n",
           vertnum, vertnum, Rnew, depth);
    vertnum++;
  }
  if (Rnew < RMAX) {
    printf("v%d = .pinned(%f, %f, .in);\n",
           vertnum, (double)p->x1/Rnew, (double)p->y1/Rnew);
    printf("c%d = .c.vf(v%d, %f, .l%don);\n", vertnum,
           vertnum, 1.0/Rnew, depth);
    vertnum++;
  }
}

struct point O1, O2, O22, O3, O33;

main()
{
  int n = 4;
```

```
    O1.x1 = O1.y1 = 0; O1.x2 = O1.y2 = 1; O1.R = 1;
    O2.x1 = 1; O2.x2 = 2; O2.y1 = 0; O2.y2 = 1; O2.R = 2;
    O22.x1 = -1; O22.x2 = 2; O22.y1 = 0; O22.y2 = 1; O22.R = 2;
    O3.x1 = 0; O3.x2 = 1; O3.y1 = 2; O3.y2 = 3; O3.R = 3;
    O33.x1 = 0; O33.x2 = 1; O33.y1 = -2; O33.y2 = 3; O33.R = 3;

    printf(".l0;\n");
    printf("v0 = .pinned(0,0, .in);\nc0 = .c.vf(v0, 1.0);\n");
    printf("v1 = .pinned(0.5,0, .nomark, \"2\", .l1on);\n"
           "c1 = .c.vf(v1, 0.5, .l1on);\n");
    printf("v2 = .pinned(-0.5,0, .nomark, \"2\", .l1on);\n"
           "c2 = .c.vf(v2, 0.5, .l1on);\n");
    printf("v3 = .pinned(0,.66666666, .nomark, \"3\", .l2on);"
           "\nc3 = .c.vf(v3, 0.3333333333, .l2on);\n");
    printf("v4 = .pinned(0,-.666666666, .nomark, \"3\", .l2on);"
           "\nc4 = .c.vf(v4, 0.3333333333, .l2on);\n");
    newcirc(&O2, &O22, &O3, 3);
    newcirc(&O2, &O22, &O33, 3);
    newcirc(&O1, &O2, &O3, 3);
    newcirc(&O1, &O22, &O3, 3);
    newcirc(&O1, &O2, &O33, 3);
    newcirc(&O1, &O22, &O33, 3);
}
```

Each circle has a point at its center rational coordinates, so the x- and y-coordinates of each center point are listed with two integers: the numerator and denominator. A radius (actually the inverse of the radius) is also part of the data structure.

Rather than calculate the coordinates of the center which requires some fairly messy calculations, the program simply tries a small grid of possible centers. The routine rightdist returns 1 if the given center point is the proper distance from a proposed pair of rational coordinates. All calculations are done in integers, so exact results are obtained in every case.

The routine newcirc generates the circle that lies in the gap between three circles, and in addition, a parameter depth indicates how many levels down that circle is generated. It works by calculating the new radius and then checking that the center is the appropriate distance from all three given circle centers. Finally, it calls itself recursively to create three more circles in the newly-created gaps. The recursion halts when the circles get too small. In addition, numeric labels are printed on the diagram if the circles are somewhat larger.

The main routine simply draws enough circles to get the recursion started (the first five are necessary), and then newcirc is called on each triple of the initial circles. As the circles are printed in the form of **Geometer** commands, the layer upon which each appears depends on the depth at which that circle is generated. Thus the initial drawing is only a large circle, but each time the *Next* command is issued, another level of circles is drawn. In the sample code above, all circles of radius greater than $1/400$ of the original radius are drawn.
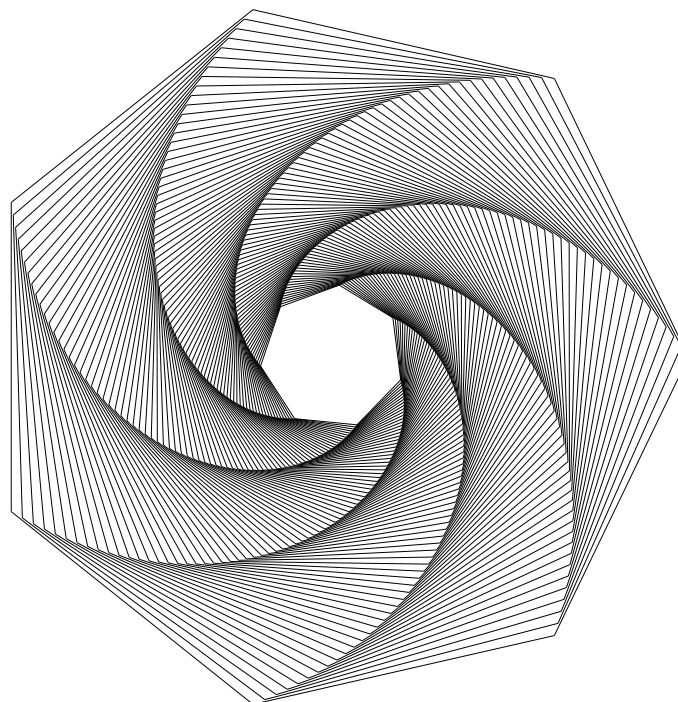
Geometer Proofs

This chapter contains a list of interesting theorems whose proofs are available on the CD as **Geometer** files. Some of the theorems are those usually presented in high school classes and could be used for classroom presentations by teachers. Others are proofs of more complicated and generally more interesting theorems. In the text, each theorem will be stated with an accompanying diagram. To see the proof, load the diagram into **Geometer** and use the *Next* and *Prev* buttons to move forward and backward through the steps.

To save space, unlike the rest of the book the diagrams here are untitled, and the name of the **Geometer** file is listed with the problem description, not with the figure.

The more interesting and beautiful theorems are listed with an accompanying diagram; at the end of the chapter is a long list of problems without diagrams but for which **Geometer** diagrams and proofs appear on the CD.

12.1 Illustrated Theorems

12.1.1 Adams' Circle
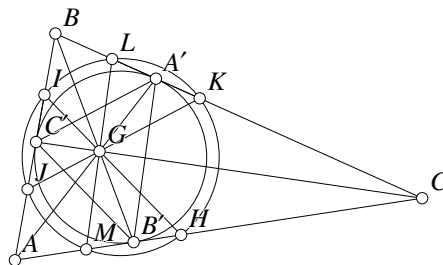
If $\triangle ABC$ is any triangle, let A', B' and C' be the points of tangency of the in-circle with the sides of the triangle. Let G (The Gergonne point) be the intersection of AA', BB' and CC'. The lines through G parallel to the sides of the triangle $\triangle A'B'C'$ intersect $\triangle ABC$ at six concentric points on Adams' Circle.

File: `GeomProofs/Adams.T[P]`

12.1.2 Rectangular Circle Centers
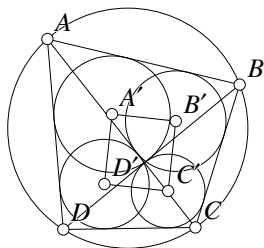
If $ABCD$ is a cyclic quadrilateral, show that the incenters of the triangles $\triangle ABC$, $\triangle BCD$, $\triangle CDA$ and $\triangle DAB$ form a rectangle.

File: `GeomProofs/Circlecenters.T[P]`

12.1.3 Fagnano's Problem
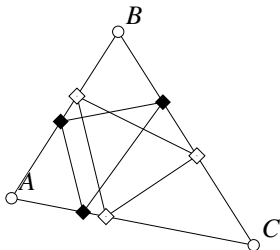
In any acute-angled triangle $\triangle ABC$, show that if three points are selected inside the segments AB, BC and CA the perimeter of that triangle is greater than or equal to the perimeter of the triangle formed when those points are the feet of the altitudes. This minimizing triangle is called the **orthic triangle**. In the figure the orthic triangle is drawn with solid black vertices.

File: `GeomProofs/Fagnano.T[P]`

12.1.4 The Euler Line
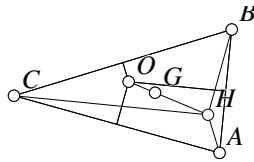
In any triangle $\triangle ABC$, show that the center of mass (G—the intersection of the three medians), the circumcenter (O—the center of the circumscribed circle) and the orthocenter (H—the intersection of the altitudes) all lie on a line called the **Euler line**.

File: `GeomProofs/Euler.T[P]`

File: `GeomProofs/Euler1.T[P]`

12.1.5 Similar Exterior Triangles
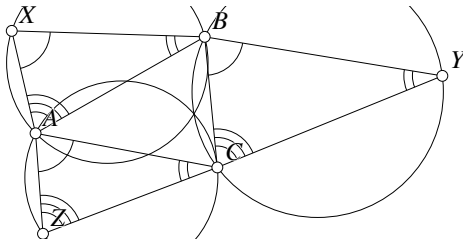
If three similar triangles are drawn exterior to the sides of $\triangle ABC$ as shown in the diagram on the left, show that the circumcircles of all three similar triangles meet at a point.

File: `GeomProofs/Simcircle.T[P]`

12.1.6 The Fermat Point
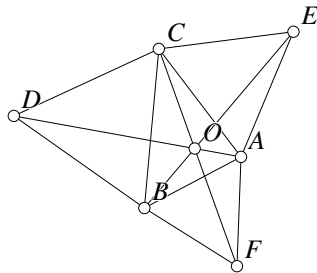
Erect outward-pointing equilateral triangles on all three sides of $\triangle ABC$. The lines connecting the vertices of $\triangle ABC$ with the tips of the equilateral triangles erected on the opposite sides meet at a point called the **Fermat point**.

File: `GeomProofs/Fermat.T[P]`

12.1.7 Four Bisectors
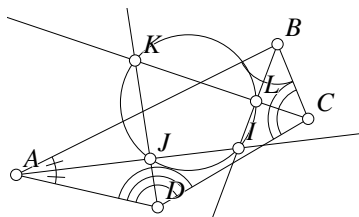
Let $ABCD$ be an arbitrary convex quadri-
lateral. If the interior angle bisectors are
constructed at A, B, C and D show that the
intersections of adjacent pairs of bisectors
are concyclic (all lie on the same circle).

File: `GeomProofs/Fourbisectors.T[P]`

12.1.8 Incenter Problem
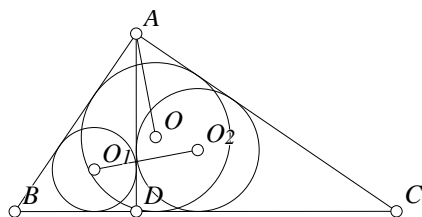
If $\angle BAC = 90°$ in $\triangle ABC$, and AD is
the altitude from A to BC, show that the
line connecting the centers of the incircles
of $\triangle ADB$ and $\triangle ADC$ is perpendicular to
the segment connecting A to the incenter of
$\triangle ABC$, and has the same length.

File: `GeomProofs/IncenterDist.T[P]`

12.1.9 Three Squares
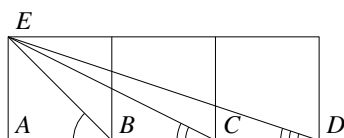
Show that if three squares are arranged
as in the diagram then:

$$\angle EBA + \angle ECB + \angle EDC = 90°.$$

File: `GeomProofs/Squares.T[P]`

12.1.10 Cyclic Quadrilateral Angle Bisectors
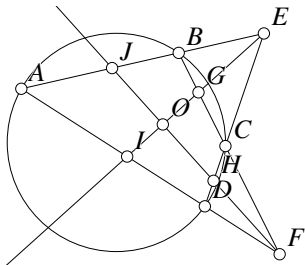
If $ABCD$ is a cyclic quadrilateral and sides AB and CD are extended to meet at E and BC and AD are extended to meet at F, show that the angle bisectors of $\angle AED$ and $\angle BFA$ are perpendicular.

File: `GeomProofs/Quadanglebisect.T[P]`

12.1.11 Japanese Temple Problem
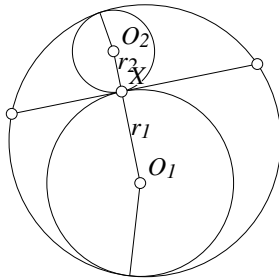
Given a circle and a chord through that circle, select a point X on the chord. There are two circles that are mutually tangent to the chord at X and also tangent to the larger circle. Show that the ratio of the radii of those two circles is independent of the location of X. What if X is on the chord but outside the circle?

File: `GeomProofs/Temple.T[P]`

12.1.12 Von Aubel's Theorem
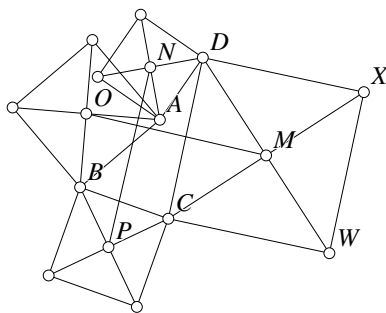
Given a quadrilateral $ABCD$, if squares are erected on the exterior of each side, show that the line segments connecting the centers of the squares erected on opposite sides have the same length and are perpendicular.

File: `GeomProofs/Vonaubel.T[P]`

$\boxed{12.2}$ **Proof Exercises**

This section contains a collection of problems, some easy, and some quite difficult. Most of them are more difficult than problems found in standard high school geometry textbooks. More difficult problems are preceded by the warning symbols: ♦, or ♦♦. The solutions appear in the **Geometer** files corresponding to the code in front of each file like E15 or C21. For example, to find the solution to problem E15, view the file E15.T in the GeomProofs directory.

1. **[E1]** ♦ If four straight lines emanate from a point, and if a line that intersects all four lines is divided harmonically by the points of intersection, then any line that intersects all four lines will also be divided harmonically by its points of intersection.

2. **[E2]** If a line is tangent to a circle at point P and parallel to a chord $\underline{AB}$ of that circle, then the point of tangency P will be the midpoint of the arc $\overgroup{AB}$ cut off by the chord.

3. **[E3]** ♦ Given two points A and B outside a circle K, show that if point C lies on the circle, then the sum $AC + BC$ is minimized when AC and BC make equal angles with the tangent to K at C.

4. **[E4]** If circle K has center O, and circle K_1 has diameter OR, where R is on K, and let L be a line from R that cuts both K and K_1 at P and P_1, respectively. Show that $RP/2 = RP_1$.

5. **[E5]** If two circles (possibly of different sizes) cut each other at two points, show that the line segment joining their points of intersection is bisected by the line joining their two centers.

6. **[E6]** If K and K_1 are two concentric circles with K_1 lying inside K, and if a line L is tangent to K_1 at a point P, and L cuts K at points A and B, then $AP = BP$.

7. **[E7]** Suppose that K_1 and K_2 are two circles that do not intersect, and neither lies inside the other. The two common internal tangent lines and the two common external tangent lines meet in four points. Show that these four points all lie on a circle.

8. **[E8]** Let K_1 and K_2 be two circles which are tangent internally at a point P. Let L be the common tangent line to the two circles at P. Let O be any point on L, and let K be a circle centered at O that cuts both K_1 and K_2. Show that the lines connecting O to the points of intersection of K and K_1 and K and K_2 cut off chords of K_1 and K_2 that are all of equal length.

9. **[E9]** Given a circle K centered at O, construct the two lines tangent to K at the endpoints of some diameter. Let P be any point of K other than one of the endpoints of that diameter. The line tangent to K at P intersects the other two tangent lines at A and B. Show that $OP \cdot OP = PA \cdot PB$.

10. **[E10]** If two circles K_1 and K_2 having diameters d_1 and d_2 are externally tangent, and L is a line that is externally tangent to both K_1 and K_2 at points P_1 and P_2, then $P_1P_2 \cdot P_1P_2 = d_1 \cdot d_2$.

11. **[E11]** Show that any side of a triangle is longer than the difference of the lengths of the other two sides.

12. **[E12]** Show that in any right triangle, the line segment connecting the right angle to the midpoint of the hypotenuse is equal to half the hypotenuse.

13. **[E13]** If P is any point in an equilateral triangle, and perpendiculars are dropped from P to each of the three sides, intersecting the sides at points A, B, and C, show that $PA + PB + PC$ is equal to the altitude of the triangle.

14. **[E14]** ♦ Let $\triangle ABC$ be an isosceles triangle with $\angle ABC = \angle ACB$. Let D and E be points on AB and AC such that $DB = EB$. Let P be the intersection of lines BC and DE. Show that $\angle CEB = 2\angle DPB$.

15. **[E15]** Show that for any $\triangle ABC$, let P be the point that is the intersection of the exterior angle bisectors at B and C. Show that PA bisects $\angle BAC$.

16. **[E16]** Show that if the opposite angles of a quadrilateral are equal, then the quadrilateral is a parallelogram.

17. **[E17]** Show that a line connecting opposite vertices of a parallelogram divides the parallelogram into congruent triangles.

18. **[E18]** Show that if $ABCD$ is a parallelogram, and P is any point interior to $ABCD$, then $\mathcal{A}(\triangle APB) + \mathcal{A}(\triangle DCP) = \mathcal{A}(\triangle BPC) + \mathcal{A}(\triangle DPA)$.

19. **[E19]** Let $ABCD$ be any convex quadrilateral, and let P be a point interior to $ABCD$. Show that $AP + BP + CP + DP \geq AC + BD$.

20. **[E20]** If the sides of a regular pentagon are extended to form a five-pointed star, show that the sum of the angles formed by the five tips of the star add to two right angles. What if the pentagon is not regular?

21. **[E21]** Let $\triangle ABC$ be a right triangle with the angle at A being the right angle. Let P be any point on the segment AB. Show that $CP^2 + AB^2 = BC^2 + AP^2$.

22. **[E22]** Let M be the centroid (the point where the three medians intersect) of $\triangle ABC$. Show that $AM^2 + BM^2 + CM^2 = (AB^2 + BC^2 + CA^2)/3$.

23. **[E23]** Let $ABCD$ be a rectangle, and let P be any point in the interior of $ABCD$. Show that $PA^2 + PC^2 = PB^2 + PD^2$.

24. **[E24]** Given any $\triangle ABC$, let P be an arbitrary point inside the triangle, and construct lines through P parallel to all three sides of $\triangle ABC$. If the line parallel to AB intersects BC at point E and AC at point I, the line parallel to BC intersects AB at J and AC at F, and if the line parallel to AC intersects AB at G and BC at H, show that $PE \cdot PF \cdot PG = PH \cdot PI \cdot PJ$. Is it necessary for P to be inside $\triangle ABC$ for this to hold?

25. **[E25]** ♦ If $\triangle ABC$ is a right triangle with right angle at A, and d is the diameter of the inscribed circle of $\triangle ABC$, show that $d = AB + AC - BC$.

26. **[E26]** Let K be a circle and P be a point outside K. The two tangent lines from P are tangent to K at points A and B. Let C be any point on the arc from A to B, and construct a tangent to K at C. The three tangent lines form a triangle. Show that the perimeter of the triangle is $PA + PB$.

27. **[E27]** ♦ Let K be the inscribed circle of any $\triangle ABC$, and let I be the center of K. From B, drop a perpendicular to line AI, intersecting it at E. Show that E lies on the line connecting the points of tangency of K with BC and AC.

28. **[E28]** Let line L_1 intersect a circle K at A and B, and let line L_2 intersect K at C and D. If $L_1 \perp L_2$, then show that $AC^2 + BD^2 = AD^2 + BC^2$.

29. **[E29]** Let $ABCD$ be a cyclic quadrilateral. Let G be the intersection of its diagonals. Let $GH \perp AD$ with H on AD, and $GI \perp BC$ with I on BC. Show that $GH : GI = AD : BC$.

30. **[E30]** Let K be a circle, and let A and A' be inverse points with respect to K, and similarly, B and B' are another pair of inverse points with respect to K. Show that A, A', B, and B' either all lie on a straight line, or $ABB'A'$ is a cyclic quadrilateral.

31. **[E31]** If O is the incenter of an arbitrary $\triangle ABC$, and X and Y are the points of tangency of the incircle with edges AB and AC of $\triangle ABC$, show that O, X, Y, and A all lie on a circle.

32. **[E32]** Given an equilateral triangle, divide one side arbitrarily into two pieces of length x and y. Construct two more equilateral triangles using the pieces of lengths x and y as one side of each new triangle, and construct the new triangles so that both are on the outside of the original triangle. Show that the triangle connecting the centers of the three equilateral triangles is itself an equilateral triangle.

33. **[E33]** If A and B are points at opposite ends of the diameter of a circle, and C is any point on the circle, show that $AC^2 + BC^2$ is a constant.

34. **[E34]** Let AB be the diameter of a semicircle, and let C be a point on AB. The perpendicular to AB through C intersects the semicircle at D and let E be the midpoint of the semicircle. If r is the radius of the semicircle, show that $CD^2 + CE^2 = 2r^2$.

35. **[E35]** ♦ Let AB be the diameter of a circle, and let CD be a chord of the circle parallel to AB, where C is closest to A. Let X be any point between A and B. Show that $XC^2 + XD^2 = XA^2 + XB^2$.

36. **[E36]** Suppose that the diameter of a circle is divided into a large number of pieces, and that circles are constructed with each of those pieces as its diameter. Show that the sum of the circumferences of the smaller circles is equal to the circumference of the original circle.

37. **[E37]** If three congruent circles are mutually tangent, what is the ratio of the area of the triangle connecting their centers with the area of the triangle that connects their points of tangency?

38. **[E38]** ♦ Let K be a circle and let P be a point outside K. From P draw the two tangents to K having points of tangency A and B. Let CD be any diameter of K. If AC and BD intersect at a point Q, show that $PQ \perp CD$.

39. **[E39]** If K_1 and K_2 are two circles that are externally tangent to each other, and AB and CD are parallel diameters of K_1 and K_2, show that AD and BC pass through the point of tangency of K_1 and K_2.

40. **[E40]** ♦ Let D be the midpoint of AC in $\triangle ABC$, and let E be any point on BC. The line DE intersects AB at G, and intersects the line through B that is parallel to AC at point F. Show that $\mathcal{H}(GE, DF)$. In other words, show that the four points G and E divide D and F harmonically.

41. **[E41]** Given an angle, construct its bisector and a line perpendicular to the bisector passing through the vertex of the angle. Let L be any line intersecting all four of the lines in the figure. Show that this L intersects the four lines in a harmonic set.

42. **[E42]** Given two circles of radius 1 tangent to each other, and each tangent to a line that does not intersect the circles at their point of tangency. What is the radius of a circle that is tangent to both circles and to the line?

43. **[E43]** Given two circles that do not intersect and such that one does not lie inside the other. Show that the four points where the common external tangents touch the circles lie in a circle K_1. Also show that the four points where the common internal tangents touch the circles lie in another circle K_2. Show that the four points of intersection of the common external tangents and the common internal tangents also lie on a circle K_3. Finally, show that K_1, K_2, and K_3 are concentric.

44. **[E44]** ♦ Given two circles that do not intersect and do not lie one inside the other, if the common internal and external tangents are constructed, show that the length of an internal tangent is equal to the distance between the points on the external tangent where it is intersected by the two internal tangents.

45. **[E45]** ♦ Given an arbitrary $\triangle ABC$. Let D lie on BC such that AD bisects $\angle BAC$. Show that $AD^2 = AB \cdot BC - BD \cdot CD$.

46. **[E46]** If $\triangle ABC$ is a right triangle with the right angle at A, and a perpendicular to BC is dropped from A that meets BC at D, show that $AB^2 : AC^2 = BD : BC$.

47. **[E47]** Given a circle K and a point A outside of K, let AX and AY be the two tangents from A to K, where X and Y lie on K. Let P be any point on K, and extend line AP to Q on K. Show that $XP/XQ = YP/YQ$.

48. **[E48]** Given any $\triangle ABC$, show that it is possible to construct a rectangle with one vertex at A, and the other three vertices on the circles whose diameters are AB, AC, and BC.

49. **[E49]** ♦ Suppose hexagon $ABCDEF$ is inscribed in a circle, and the diagonals AD, BE, and CF meet at a point. Show that $AB \cdot CD \cdot EF = BC \cdot DE \cdot FA$.

50. **[E50]** ♦♦ Suppose a circle intersects all three sides of $\triangle ABC$ twice. It intersects AB at C' and C'', it intersects BC at A' and A'', and it intersects CA at B' and at B''. Suppose further that the lines AA', BB', and CC' meet at a point. Then the lines AA'', BB'', and CC''' also meet at a (different) point.

51. **[E51]** ♦♦ Given $\triangle ABC$ and line L, let L intersect the sides of AB, BC, and CA at C', A', and B', respectively (some or all of the intersections may be outside the triangle's edges). Show that the midpoints of the segments AA', BB', and CC' all lie on a straight line. (This line is called the **Gauss line**.)

52. **[E52]** ♦ Given a right triangle $\triangle ABC$, where $\angle B = 90°$, construct a circle internally tangent to the legs AB and BC, and to the arc $\overset{\frown}{AC}$. Show that the diameter of this circle is twice the diameter of the inscribed circle of $\triangle ABC$.

53. **[E53]** Given $\triangle ABC$, let A', B', and C' be the feet of the altitudes opposite A, B, and C, respectively. If perpendiculars are dropped from A' to AB, AC, BB', and CC' meeting them at K, N, L, and M, then K, L, M, and N lie in a straight line.

54. **[E54]** ♦ Given two squares $ABCD$ and $AEFG$ with centers P and Q that share a common point A, let J and K be the midpoints of the segments DG and BE, respectively. Show that $PJQK$ is a square. This is called the "Finsler-Hadwiger theorem".

55. **[E55]** Pick four points A, B, C, and D in clockwise order on a circle. Let W, X, Y, and Z be the midpoints of the arcs $\overarc{AB}$, $\overarc{BC}$, $\overarc{CD}$, and $\overarc{DA}$, respectively. Show that $WY \perp XZ$.

56. **[E56]** ♦ Let M be the midpoint of the longer arc $\overarc{CD}$ of a circle K. Pick a point X at random on K. Construct XC, XD, and drop a perpendicular from M to XD, meeting it at Y. Show that $XC + XY = YD$. This is called "Archimedes' midpoint theorem".

57. **[E57]** ♦ In any $\triangle ABC$, let M be the midpoint of AC and let MN be parallel to the angle bisector of $\angle ABC$, where N is the other intersection with the triangle (on either AB or BC, depending on the shape of $\triangle ABC$). Show that MN divides the perimeter into two equal areas. (Hint: See problem 56.)

58. **[E58]** If an altitude of a triangle is extended to intersect the circumcircle, then the length of the segment from the orthocenter to the base of the triangle is equal to the length of the segment from the base to the circumcircle.

59. **[E59]** In $\triangle ABC$ if the excircle opposite $\angle A$ touches the side BC at D, then AD bisects the perimeter of $\triangle ABC$.

60. **[E60]** Given two circles and a line that are all tangent at a common point X with both circles on the same side of the line, pick any other point on that line and draw the other tangent to the inner circle. That line intersects the outer circle twice, at A and B, and is tangent to the inner circle at P. Show that $\angle AXP = \angle PXB$.

61. **[E61]** Show that if $\angle BAC = 30°$ in $\triangle ABC$, then BC is equal to the length of the radius of the circumscribed circle of $\triangle ABC$.

62. **[E62]** If circle K has its center on another circle, then the common chord of the two circles is equal to the length of the chord of K passing through the intersection of the two circles and tangent to the circle on which its center lies.

63. **[E63]** Let A, B, and C be three non-collinear centers of circles, none of which intersect, and none of which lie inside another. Let A', B', and C' be the intersections of the common internal tangents to the circles centered at B and C, at C and A, and at A and B, respectively. Show that AA', BB', and CC' meet at a point.

64. **[E64]** ♦ Given $\triangle ABC$, construct two right triangles $\triangle BAD$ and $\triangle BCE$ on the outside of $\triangle ABC$ such that $\angle ABD = \angle CBE$. Let F be the foot of the altitude from B to AC. Show that the lines AE, CD, and BF are concurrent.

65. **[E65]** Draw a pentagram (a five-pointed star) by choosing five points A, C, E, G, I and forming the segments AE, EI, IC, CG, and GA. Let AE and IC intersect at B, CG and AE at D, EI and CG at F, GA and EI at H, and IC and GA at J. Show that

$$\frac{AB}{BC} \cdot \frac{CD}{DE} \cdot \frac{EF}{FG} \cdot \frac{GH}{HI} \cdot \frac{IJ}{JA} = 1.$$

66. **[E66]** Construct a square on the hypotenuse of a right triangle and draw a line from the vertex opposite the hypotenuse to the center of the square. Show that this line splits the right angle into two $45°$ angles. Does it matter on which side of the hypotenuse the square is constructed?

67. **[E67]** Let $\triangle ABC$ be a right triangle with $\angle ABC = 90°$. Let $AB = 3 \cdot BC$. If M is the midpoint of AC, and D and E trisect AB, show that $\triangle DME$ is an isosceles right triangle.

68. **[E68]** A circle with diameter PP' is tangent to a line L at P'. Let XY be a chord of the circle parallel to PP'. If Q is the midpoint of the arc $\overarc{XY}$, and if lines PX, PQ and PY are extended to intersect L at X', Q', and Y', respectively, show that $\mathcal{H}(P'Q', X'Y')$.

69. **[E69]** ♦ If a quadrilateral is circumscribed around a circle and if the lines connecting the opposite points of tangency are perpendicular, then the quadrilateral can be inscribed in a larger circle.

70. **[E70]** If the altitudes of $\triangle ABC$ are h_a, h_b, and h_c, and the radius of the inscribed circle is r, show that $1/h_a + 1/h_b + 1/h_c = 1/r$.

71. **[E71]** If $ABCD$ is a trapezoid with $AB \parallel CD$ and squares are erected on the exterior of sides AB and CD, the line connecting the midpoints of the squares passes through the intersection of lines AC and BD.

72. **[E72]** ♦Let I be the incenter of $\triangle ABC$. If O is the circumcenter of $\triangle BIC$, show that A, I, and O lie in a straight line.

73. **[E73]** In a semicircle with center O inscribe a circle K tangent at O. Inscribe a second circle K_1 in the semicircle tangent to the semicircle, its diameter, and to K. What is the ratio of the radius of K_1 to the radius of the semicircle?

74. **[E74]** Three circles of radius 1 are tangent to each other and their diameters are all part of the same line. A line is drawn from the left end of the diameter of the leftmost circle which is tangent to the rightmost circle. This tangent line cuts a chord from the middle circle. Find the length of that chord.

75. **[E75]** Given two circles of radius R and r tangent to each other, and both tangent to a line that does not pass through their point of common tangency, construct a circle tangent to both circles and the line. Find the radius of the new circle. Suppose that you begin with two circles of radius 1, inscribe the circle between them and their tangent line, then inscribe circles between both pairs of the three circles, then between all pairs of the set of 5 circles, and so on. What are the radii of the circles so generated? In fact, if two adjacent circles have radii $1/m^2$ and $1/n^2$, where m and n are integers, what's the radius of the new circle? (See Exercise 42.)

76. **[E76]** ♦Given two triangles $\triangle ABC$ and $\triangle DEF$ inscribed in a circle such that the vertices alternate: A, D, B, E, C, F, show that an ellipse can be inscribed in the interior hexagon. What if the two triangles are inscribed in an ellipse instead of in a circle?

77. **[E77]** In $\triangle ABC$, let A', B', and C' lie on lines BC, CA, and AB, respectively such that AA', BB', and CC' meet at a point. Show that

$$\left(\frac{\sin \angle BAA'}{\sin \angle CAA'} \right) \cdot \left(\frac{\sin \angle CBB'}{\sin \angle ABB'} \right) \cdot \left(\frac{\sin \angle ACC'}{\sin \angle BCC'} \right) = 1$$

.

78. **[E78]** In $\triangle ABC$, let A', B', and C' lie on lines BC, CA, and AB, respectively such that AA', BB', and CC' meet at a point P. If the lines AA', BB', and CC' are reflected across the corresponding interior angle bisectors, those reflected lines will also meet at another point Q. Points P and Q are called **isogonal conjugates**. (See Exercise 77.)

79. **[E79]** Show that in any $\triangle ABC$, the orthocenter and the circumcenter are isogonal conjugates. See Exercise 78 for the definition of isogonal conjugate.

80. **[E80]** Show that the diameter of the nine-point circle of a triangle is half the diameter of the circumcircle of that triangle.

81. **[E81]** Show that if a ray is drawn from the orthocenter H of a triangle that intersects the nine-point circle at P and the circumcircle at Q, that $HP = PQ$.

82. **[E82]** ♦Let AA', BB', and CC' be the altitudes of $\triangle ABC$, where A', B', and C' lie on BC, CA, and AB, respectively. Show that AA', BB', and CC' are the angle bisectors of $\triangle A'B'C'$.

83. **[E83]** Let P, A, B, and C lie on a circle. Construct the three circles with diameters PA, PB, and PC. In pairs, those three circles meet in three points other than point P where they all meet. Show that those three intersection points lie on a straight line.

84. **[E84]** If three circles of equal radius centered at A, B, and C all pass through the same point O, and no pair are tangent, let A', B', and C' be the pairwise intersections of the circles centered at B and C, at C and A, and at A and B, respectively, show that AA', BB', and CC' meet at a point.

85. **[E85]** Let A, B, C, D and E be 5 concyclic points let A', B', C', D' and E' be the midpoints of CD, DE, EA, AB and BC, respectively, and let V, W, X, Y and Z be the midpoints of $B'E'$, $A'C'$, $B'D'$, $C'E'$ and $E'A'$, respectively. Show that V, W, X, Y and Z are concyclic.

86. **[E86]** Let A, B and C lie on a line with C between A and B such that $AC = 3CB$. Two circles are drawn with diameters AC and CB. The common exterior tangents of those two circles meet at D. Show that $2BD = BC$.

87. **[E87]** Let $\triangle ABC$ contain only acute angles. Construct circles having diameters AB, BC and CA. From each vertex of the triangle, draw the two tangent lines to the circle whose diameter is the opposite side. Show that the six points of tangency all lie on a circle.

88. **[E88]** ♦Show that if the angle bisectors of a triangle are equal, then the triangle is isosceles. This is the Steiner-Lehmus theorem.

89. **[E89]** If two circles intersect at points A and B and a line that passes through the segment AB intersects the circles at the points W, X, Y and Z in that order, show that $\angle ZAY = \angle WBX$.

90. **[E90]** If the quadrilateral $ABCD$ has perpendicular diagonals and if W, X, Y and Z are the midpoints of the sides, then if perpendiculars are dropped from W, X, Y and Z to the points W', X', Y' and Z' on the opposite side, then all eight points W, X, Y, Z, W', X', Y' and Z' lie on the same circle. This is called the eight-point circle theorem.

91. **[E91]** If a circle is drawn whose diameter is the segment connecting the incenter with one of the excenters, then that circle passes through two vertices of the triangle. This is called the incenter-excenter theorem.

92. **[E92]** Let A be a point outside circle O and choose a point P such that the tangent PT to the circle has the same length as PA. If C is any point on O, let B and D be the intersections of PC and CA with O, respectively. Let E be the intersection of AB with O. Show that $DE \parallel AP$.

93. **[E93]** Show that if the cevians AD, BE and CF of $\triangle ABC$ are concurrent, then the other intersections of the triangle with the circle passing through D, E and F are the endpoints of another set of concurrent cevians.

94. **[E94]** Show that if two chords of a circle are perpendicular, then the sum of the squares of the chord segments is equal to the square of the diameter of the circle.

95. **[E95]** If three circles of equal radius R meet at a point, then the pairwise intersections of those circles lie on a circle that also has radius R. This is known as Johnson's theorem.

96. **[E96]** Given two lines AB and AC, choose D and E to lie on lines AB and AC, respectively. Let F be the intersection of CD and BE. Show that the four circumcircles of $\triangle ABE$, $\triangle ADC$, $\triangle BDF$ and $\triangle CEF$ are concurrent. (Hint: consider the Simson line.)

97. **[E97]** Show that if $\triangle ABC$ is a right triangle with $\angle C = 90°$, then the altitude through C is equal to the median through C reflected across the angle bisector through C. Notice that this proves that the three symmedians of a right triangle meet on the altitude through the right angle, since we know that the three symmedians meet at a point and this exercise shows that one of the symmedians is the altitude through C.
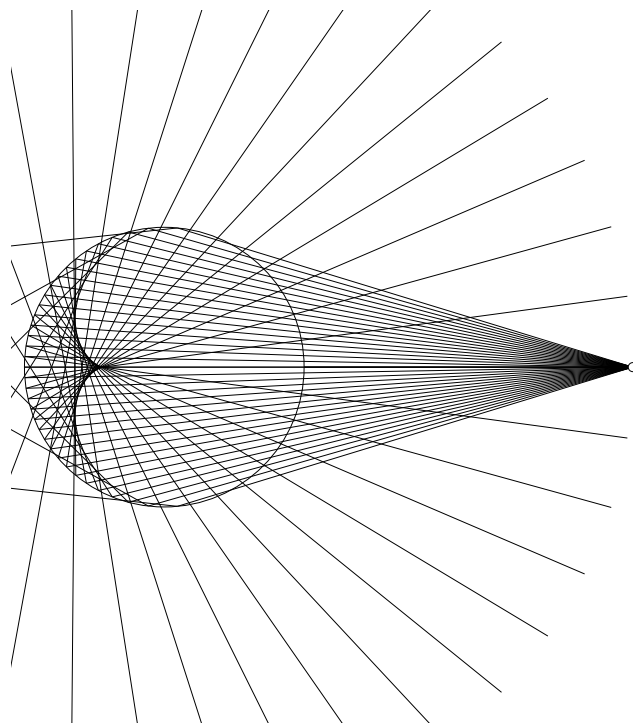
Mathematics Review

This book is about solving geometric problems, so this chapter reviews mainly geometric topics, but not entirely. Vectors, complex variables, and analytic geometry, for example, can sometimes provide quicker and better solutions for problems that at first appear to be completely geometric.

If you feel even a little rusty with respect to high school mathematics, it may be worthwhile to skim through this appendix to get a feeling for what is important.

Since this appendix is a review, in most cases proofs of the results are not included. However, many of the more important theorems do exist as proofs in **Geometer** files. These may be useful for teachers of geometry to present the proofs to a class in a more interactive way. When such a proof exists for a particular result, it can be found in the directory Proofs on the CD, and a reference to it will be included in the text accompanying the statement of the result in this chapter.

No attempt has been made to present the material in this chapter in a "logical" order such that each topic is defined and discussed before it is used. The ordering here

is designed to show the reader the relationships among the various parts of geometry. Almost any high school geometery text can be consulted for a logically consistent order.

A.1 Types of Geometric Problems

Here is a list of the typical sorts of problems that are posed in geometry. It is difficult to divide geometry problems cleanly into sets, but here are some general categories.

- **Finding the measure of something.** This is why geometry was invented—given some measurements (lengths, angles, areas), calculate some other measurements.

  The most useful tools to solve these sorts of problems are formulas like the pythagorean theorem, the law of sines and the law of cosines, Stewart's theorem, the formulas for areas of objects, et cetera, but there are other techniques available.

- **Geometric constructions.** The most famous of these problems require constructions of geometric objects given certain initial objects and a straightedge and compass[1]. Sometimes construction with other sets of allowable tools is studied, but the most famous set consists of a straightedge and compass. Construction problems are somewhat artificial, but the exercise of solving them can be very beneficial to one's geometric intuition.

- **Finding and proving relationships.** Quite often these problems simply require that you prove some relationship, but problems that require you to study a situation to find the relationship first and then to prove it are more instructive.

  The word "relationship" can mean many things. Here are a few examples. Are certain lengths or areas or ratios or angles equal? Is one always greater than the other? Are two figures congruent or similar? Do three lines or circles meet at a point, or do four points lie on a circle? Are two lines parallel or perpendicular? Is a line tangent to a circle? Are two circles concentric or orthogonal?

- **Finding a locus of points.**

  Everyone knows that a circle is the set of points at a fixed distance from a given point, but what is the set of points equidistant from a point and a line? Given a circle and a point, what is the locus of points at the intersection of a line tangent to a circle and the line perpendicular to it passing through the point? Such easily-stated problems can lead to the study of fantastic curves.

A.2 Congruence

---

[1]Mathematicians assume that the tools are ideal in the sense that they can be used to draw infinitely accurate lines and circles.

Congruence is technically an undefined term. In theory all we know about congruence is from its properties as stated in the postulates, but it is easy to understand intuitively. Two geometric figures are **congruent** if they are exactly the same shape and size. In other words all measurements are the same, be they lengths or angles. If five lines meet at a point in one figure, the corresponding five lines meet at the corresponding point at the same angles in a congruent figure, et cetera.

Imagine two geometric figures drawn on transparent plastic. They are congruent if one of the pieces of plastic can be placed on top of the other in such a way that the two figures match exactly. When doing this matching, it is OK to turn one of the sheets upside-down.

Figure A.1: Congruent Triangles
`Review/Congruent.T` [M]

♦ If a sheet can be flipped to make a match, the two figures are said to be mirror images of each other. Notice that if a figure has certain symmetry properties, it can be a mirror image of itself. The concept of congruence is not limited to two-dimensional figures; the same idea makes sense in three (or more) dimensions, but it is harder to imagine "drawings" on three dimensional plastic sheets and especially in making them coincide. But what is especially hard to imagine is turning these three dimensional sheets "over" to obtain the mirror image so that, for example, your left hand and right hand would be approximately congruent. The flipping of this three-dimensional sheet of plastic would require movement through a fourth dimension.

The symbol "≅" is used to indicate congruence. Figure A.1 shows four congruent triangles. When we list the vertices in the equation that shows the congruence, we list them in the same order. Thus $\triangle ABC \cong \triangle A'B'C'$, but it is *not* true that $\triangle ABC \cong \triangle B'A'C'$. This second formula would indicate that A in the first triangle matches B' in the second, B matches A', and so on.

If the geometric figures have certain amounts of symmetry, there may be multiple ways to set up a congruence. Every figure, of course, is congruent to itself, but for something like an equilateral triangle, there may be many self-congruences. If $\triangle ABC$ is equilateral, then we have:

$$\triangle ABC \cong \triangle BCA \cong \triangle CAB \cong \triangle ACB \cong \triangle CBA \cong \triangle BAC.$$

A regular hexagon would be congruent to itself in 12 ways.

♦ Notice in figure A.1 that although they are congruent, $\triangle ABC$ and $\triangle A'''B'''C'''$ are mirror images of each other—it is impossible to move one of them in the plane until it lies exactly on top of the other; one of them must be flipped over before they can be made to match. Notice also that when a triangle is flipped, the clockwise/counterclockwise ordering of the points reverses. In $\triangle ABC$ the points A, B, and C appear in a clockwise orientation; in $\triangle A'''B'''C'''$ the corresponding points A''',
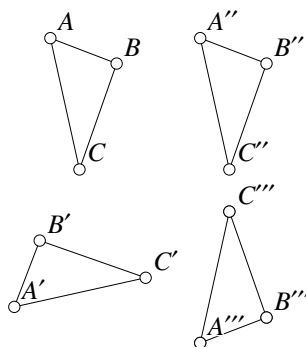
B''' and C''' appear in a counterclockwise order. Any mirror image of an object will convert clockwise to counterclockwise orientations and vice-versa.

Although high-school geometry books concentrate on the congruence of triangles, any sorts of geometric figures can be compared to see if they are congruent. Two line segments are congruent if they are the same length. Two circles are congruent if they have the same radius, two angles are congruent if they have the same measure, et cetera.

Congruence is called an **equivalence relation** because it has the following three properties:

1. The relation is reflexive; every figure is congruent to itself: $A \cong A$.

2. The relation is symmetric: if $A \cong B$ then $B \cong A$.

3. The relation is transitive: if $A \cong B$ and $B \cong C$ then $A \cong C$.

A.2.1   Similarity

Similarity is like congruence, but possibly using a magnifying glass. Two geometric figures are similar if they have the same shape, but not necessarily the same size. In other words, two figures are similar if it is possible to make them congruent by magnifying one of them.

The symbol "$\sim$" is used to indicate similarity. If $\triangle ABC$ is similar to $\triangle DEF$, we write $\triangle ABC \sim \triangle DEF$. In the same way that the vertices of congruent objects are listed in a corresponding order, so are the vertices of similar objects. In the example above, this means that if we magnified $\triangle DEF$ to a new triangle $\triangle D'E'F'$, then $\triangle D'E'F' \cong \triangle ABC$, where D' corresponds to A, E' to B, and F' to C.

Similarity, like congruence, is an equivalence relation. Remember that congruent objects are also similar—the magnification is by a factor of 1.

Similarity is used more often than congruence. If it is known that two figures are similar, then all the ratios of the corresponding lengths are the same. In other words, if one drawing is twice as big as another (in length), then every

Figure A.2: Proof of Chord Theorem

        `Review/Crossproof.T [P]`

line segment in the first drawing is exactly twice as long as the corresponding segment in the other. *In almost any geometric problem where there is a product of lengths or a ratio of lengths involved, the solution will probably follow from similarity arguments.*

As an example of this technique, consider figure A.2. Suppose we are given a circle with two lines passing through a point P inside it. If A, B, C and D are the

intersections of the lines with the circle as in the figure, then $AP \cdot PB = CP \cdot PD$ (or equivalently, $st = uv$).

If we are going to show that $st = uv$ using similarity arguments, we had better form some triangles and there are only a few possibilities. Here we construct the segments AC and DB although it also works to connect AD and BC. Similarity lets us deal with ratios, so if $st = uv$ that is equivalent to $s : v = u : t$ (or, if you prefer, $s/v = u/t$). This will be true if $\triangle BDP \sim \triangle CAP$.

Two triangles are similar if they have two equal angles. $\angle BPD = \angle CPA$ because they are vertical angles. Also, $\angle CAP = \angle BDP$ since they both are inscribed in the same arc $\overset{\frown}{CB}$. Therefore $\triangle BDP \sim \triangle CAP$, so $s/v = u/t$, so $st = uv$.

The entire field of trigonometry is based on similarity. See Section A.10.1

A.2.2 Measure

The word "geometry" is derived from the Greek $\gamma\epsilon\omega\mu\epsilon\tau\rho\iota\alpha$ which comes from $\gamma\epsilon\omega-$, the combining form of $\gamma\eta$ (earth) plus $-\mu\epsilon\tau\rho\iota\alpha$ (measure). The original application of geometry was to measure the earth—to find areas of fields, distances between cities, and so on.

In spite of this, in pure Euclidean geometry, there is no such thing as an absolute measure of anything. Two line segments can be compared with the result that they are the same length, or that one is twice as long as the other, or that one is $\sqrt{2}$ times as long as the other, but the process is always one of comparison. There is no particular segment that is "one inch" or "one meter" long. If a diagram for a theorem in Euclidean geometry is put in a copy machine with magnification on, the resulting diagram will illustrate the theorem equally well.

Geometric figures are scale drawings where the scale is never stated. If a diagram contains line segment A that is twice as long as line segment B, then if A is two inches, B is one inch, or if A is two centimeters, B is one centimeter, or if A is two light-years, B is one light-year, et cetera. All the measurements in a diagram are relative.

Of course it is possible to talk about ratios of areas (and of volumes, in solid geometry) but the types must match in comparisons. Areas can be compared with areas, volumes with volumes, et cetera. Just as in physics you cannot compare a time with a distance you cannot compare a geometric area with a volume. You can compare ratios like $area : length = area : length$, but notice that the types on both sides are the same—an area to length ratio in this case.

In physics, a velocity can be measured in feet per second, miles per hour, or furlongs per fortnight. All these are different units, but they all have something in common: each is a length divided by a time, and all can be interconverted. But a distance cannot be converted to a velocity—one has units of length and the other has units of length divided by time.

Similarly, in geometry, lengths cannot be compared to areas or volumes, since areas have units of "length squared" and volumes of "length cubed". A ratio of two lengths has units of length divided by length, so it is dimensionless, and cannot be compared to
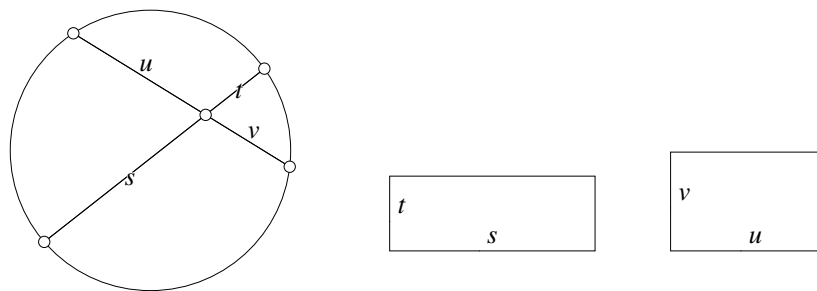
Figure A.3: The Chord Theorem II
`Review/Cross.D [D]`

lengths or areas. But a ratio of an area to a length is just a length—its units are "length squared" divided by length, or just length.

Statements about ratios of lengths can be converted to equivalent statements about areas by cross-multiplying. In the previous section we proved the chord theorem which states that $s{\cdot}t = u{\cdot}v$ (see figure A.3). The same theorem can equally well be stated as an equality of ratios: $s : v = u : t$ (which is equivalent to $s/v = u/t$, unless $v = t = 0$). Multiply this equation on both sides by v and t to obtain the area equality $s \cdot t = u \cdot v$. A **Geometer** proof of the chord theorem can be found in `Proofs/CrossProd.T` and `Proofs/Crossproof.T`.

Of course the product of two lengths can be interpreted as an area—the area of a rectangle whose sides are those two lengths. Saying that $s \cdot t = u \cdot v$ is equivalent to stating that two rectangles, one with sides s and t and the other with sides u and v, have the same area. Those two rectangles are shown in the right part of figure A.3.

This idea of converting a ratio equality to an area equality or vice-versa can be extremely powerful. Ratios arise very naturally from similarity (see Section A.2.1), so if you need to prove something about the product of two lengths and you do not have any idea how to proceed, convert those products into ratios and see if you can show that they are parts of similar geometric figures[2].

A detailed examination of the concept of area is beyond the scope of this book. Even if we say that the area of a circle having radius r is πr^2 we have moved beyond simple geometry to calculus—what does it really mean to talk about the area of a geometric figure with curved edges when area was originally defined only for shapes whose boundaries are straight lines?

Like ratios, angles are dimensionless. The "30" in a $30°$ angle is effectively part of a ratio: it is $30°/360° = 1/12$ of a $360°$ circle. In radians, the same angle is still $1/12$ of a circle, but this time the measure of $\pi/6$ is in relation to a circle whose circumference is 2π. The unit of measure of an angle is simply a comparison to the somewhat arbitrary measure that is assigned to a full circle. The pros and cons of various angle measures are discussed in detail in Section A.5.1.

---

[2]Sometimes it is convenient to consider positive and negative ratios (this is sometimes called a **directed ratio**—see Section 5.2.1).

A.3 Points

The term **point** itself is technically undefined—a point is simply an object that happens to satisfy the properties specified by the postulates. It is useful to think of a point as an infinitely small dot or position on the plane (or in space, in the case of solid geometry).

Sometimes it is useful to think of other geometric figures—lines, circles, parabolas, et cetera, as collections (or sets) of points. In fact, a definition of a circle might be: "A circle is the set of all points that are equidistant from a given point." Or equivalently: "A circle is the locus of points equidistant from a given point."

A point has no size—zero length and zero area. (In *Elements* ([Euclid, 1956]), Euclid tried to define the term as follows: "A point is that which has no part.")

Here are some definitions of special types of points:

Midpoint. A point M is the **midpoint** of two other points A and B if it lies on the line AB and $AM = MB$.

Collinear Points. A set of points is **collinear** if all of them lie on the same line.

Concyclic Points. A set of points is **concyclic** if they all lie on the same circle.

Point of Intersection. When two lines, or a line and a circle, or two circles (or various other curves, like conics sections) have a point in common, that common point is called the **point of intersection**. Depending on the figures, there may be more than one point of intersection. Quite often the terms "intersection of two lines", "intersection of two circles", et cetera, is used in place of "the point of intersection of two lines", "the point of intersection of two circles", et cetera.

Point of Tangency. When a line is tangent to a curve (in other words, when the line and the curve locally share a single point and have the same slope), that point in common between the line and curve is called the point of tangency. See the definition of tangency in the next section.

There are very few theorems that deal exclusively with points; almost all deal with points in relation to other geometric figures, and those theorems are covered in the sections that follow.

A.4 Lines

Like "point", the term **line** is technically undefined, but most people have a good idea what it signifies. In Euclidean geometry, "line" will refer to a straight line that extends forever in both directions. It is "infinitely thin", in the sense that if you pass across it, its width is only a single point. A line is composed of an infinite number of points.

If A and B are two different points, "the line AB" refers to the particular line that passes through those two points (the postulates guarantee that there is only one such

line). In some books the symbol $\overleftrightarrow{AB}$ is written to mean "the line AB", where the arrow above emphasizes that the line continues forever in both directions.

Given two points, the line segment connecting them is the shortest path between them.

A common misconception is that one can somehow talk about points on that line that are "next to" each other. Between any two different points on the line there are other points (in fact an infinite number of other points).

Here are some definitions of special lines, and of portions of lines:
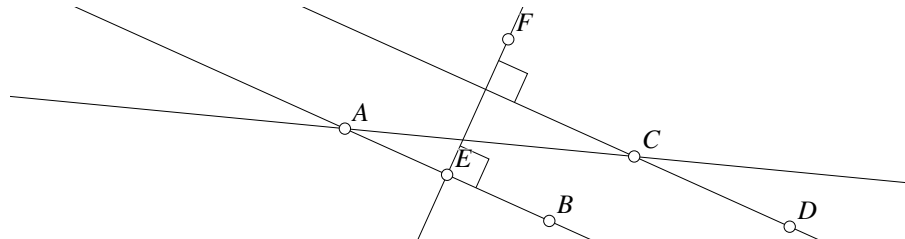
Figure A.4: Lines
Review/Lines.D [D]

Segment. A **segment** is a piece of a line, consisting of two points on that line (the endpoints of the segment) and all the points that lie between those endpoints on the line. Sometimes the segment AB is written as follows: $\overline{AB}$. A segment has a length, but its area is zero.

Ray. Intuitively, a **ray** is half of a line—pick a point on a line and all the points on one side of it and you have a ray. This can be made precise as follows: Given a segment AB of a line L, the ray $\overrightarrow{AB}$ consists of all the points of the segment AB, together with all the points C on L such that B lies between A and C. The ray $\overrightarrow{AB}$ is said to emanate from point A.

Parallel Lines. Two lines are **parallel** if they lie in the same plane and have no points in common. (In plane geometry there is only one plane, so we merely need to say the lines have no points in common.) If L_1 and L_2 are two parallel lines, we write $L_1 \parallel L_2$. In figure A.4, lines AB and CD are parallel. In Euclidean geometry, parallel lines remain the same distance apart.

Transversal Line. Given a pair of parallel lines, a line that cuts through both of them is said to be a **transversal**. In figure A.4, lines AC and EF are transversals to the pair of parallel lines AB and CD.

Tangent Line. A line is **tangent** to a circle if it has exactly one point in common with the circle. In other words, it just touches the edge of the circle. All the points of the circle except for the point of tangency lie on one side of the line. The general concept of tangency of a line and a curve requires calculus to define
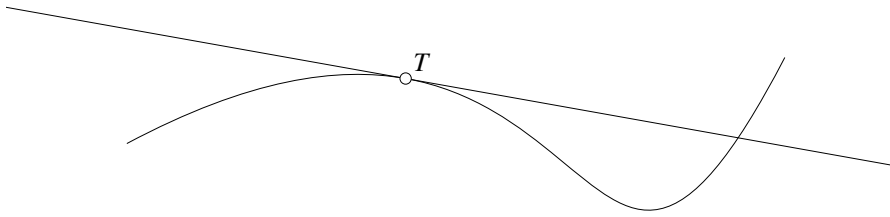
Figure A.5: Tangent Line
`Review/Tangent.D [D]`

what is meant by the slope of a curve at a point. A line is tangent to a curve if it passes through a point on the curve and the slope of the line is the same as the slope of the curve at that point. Note that a line can be tangent at one point and then pass through the curve at another or that the tangent line can pass through the curve at the point of tangency. Figure A.5 illustrates a line tangent to a curve at the point T that passes through the curve at another point.

Two curves (including circles, of course) are tangent at a point if they both are tangent to the same line at that point.

Perpendicular Lines. Two lines are **perpendicular** if they form a right angle (a $90°$ angle). If L_1 and L_2 are perpendicular, we write $L_1 \perp L_2$. In figure A.4, $EF \perp AB$. (Of course $EF \perp CD$ as well.)

Perpendicular Bisector. A line that is perpendicular to a segment and divides the segment into two congruent segments is called the **perpendicular bisector** of the segment. In figure A.4, point E is the midpoint of AB, so $AE = EB$. Since line EF is perpendicular to line AB, then EF is the perpendicular bisector of the segment AB.

Here are some key theorems about lines in plane Euclidean geometry. All can be proved from the postulates but many of them are simply assumed in a standard high school geometry course.

- Two distinct lines can intersect in at most one point.

- Every line contains an infinite number of points.

- Given any four distinct points on a line, it is possible to label them A, B, C, and D such that B lies between A and C and between A and D, and also that C lies between A and D and between B and D.

- Every line divides the plane into two regions such that if two points not on the line are in the same region, then the segment connecting them does not intersect the line, and if the two points are not in the same region, then the segment connecting them does have a point in common with the line.

A.5 Angles

It is a bit tricky to define the term **angle**. See Section A.5.1 for a more thorough discussion. Here is a usable definition: Two rays in the plane $\overrightarrow{OA}$ and $\overrightarrow{OB}$ both emanating from the same point O, and not lying in the same line, form an angle ($\angle AOB$ or $\angle BOA$). The angle divides the rest of the plane into two regions. A point P of the plane is on the interior of $\angle AOB$ if a line exists through P that intersects both rays such that P lies between the two points of intersection. If no such line exists then P is exterior to the angle.
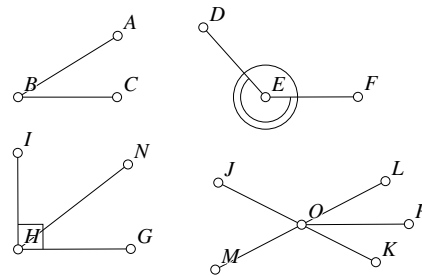
Figure A.6: Angles
`Review/Angles.D [D]`

Notice that this definition requires angles to be larger than $0°$ and smaller than $180°$. With this definition, straight angles (angles measuring $180°$) and reflex angles (angles larger that $180°$) do not exist. The discussion and definitions below sometimes ignore these problems.

Supplementary Angles. Two angles that add to a straight angle (to $180°$), are called **supplementary angles**. They need not share a vertex. In figure A.6, $\angle MOJ$ and $\angle JOL$ are supplementary.

Right Angle. An angle is a **right angle** if it is congruent to its supplementary angle. Intuitively, right angles are made by perpendicular lines, or lines that meet at $90°$. In figure A.6, $\angle GHI$ is a right angle.

Complementary Angles. Two angles are called **complementary angles** if they add together to make a right angle. They need not share a common vertex. In figure A.6, $\angle GHN$ and $\angle NHI$ are complementary.

Vertical Angles. When two lines cross, four angles are formed, and the opposite pairs are known as **vertical angles**. In figure A.6, $\angle LOK$ and $\angle MOJ$ are vertical angles, as are $\angle LOJ$ and $\angle MOK$.

Acute Angle. An **acute angle** is an angle that is less than a right angle. Without using the concept of the measure of an angle, we can simply say that an acute angle is an angle that can be obtained inside a right angle. In figure A.6, $\angle ABC$, $\angle GHN$, $\angle NHI$, $\angle LOK$, and $\angle JOM$ are acute angles, among others .

Obtuse Angle. An **obtuse angle** is an angle that can contain a right angle. In other words, an angle that is larger than $90°$ (and less than $180°$) is obtuse. In figure A.6, $\angle DEF$, $\angle JOL$, and $\angle KOM$ are examples of obtuse angles.

Angle Bisector. A ray emanating from the same point as the two rays defining an angle that divides the angle into two congruent angles is called the **angle bisector**. In figure A.6, the ray $\overrightarrow{OP}$ is the angle bisector of $\angle KOL$. In other words, $\angle KOP = \angle POL$.

Straight Angle. A **straight angle** is not a proper angle as defined above, but you can think of it as a $180°$ angle. In figure A.6, $\angle MOL$ is a straight angle.

Reflex Angle. A **reflex angle** is not a proper angle as defined above, but you can think of it as the "outside" of the angle in question. In other words, a $30°$ angle will have a reflex angle of $330°$. In figure A.6, $\angle DEF$ and its reflex angle are shown with one ring and two rings. They would both be written the same way: $\angle DEF$, so there is clearly some deficiency in the naming convention that can be avoided with directed angles (see Section A.5.2).

A.5.1 More on Angles

The previous definition of angle restricts the concept of angle to those having measures between (but not including) $0°$ and $180°$. For much of geometry, this is sufficient, but a much more general definition of angle is often useful. It is sometimes nice to allow angles of any size—positive, negative, or zero.

A common approach is to define the measure of an angle as a sort of total turning relative to a fixed ray. In the Euclidean plane with a standard Cartesian coordinate system, that fixed ray is usually the positive x-axis, and positive angles are measured in a counterclockwise direction from that. Negative angles, of course, represent clockwise turning. Angles can thus have any real measure, positive, negative, or zero, and can be added or subtracted without worry—another valid angle is certain to result.

If angles can have any positive or negative measure, one has to decide what is meant by congruence. Is a $10°$ angle equal to a $370°$ angle? There is nothing particularly tricky going on, but it is good to keep in mind that alternative definitions have both pros and cons.

In this text, we will measure angles in degrees instead of radians. This is in line with what appears in most geometry textbooks, and it is easy to convert back and forth. To convert between degrees and radians, π radians is equal to 180 degrees.

A.5.2 Directed Angles

If we wish to refer to reflex angles there is a labeling problem since an angle and its reflex angle share the same vertices. For example, in figure A.6, the points D, E, and F in reality determine two angles, one of which is marked with a single ring and the other with two rings. If one has measure θ, the other has measure $360° - \theta$, but both would be written $\angle DEF$.

One way around this is the idea of **directed angles**—assume that "$\angle DEF$" implies that the ray $\overrightarrow{ED}$ is the first one, and the ray $\overrightarrow{EF}$ is the second. The angle goes

counterclockwise from the first ray to the second so in figure A.6, $\angle FED$ is the one that is less than $180°$ and $\angle DEF$ would be its reflex angle—an angle larger than $180°$. We seldom encounter situations where this distinction is necessary, but the distinction can be made. For a more complete discussion, see [Johnson, 1917].

$\boxed{\text{A.6}}$ Triangles

If A, B, and C are any three non-collinear points, then the **triangle** $\triangle ABC$ consists of the three segments AB, BC, and CA. The points A, B, and C are called the **vertices** of the triangle and the angles $\angle ABC$, $\angle BCA$, and $\angle CAB$ are its angles. Sometimes it is useful to think of a triangle as the union of the three lines of which the segments are parts. A triangle is just a special case of a polygon, so some general polygon definitions, like "perimeter" and "area", for example, will apply to triangles as well.

Equilateral Triangle. If the three segments composing a triangle are congruent, the triangle is **equilateral**.

Isosceles Triangle. If two of the three segments composing a triangle are congruent, the triangle is an **isosceles triangle**. The third side does not necessarily have to be different—an equilateral triangle is a special case of an isosceles triangle.

Scalene Triangle. A **scalene** triangle is a triangle none of whose segments are congruent. Equivalently, no pair of its three angles are congruent.

Right Triangle. If two of the segments composing a triangle are perpendicular, it is called **right triangle**. The two sides forming the right angle are called **legs** and the third side is called the **hypotenuse**. Although most right triangles are technically scalene triangles, they are hardly ever referred to in that way.

Acute Triangle. If the three angles of a triangle are all acute (less than $90°$), it is called an **acute triangle**.

Obtuse Triangle. If one of the angles of a triangle is obtuse (greater than $90°$), it is called an **obtuse triangle**.

Median. A segment connecting a vertex of a triangle to the midpoint of the opposite side is called a **median** of the triangle. Every triangle has three medians and they meet at a point called the **centroid**, **center of gravity**, or **center of mass**. The median of $\triangle ABC$ in figure A.7 is G. A **Geometer** proof appears in `Proofs/Medians.T`

Centroid. The intersection of the three medians. The centroid is $2/3$ of the distance from any vertex to the midpoint of the opposite side.

Inscribed Circle. A circle inscribed inside a triangle and tangent to all three sides. It is also called the incircle.
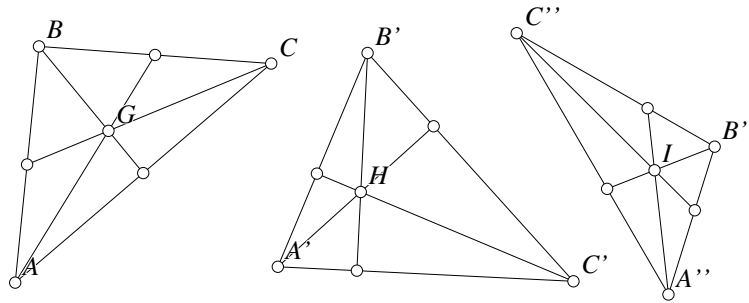
Figure A.7: Medians, Angle Bisectors, and Altitudes
`Review/Tricenters.T [M]`
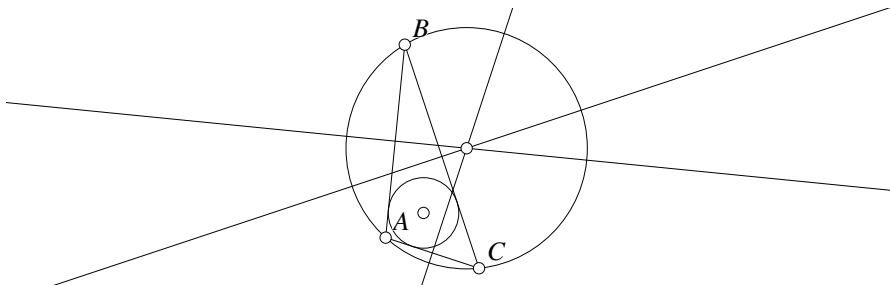
Figure A.8: Inscribed and Circumscribed Circles
`Review/Trianglecircs.T [M]`

Incenter. The three angle bisectors of the angles of a triangle meet at a point called the **incenter**. The incenter is the center of the **inscribed circle** or **incircle** of the triangle—a circle that is tangent to all three segments making up the sides and inside the triangle. The incenter of $\triangle A''B''C'''$ in figure A.7 is I. See also figure A.8. A **Geometer** proof of this fact can be found in the file `Teaching/ThreeIn.T`.

Altitude. An **altitude** of a triangle is a line passing through one of its vertices and perpendicular to the side opposite that vertex. The three altitudes of a triangle meet at a point called the **orthocenter**. The orthocenter of $\triangle A'B'C'$ in figure A.7 is H. Be careful—although both the centroid and incenter are always inside the triangle, the orthocenter may be outside it. A **Geometer** proof that the three altitudes meet at a point can be found in `Proofs/Orthocenter.T`.

Orthocenter. The point where the three altitudes of a triangle intersect.

Excenter. If a triangle is considered to be three lines rather than three segments, there are four circles tangent to all three lines. One, the incircle, is tangent to the three segments. Each of the others, called an **excircle**, is tangent to the interior of one segment, and to the other two lines exterior to the triangle. The center of an excircle is called an **excenter**. Every triangle has three excircles and three

excenters. See Section A.6.8.

Circumcircle. The circle passing through all three vertices of a triangle. See figure A.8.

Circumcenter. The **circumcenter** of a triangle is the center of the circle (called the **circumcircle**) that passes through the three vertices of the triangle. The center of the circumcircle can be found at the intersection of the three perpendicular bisectors of the segments that make up the sides of the triangle. See figure A.8.

A.6.1 | Congruence and Similarity of Triangles

There are a few standard methods to prove that two triangles are congruent. In high school textbooks, they are usually listed as an abbreviation like "SAS", "ASA", "SSS", or "AAS". These stand for "Side-Angle-Side", "Angle-Side-Angle", et cetera.

"SAS", or "Side-Angle-Side" stands for the idea that if two of the sides in one triangle have the same lengths as two sides in another triangle, and further that the angles between those two sides are equal in both triangles, then the triangles themselves are congruent. Because they are congruent, you then automatically know that all the measurements in the two triangles are the same—the other side and the

Figure A.9: SSA Failure
`Review/Ssa.D` [D]

other two angles. (And in fact you also know that corresponding medians will be equal, that corresponding altitudes will be equal, that the areas of the two triangles will be equal, and so on. The order "SAS" is important, and can be interpreted as the order that you read off the parts of the triangles as you go around. First there is a side, then an angle, then another side. There is no "SSA" method because if the angle is not the included angle—the angle between the two sides—then the triangles are not necessarily congruent.

In figure A.9 are two triangles that are clearly not congruent, but they have two pairs of equal sides and an equal angle, but not the angle between them. The two triangles are $\triangle ADB$ and $\triangle ACB$. Side AB is in both of them, so that is one match. Sides AD and AC are also clearly equal since they both lie on a circle centered at A, and both triangles share the angle at B.

Although "SSA" is not in general a valid method of proof of the congruence of a pair of triangles, it is valid in the special case that the angle is at least $90°$, and in the special case where it is exactly $90°$ the method is known as the "hypotenuse-leg" method. It is quite instructive to play around with the **Geometer** illustration for figure A.9 and try to adjust the vertices so that the angle at B is a right angle or larger and see why in this case an "SSA-where-A-is-a-right-angle-or-larger" congruence theorem is true.

"ASA" states that two triangles are congruent if two angles and the side between them are congruent. Both "ASA" and "AAS" are equivalent for the simple reason that if two of the angles of a triangle are equal to the two corresponding angles in another, then the third angles of both must be equal as well, since the three angles of a triangle always add to $180°$.

Finally, if all three sides in a triangle are equal in length to the three sides of another triangle, then the triangles are congruent, so the corresponding angles will also be equal. This is known as the "SSS", or "Side-Side-Side" method[3].

Proving the similarity of two triangles is almost always done by showing that they have two equal angles (and hence, three equal angles). This is called "Angle-Angle" similarity, or just "AA".

There are also the "SAS" and "SSS" similarity conditions: If the included angles between pairs of sides in two triangles are equal, and if the ratios of the lengths of the sides are equal, then the triangles are similar. Finally, if all three sides of two triangles are in the same ratio, the triangles are similar. In practice, these last two methods are seldom used.

A.6.2 General Properties of Triangles

- The lengths of any two sides of a triangle added together are longer than the third side. In other words, if a triangle's sides have lengths a, b, and c, then $a + b > c$ (and or course $b + c > a$ and $c + a > b$). This is known as the **triangle inequality**.
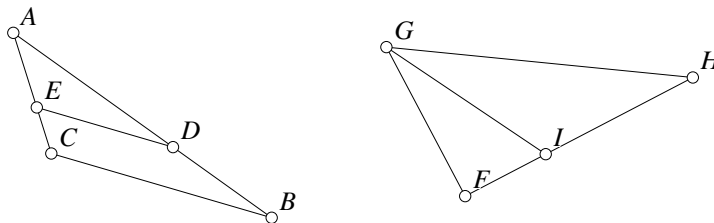
Figure A.10: Elementary Triangle Properties
`Review/Triprop.D [D]`

- If a line is parallel to one side of a triangle, it divides the other two sides proportionally. In figure A.10, if $DE \parallel BC$ then $AE : EC = AD : DB$, $AE : AC = AD : AB$, et cetera.

- A line parallel to one side of a triangle cuts off another triangle that is similar to the original triangle. In figure A.10 if $DE \parallel BC$ then $\triangle ABC \sim \triangle ADE$.

---

[3]The fact that three lengths determine a triangle is key to many mechanical construction techniques. If three bars are bolted together to form a triangle, that triangle cannot change shapes unless the bars bend or break. Geodesic domes are constructed entirely of triangles, and many rectangular features in architecture are braced across the diagonal, essentially turning the flexible rectangle into two non-flexible triangles.

- The median of a triangle divides the triangle into two triangles of equal area. More generally, a line from a vertex of a triangle to the opposite side divides the triangle into two triangles with areas proportional to the lengths into which that side was divided. In figure A.10, $FI : IH = \mathcal{A}(\triangle FIG) : \mathcal{A}(\triangle HIG)$, no matter where I lies on the interior of the segment FH.

- The angle bisector of a triangle divides the side opposite the bisected angle into two segments proportional to the lengths of the adjacent sides of the original triangle. In figure A.10, if $\angle FGI = \angle IGH$ then $FI : IH = GF : GH$. A **Geometer** proof appears in `Proofs/Anglebisect1.T`.

- In $\triangle ABC$, if the lengths of the sides opposite the angles A, B, and C are a, b, and c, respectively, then if $\angle A = \angle B$ we know that $a = b$. Conversely, if $a = b$, then $\angle A = \angle B$. Finally, if $\angle A > \angle B$, then $a > b$, and conversely.

In other words, in any triangle, the larger sides are opposite larger angles. So in the $30°$–$60°$–$90°$ triangle whose sides have lengths 1, $\sqrt{3} \approx 1.7320508$, and 2, the side of length 1 is opposite the $30°$ angle, the side of length $\sqrt{3}$ is opposite the $60°$ angle, and the side of length 2 is opposite the right angle.

This theorem is usually proved in the case of equality of the sides or angles by dropping a perpendicular line to divide the triangle into two smaller ones that are shown to be congruent by SAS or AAS.

Here is a beautiful proof that sides opposite equal angles are equal and vice-versa. Suppose that in $\triangle ABC$, $AB = AC$. Then by SAS, $\triangle BAC \cong \triangle CAB$, so $\angle ABC = \angle ACB$. Similarly, if $\angle ABC = \angle ACB$, then $\triangle BAC \cong \triangle CAB$ by ASA, so we can conclude that $AB = AC$. Both proofs show that a triangle is congruent to the mirror-image of itself.

A.6.3 Area of a Triangle

The area of a triangle is half of the base times the altitude. Remember that the altitude is the perpendicular line to the base from the vertex opposite the base. The altitude may not even intersect the base inside the triangle. For any triangle, there are three base-altitude pairs that can be used to calculate the area, and all of them give the same result, so one pair may be easier to use than another.

Figure A.11 shows the three ways to calculate the area of $\triangle ABC$ using the base-altitude formula. The three altitudes are AA', BB', and CC'. Notice that two of the altitudes, AA' and BB' lie outside the original triangle $\triangle ABC$, although in a general triangle they are
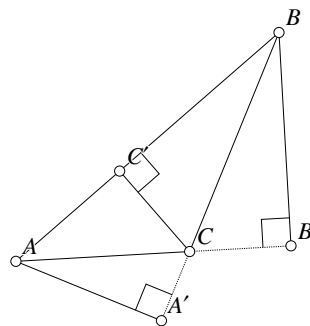
Figure A.11: Area of a Triangle
`Review/Trianglearea.D` [D]

sometimes in and sometimes out. Any of the following expressions gives the area of $\triangle ABC$:

$$\mathcal{A}(\triangle ABC) = \frac{AB \cdot CC'}{2} = \frac{BC \cdot AA'}{2} = \frac{CA \cdot BB'}{2}.$$

From the fact that the area of a rectangle is equal to its base times its height, it is very easy to conclude that the area of a triangle is half the base times the height. In figure A.12, suppose the problem is to determine the area of the triangle at A. Make an exact copy of the triangle, but flip it over and place it against itself as in B. The area of this parallelogram will be twice the area of the original triangle. To find the area of the parallelogram, cut off the right triangle at the right end of the parallelogram, as in C. Move that part to the left as in D. Thus the area of the rectangle at D is twice the area of the original triangle. The base of the rectangle at D is the same as the base of the original triangle, and it is also easy to see that the altitude of the rectangle is the same as the altitude of the original triangle. The area of the rectangle is its base times its height, which is the same as the base times height of the original triangle, but since the rectangle's area is made up of two copies of the triangle, the triangle's area is half that of the rectangle.

The argument in the paragraph is intuitive, but is not quite right—if the altitude hits the base very far outside the segment which is the base, it may not be possible to slice the parallelogram as described. If the top point of the triangle written A in figure A.12 is moved to the right by a long distance, the argument does not work. The argument can be made rigorous most easily with an appeal to Cavalieri's principle which states that if two plane figures can be placed so that all lines parallel to a given line cut equal length segments in both figures, the two figures have equal area[4].
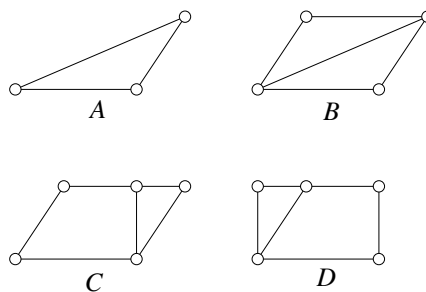
Figure A.12: Area of a Triangle
`Review/Trianglearea1.D [D]`

A.6.4 Angles in a Triangle

The three interior angles of a triangle add to $180°$. Figure A.13 contains most of a proof of this fact. In any $\triangle ABC$, find the midpoints of the sides (X, Y, and Z) and connect them. It is easy to show that all four of the smaller triangles produced are congruent. That means that the three angles that meet at Z: $\angle AZY$, $\angle YZX$, and $\angle XZB$ are equal to the vertex angles in the original $\triangle ABC$. But they also fit together perfectly on the line AB, so the sum of the three angles of any triangle is $180°$. A **Geometer** proof of this fact appears in `Proofs/Sum3angles.T`.

---

[4]The three-dimensional version of Cavalieri's principle states that if two solids can be arranged so that all planes parallel to a given plane intersect the solids with equal areas, the solids have equal volumes.
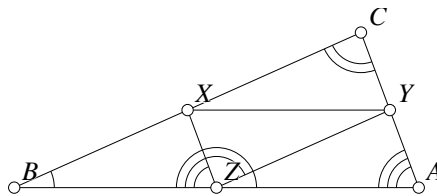
Figure A.13: Sum of Angles in a Triangle
`Review/Sum3angles.T [P]`

A trivial (but often very useful) consequence of this is that any external angle of a triangle is equal to the sum of the other two interior angles. Figure A.13 shows this as well. In $\triangle BXZ$, the exterior angle at Z is $\angle AZX$, and that is clearly equal to the sum $\angle AZY + \angle YZX$ which are equal to the other two angles in the original $\triangle BXZ$ (remember that all four of the smaller triangles in this figure are congruent).

In a right triangle, of course, this means that the two acute angles must add to make a right angle (in other words, they are complementary) since the exterior angle of the right angle is also a right angle.

♦ The theorem above cannot be proved without the so-called "parallel postulate" which states that given a line and a point not on the line, there exists exactly one line passing through that point that is parallel to the given line. In other geometries, Riemannian geometry, projective geometry, and others, the parallel postulate does not hold, and the angles of triangles in those geometries need not add to $180°$.

♦ To see why it might be reasonable to consider a geometry where a triangle's angles do not add to $180°$, imagine that you live on a sphere (that should not be too hard) and that you measure angles of triangles on its surface. If the triangles are small, the angles will add to almost $180°$, but as they get larger, the error becomes greater.

Consider a trip along the triangle that begins at the north pole and goes due south along the prime meridian ($0°$ longitude) through the observatory at Greenwich, England, then through Spain near Valencia, then through north Africa, finally exiting that continent through Ghana. Just after you enter the Atlantic Ocean, you will arrive at the equator, where you then take a $90°$ turn to the west and go a quarter of the way around the earth through South America and entering the Pacific Ocean from Ecuador. After you plunge into the Pacific a bit north of Guayaquil,

Figure A.14: Pythagorean Theorem
`Review/Pythagorean.D [D]`

continue a few hundred kilometers along the equator to the $90°$ line of latitude somewhere near the middle of the Galápagos Islands, and finally take another $90°$ turn due north passing near Guatemala City, through the Yucatan in Mexico, and very near New Orleans, Louisiana. Continue along this line all the way to the north pole where you close the triangle. All the paths you took were as straight as you could make them, and all three angles in that giant spherical triangle measure $90°$, so the total of the interior angle measurements is $270°$. If you are willing to look at even larger triangles that cover more than $1/8$ of the globe, you can make the total measure as close to $540°$ as you want. For example, if you go past the Galápagos Islands and across the Pacific to near the international date line before you turn north, you can make the angle at the north pole as close to $180°$ as you would like.

A.6.5   The Pythagorean Theorem

Certainly the most famous, and probably the most useful theorem in all of Euclidean geometry is the pythagorean theorem. Not only is it useful in standard geometric proofs, but it is also the basis for coordinate geometry, or so-called analytic geometry (see Section A.11).

Here is a statement of the theorem:

Theorem A.1 (Pythagorean Theorem) *The square of the length of the hypotenuse of a right triangle is the sum of the squares of the lengths of the opposite two sides and conversely, if the sum of the squares of two sides of a triangle is equal to the square of the third side, the triangle is a right triangle.*

The theorem is illustrated in figure A.14. In $\triangle ABC$, where $\angle C = 90°$, squares are constructed on each of the sides. The length of sides AB, BC, and CA are c, a, and b, respectively. Thus the squares have areas of a^2, b^2, and c^2. The pythagorean theorem simply states that $a^2 + b^2 = c^2$. There are literally hundreds of proofs of this theorem; what follows are two proofs, the first of which is computational, and the second is more visual.

Figure A.15: The Pythagorean Theorem Proof

`Review/Pyproof1.D [D]`

In figure A.15, $\triangle ABC$ is a right triangle with $\angle C = 90°$. Drop an altitude from C that meets the segment AB at point D. Because of the AA similarity theorem, all three of the triangles in the figure are similar: $\triangle ABC \sim \triangle CBD \sim \triangle ACD$.

By similarity, we have $AD/AC = AC/AB$ and $DB/CB = CB/AB$. Multiply these out, giving: $AC^2 = AB \cdot AD$ and $CB^2 = AB \cdot DB$. Add them: $AC^2 + CB^2 = AB \cdot DB + AB \cdot AD = AB \cdot (DB + AD) = AB \cdot AB = AB^2$, which completes the proof.

Here is another proof which combines a nice visual and analytic approach:
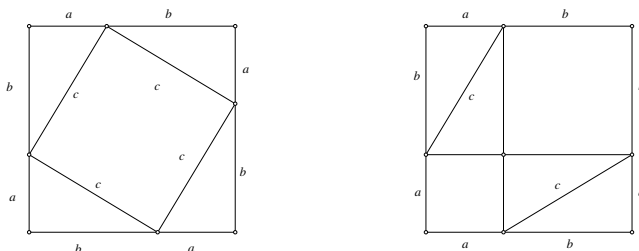
Figure A.16: Proof 2 of the Pythagorean Theorem
`Review/Pyproof2.D, Review/Pyproof3.D [M]`

Make four copies of the right triangle whose sides have lengths a, b, and c and arrange them to form a square as shown in the left of figure A.16.

We can calculate the area of the square in two ways. If we just square the length of a side, we get $(a + b)^2$. If we notice that the square is composed of a central square and 4 equal triangles, we get: $c^2 + 4(ab/2) = c^2 + 2ab$. Setting those areas equal and simplifying, we obtain:

$$\begin{aligned} (a+b)^2 &= c^2 + 2ab \\ a^2 + 2ab + b^2 &= c^2 + 2ab \\ a^2 + b^2 &= c^2. \end{aligned}$$

But there is even a more visual proof. Take the squares and triangles on the left and rearrange them as in the diagram on the right in figure A.16. Both figures have four triangles, and the remaining area is either two squares with areas a^2 and b^2 or one with area c^2. Clearly $a^2 + b^2 = c^2$.

We obviously need to show that both the figures above are, indeed, squares and that the object in the center in the right side of figure A.16 is also a square, but this is easy since we know that the three angles of a triangle add to $180°$ and that in a right triangle, one of the angles is $90°$.

A.6.6 Stewart's Theorem

Although this theorem does not normally appear in high school geometry courses, it is easy to prove and it is *extremely* useful:
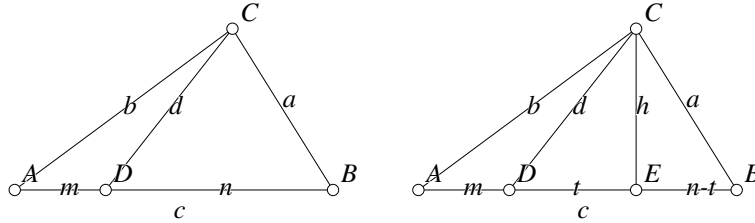
Figure A.17: Stewart's Theorem
`Review/Stewart.D [D]`

Theorem A.2 (Stewart's Theorem) *In any triangle $\triangle ABC$, let D be any point on the segment AB. If the segment lengths are: $a = BC$, $b = AC$, $c = AB$, $d = CD$, $m = AD$ and $n = BD$ then:*

$$a^2 m + b^2 n = c(mn + d^2).$$

The proof is fairly simple. Drop a perpendicular line from C to AB and suppose it intersects at E between D and B as in the diagram on the right in figure A.17. (A

complete proof requires that all possibilities of the position of E be considered, but all are similar.) Then the pythagorean theorem yields:

$$
\begin{aligned}
d^2 &= t^2 + h^2 \\
a^2 &= (n-t)^2 + h^2 \\
b^2 &= (t+m)^2 + h^2
\end{aligned}
$$

As direct results of these equations and remembering that $c = m + n$ we obtain:

$$
\begin{aligned}
a^2 m + b^2 n &= m((n-t)^2 + h^2) + n((t+m)^2 + h^2) & \text{(A.1)} \\
c(mn + d^2) &= (m+n)(mn + h^2 + t^2) & \text{(A.2)}
\end{aligned}
$$

The right sides of equations A.1 and A.2 are readily shown to be equal, so the proof is complete.

Stewart's theorem can also be proved by applying the law of cosines repeatedly.

An interesting special case of Stewart's theorem occurs if the angle at C is double the angle at A. In this case $c^2 = a(a + b)$.

A.6.7 Special Triangles

Some triangles appear over and over because they are useful, either in a practical sense, or because they serve as good examples or as examples that are easy to use for sample calculations. One set of them is interesting because the angles are nice round numbers, and the other set because the lengths of the sides are nice round numbers. (Of course it is easy to make up triangles where the angles are nice round numbers, but for the vast majority of them, it is impossible to calculate analytically the lengths of the sides and vice-versa.)
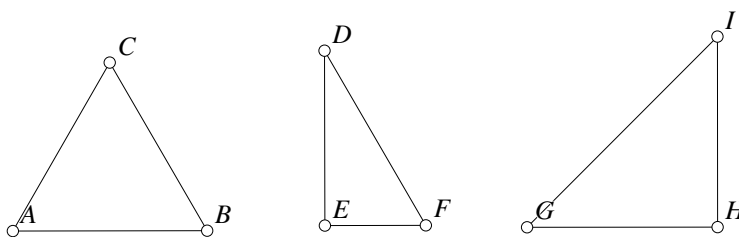
Figure A.18: Special Triangles
`Review/Special.D` [D]

Triangles with nice angles include the $60°\!-\!60°\!-\!60°$ triangle, the $30°\!-\!60°\!-\!90°$ triangle, and the $45°\!-\!45°\!-\!90°$ triangle. See figure A.18. The first one, ($\triangle ABC$) is equilateral—all three sides and angles are the same. The second ($\triangle DEF$) is just the equilateral triangle sliced in half—its three edges have lengths in the ratio $EF : FD :$

$DE = 1 : 2 : \sqrt{3}$. The $45°$–$45°$–$90°$ triangle ($\triangle GHI$) is isosceles, and the ratios of the lengths of its sides are $1 : 1 : \sqrt{2}$. Remember that the larger angles are opposite the longer sides ($\sqrt{2} \approx 1.41421$ and $\sqrt{3} \approx 1.73205$).

The other triangles you see regularly have sides whose lengths have integer ratios and at the same time are right triangles. Any three integers a, b, and h will form a right triangle with h as the hypotenuse if $a^2 + b^2 = h^2$ (this is a result of the pythagorean theorem). Here are a few examples of so-called pythagorean triplets: $(3, 4, 5)$, $(5, 12, 13)$, $(6, 8, 10)$ (of course the $(6, 8, 10)$ is similar to the $(3, 4, 5)$ triangle—just twice as big), and $(7, 24, 25)$. A **reduced pythagorean triplet** is one where the entries have no factors in common. $(6, 8, 10)$ is *not* reduced because there is a factor of 2 in every number.

The pythagorean triples do not have "reasonable" angles. For the $(3, 4, 5)$ triangle, for example, the acute angles are approximately $36.869898°$ and $53.130102°$.

♦ There are an infinte number of reduced pythagorean triplets. If m and n are any two integers with $m > n > 0$, then the following three numbers form a pythagorean triplet: $(m^2 - n^2, 2mn, m^2 + n^2)$. It is easy to show that any such set is a triplet, but it is also true that all pythagorean triplets (reduced and non-reduced) are of this form. If one of m or n is even and if they are relatively prime, then the formula will produce a reduced pythagorean triple.

♦ It is not hard to prove that all pythagorean triples have this form. The problem of finding points with rational coordinates on the unit circle $x^2 + y^2 = 1$ is equivalent to finding integral pythagorean triples. Every line that passes through a point with rational coordinates and the point $A = (-1, 0)$ has a rational slope, and vice-versa. The equation of the line passing through A and having slope n/m is $y = (n/m)(x - 1)$ and if we plug that into the equation of the circle we find that the rational point of intersection other than A is at $((m^2 - n^2)/(m^2 + n^2), 2mn/(m^2 + n^2))$ which is equivalent to the solution in the previous paragraph.

Table A.1 lists a few smaller examples of the form: $(n, m) \longrightarrow (2mn, m^2 - n^2, m^2 + n^2)$. Examples followed by an asterisk (*) are reduced pythagorean triplets.

$(1, 2) \longrightarrow (4, 3, 5)*$	$(1, 3) \longrightarrow (6, 8, 10)$	$(1, 4) \longrightarrow (8, 15, 17)*$
$(1, 5) \longrightarrow (10, 24, 26)$	$(1, 6) \longrightarrow (12, 35, 37)*$	$(1, 7) \longrightarrow (14, 48, 50)$
$(2, 3) \longrightarrow (12, 5, 13)*$	$(2, 4) \longrightarrow (16, 12, 20)$	$(2, 5) \longrightarrow (20, 21, 29)*$
$(2, 6) \longrightarrow (24, 32, 40)$	$(2, 7) \longrightarrow (28, 45, 53)*$	$(2, 8) \longrightarrow (32, 60, 68)$
$(3, 4) \longrightarrow (24, 7, 25)*$	$(3, 5) \longrightarrow (30, 16, 34)$	$(3, 6) \longrightarrow (36, 27, 45)$
$(3, 7) \longrightarrow (42, 40, 58)$	$(3, 8) \longrightarrow (48, 55, 73)*$	$(3, 9) \longrightarrow (54, 72, 90)$
$(4, 5) \longrightarrow (40, 9, 41)*$	$(4, 6) \longrightarrow (48, 20, 52)$	$(4, 7) \longrightarrow (56, 33, 65)*$
$(4, 8) \longrightarrow (64, 48, 80)$	$(4, 9) \longrightarrow (72, 65, 97)*$	$(4, 10) \longrightarrow (80, 84, 116)$

Table A.1: Some Pythagorean Triples

Students taking standardized tests in geometry, or participating in high school mathematics contests, should commit a few of the smaller examples above to memory because they appear all the time! In a recent AHSME high school math contest

(American High School Mathematics Examination for 1999) one of the questions concerned a circle circumscribed about a triangle with sides 20, 21, and 29. Those who knew that the $(20, 21, 29)$ triangle is a right triangle knew that the side of length 29 must be the diameter of the circle, and the problem was trivial[5].

A.6.8 Inscribed and Escribed Circles
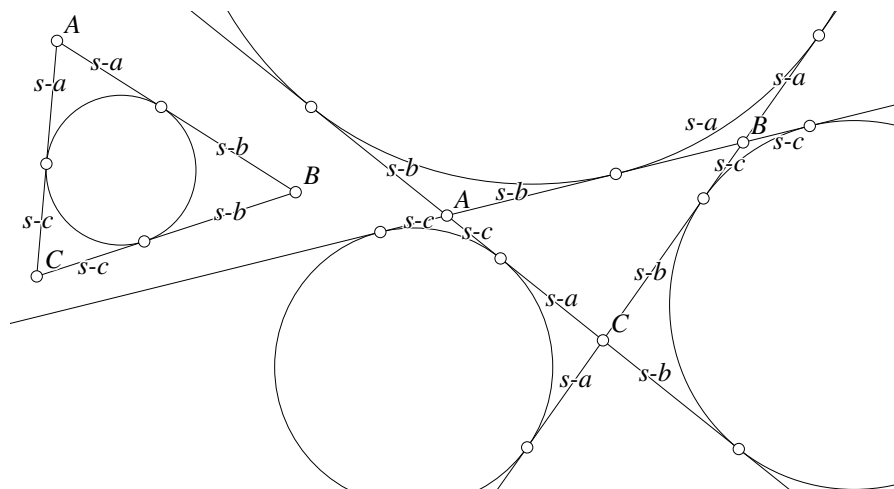
Figure A.19: Lengths to Points of Tangency
`Review/Trilengths.D` [D]

Every triangle has four circles tangent to its three sides—the incircle and three excircles. There are a number of easily derived lengths in these triangles that are extremely useful for calculations. In figure A.19 above, a, b, and c are the lengths of the sides opposite vertices A, B, and C, respectively, and $s = (a + b + c)/2$ is the semiperimeter (half the perimeter). The lengths marked $s - a$, $s - b$ and $s - c$ are measured from the vertex of the triangle to the points of tangency of the inscribed or escribed circle. The figure includes two triangles $\triangle ABC$ since otherwise it would be too cluttered to understand. If all four tangent circles were drawn in one triangle, each length, $s - a$, $s - b$ and $s - c$, would measure six different segments.

These are not hard to derive. If we do not know the lengths, call the lengths from A, B and C to the points of tangency with the incircle α, β and γ, respectively. The perimeter of $\triangle ABC = 2\alpha + 2\beta + 2\gamma = a + b + c = 2s$.

But we also know that $\alpha + \beta = c$, $\beta + \gamma = a$ and $\gamma + \alpha = b$. A little algebra shows that $\alpha = s - a$, $\beta = s - b$ and $\gamma = s - c$.

A similar observation gives us the lengths from the vertices to the excircles. In the figure on the right, let us temporarily call the sides labeled $s - a$, $s - b$ and $s - c$ α, β

---

[5]For those participating in math contests, another interesting triangle to know about is the one whose sides are 13, 14, and 15. It is just a 5-12-13 triangle back to back with a 9-12-15 triangle. The 9-12-15 triangle is the 3-4-5 triangle in disguise.

and γ, respectively. Since the distances from A to the tangent points with the excircle on the other side of BC are the same, we have $c + \gamma = b + \beta$. We also see that $\beta + \gamma = a = 2s - b - c$. If we solve these two equations for β, we obtain $\beta = s - b$. Exactly the same derivation gives $\alpha = s - a$ and $\gamma = s - c$.

There is more to say, however. Let r be the radius of the incircle. Then the area of $\triangle ABC$, $\mathcal{A}(\triangle ABC) = r(a + b + c)/2 = rs$. This is because if I is the incenter of $\triangle ABC$ and we split $\triangle ABC$ into $\triangle AIB$, $\triangle BIC$ and $\triangle CIA$. Those three triangles have bases of a, b and c and all have altitude r.

A similar analysis shows that if r_a, r_b and r_c are the radii of the excircles opposite vertices A, B and C, we have:

$$\mathcal{A}(\triangle ABC) = r_a(s - a) = r_b(s - b) = r_c(s - c) = rs. \tag{A.3}$$

A little more algebra yields the following relationship:

$$\frac{1}{r_a} + \frac{1}{r_b} + \frac{1}{r_c} = \frac{1}{r}. \tag{A.4}$$

A.7   Quadrilaterals

The term **quadrilateral** is the general name for a polygon having four sides. The situation becomes surprisingly more complicated when one goes from three-sided triangles to four-sided quadrilaterals, so there are many special names and definitions associated with quadrilaterals. With more than four sides, the situation is so complicated that mathematicians have basically given up—there are very few names for special polygons other than the regular ones for polygons with five or more sides. A quadrilateral is just a special case of a polygon, so remember that general polygon definitions (see Section A.8) apply to quadrilaterals (and even triangles) as well.

For triangles there is a special triangle symbol for notation, so the triangle with vertices A, B, and C is written $\triangle ABC$. Quadrilaterals have so many shapes that it would be somewhat misleading to write something like $\square ABCD$ for the quadrilateral with vertices A, B, C, and D (although many texts do). Here we will simply say something like, "the quadrilateral $ABCD$" or if the specific shape is known, "the square $ABCD$" or "the trapezoid $ABCD$".

Here are some specific definitions related to quadrilaterals.

Square A **square** has all four sides and all four angles equal. Of course, the angles are all right angles. A square is a regular quadrilateral. See A in figure A.20.

Rectangle A **rectangle** has four right angles, but not necessarily four equal sides. See A and B in figure A.20.

Parallelogram A **parallelogram** has as its edges two pairs of parallel line segments. See A, B, C, and D in figure A.20.

Rhombus A **rhombus** has all four sides equal. See A and D in figure A.20.
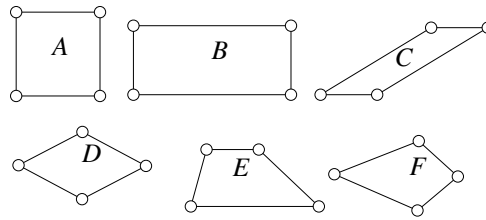
Figure A.20: Special Quadrilaterals
Review/Quadrilaterals.D [D]

Trapezoid A **trapezoid** has two sides parallel, but not necessarily the other two. See anything except F in figure A.20.

Isosceles Trapezoid An **isosceles trapezoid** is a trapezoid that is not a parallelogram, and the two non-parallel sides have the same length.

Kite A **kite** is a convex polygon both of whose pairs of adjacent sides are of equal length. See A, D, and F in figure A.20.
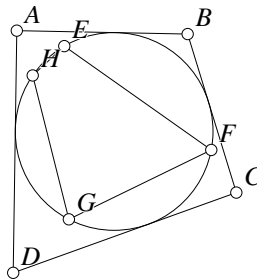
Figure A.21: Inscribed and Circumscribed Quadrilaterals
Review/Circquads.D [D]

Cyclic Quadrilateral If there exists a circle such that the quadrilateral (or any polygon, for that matter) has all its vertices on it, the quadrilateral is called a **cyclic quadrilateral** or a **concyclic quadrilateral**. Quadrilateral $EFGH$ is inscribed in the circle and hence is a cyclic quadrilateral. See figure A.21.

Circumscribable Quadrilateral Similarly, if there is a circle that is tangent to all four sides of a convex quadrilateral, it is called a **circumscribable quadrilateral**. In figure A.21, quadrilateral $ABCD$ is circumscribed about the circle.

Convex Quadrilateral A **convex quadrilateral** is a special case of a convex polygon (see Section A.8). A shape is convex if a line segment connecting any two interior points is completely inside the shape. All the quadrilaterals in figure A.20 are convex. The quadrilateral $EFGH$ in figure A.22 is not. A non-convex quadrilateral (or any non-convex shape) is sometimes called **concave**.
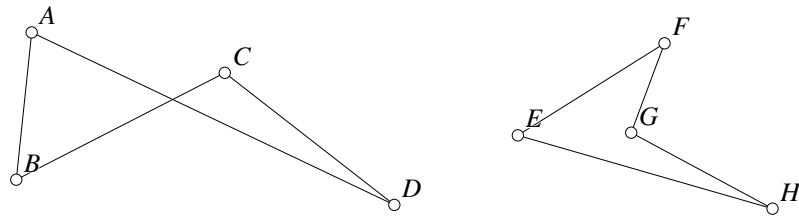
Figure A.22: Bowtie and Non-Convex Quadrilateral
`Review/Badquads.D [D]`

Bowtie Quadrilateral A **bowtie quadrilateral** (or **non-simple quadrilateral**) is a quadrilateral whose edges cross each other such as $ABCD$ in figure A.22. This cannot happen with triangles, but once a polygon has more than three sides, this is possible, and it is the bane of computer graphics programmers. When the edges of a polygon cross themselves, it no longer makes sense to talk about "inside" and "outside".

A.7.1 Quadrilateral Theorems

Many properties of quadrilaterals are simple consequences of the fact that two opposite vertices can be connected by a segment dividing the quadrilateral into two triangles whose properties are easy to calculate. For anything other than special quadrilaterals like rectangles or concyclic quadrilaterals, usually the easiest way to find the area, for example, is to divide it into two triangles and find the sum of the areas of those two triangles.

However, here are a few important facts about quadrilaterals, in no particular order. The vast majority are easy to prove:

- The four internal angles of a non-bowtie quadrilateral add to $360°$. For bowtie quadrilaterals, there is no such thing as an internal angle.

- The diagonals of a rectangle are congruent.

- The diagonals of a rhombus are perpendicular and conversely— if the diagonals of a parallelogram are perpendicular, the parallelogram is a rhombus. See `Proofs/Rhombus.T` for a **Geometer** proof.

- The diagonals of a rhombus bisect the angles at the vertices of the rhombus.

- The opposite sides of a parallelogram are of equal length and the opposite angles have equal measures. See `Proofs/EquParSides.T` for a **Geometer** proof.

- Any pair of adjacent angles in a parallelogram are supplementary.

- Either diagonal of a parallelogram divides it into two congruent triangles. See `Proofs/EquParSides.T` for a **Geometer** proof.

- The diagonals of a parallelogram bisect each other. See `Proofs/ParBisect.T` for a **Geometer** proof.

- If a parallelogram has sides of lengths a and b and diagonals of lengths x and y, then $2a^2 + 2b^2 = x^2 + y^2$.

- The area of a parallelogram is the base times the height. **Warning:** the height is the perpendicular distance from the base to the side parallel to it—it is not the distance along the other edge.

- If the midpoints of the edges of any quadrilateral are connected in order, the resulting figure is a parallelogram. This remains true even for bowtie quadrilaterals. As long as it is not a bowtie quadrilateral, then the area of the parallelogram is half the area of the quadrilateral. This is known as Varignon's theorem.

- An isosceles trapezoid can be inscribed in a circle. See `Proofs/Isotrap.T` for a **Geometer** proof.

- Isosceles trapezoids have congruent base angles. See `Proofs/IsoTrapAng.T` for a **Geometer** proof.

- The area of a trapezoid is the average of the lengths of the parallel sides times the distance between the sides.

- If a quadrilateral $ABCD$ can be circumscribed about a circle, then $AB + CD = AD + BC$ and conversely. See `Proofs/QuadEdges.T` for a **Geometer** proof.

- If a convex quadrilateral $ABCD$ can be inscribed in a circle, then $\angle A + \angle C = \angle B + \angle D = 180°$ and conversely.

$\boxed{\text{A.8}}$ General Polygons

Given n different points in the plane, $A_1, A_2, \ldots, A_n$, a **polygon** is the set of segments A_1A_2, A_2A_3, A_3A_4, $\ldots$, $A_{n-1}A_n$, and A_nA_1. A polygon has three or more sides—it is usually not very useful to think of one-sided or two-sided polygons since then your theorems usually would have the form, "For all polygons except those with one or two sides ...". There is no upper limit, of course, to the number of sides a polygon may have. The points $A_1, \ldots, A_n$ are called the vertices of the polygon, and the segments A_1A_2, $\ldots$, A_nA_1 are called the edges. Sometimes it is important to talk about the angles of the polygon which are $\angle A_1A_2A_3$, $\angle A_2A_3A_4$, $\ldots$, $\angle A_nA_1A_2$. Notice that this is a quite general definition of a polygon and allows the edges to cross each other, et cetera[6]. Special names exist for more restricted classes of polygons. There are a couple of different schemes for naming polygons (including the names for 1- and 2-sided polygons), and the more common ones are listed in Table A.2. The names used in this book are listed first, and one alternative is listed in parentheses.

---

[6]There are, however, still more general definitions of polygons that allow holes or which allow polygons to have more than one component. Although such definitions are sometimes useful in computer graphics, they will not be considered here.

1	monogon	11	hendecagon (undecagon)
2	digon	12	dodecagon
3	triangle (trigon)	13	tridecagon (triskaidecagon)
4	quadrilateral (tetragon)	14	tetradecagon (tetrakaidecagon)
5	pentagon	15	pentadecagon (pentakaidecagon)
6	hexagon	16	hexadecagon (hexakaidecagon)
7	heptagon	17	heptadecagon (heptakaidecagon)
8	octagon	18	octadecagon (octakaidecagon)
9	enneagon	19	enneadecagon (enneakaidecagon)
10	decagon	20	icosagon

Table A.2: Names of polygons

Here are some definitions relating to general polygons. Remember that triangles and quadrilaterals are special cases of polygons, so we can talk about the perimeter of a triangle or square, for example.

Perimeter The **perimeter** of a polygon is the sum of the lengths of the edges of the polygon. It is the distance around the polygon.

Simple Polygon A **simple polygon** is one whose non-adjacent edges do not meet, even at the endpoints. In other words, there are no crossing lines, or places where the polygon comes back and touches itself.

♦ Although it is "obvious" what is meant by the inside and outside of a simple polygon, in practice (for example, when you are trying to write a computer program that deals with general polygons) it can be difficult to determine whether a point is inside or outside a given polygon, especially if the polygon has hundreds or thousands of sides twisted up like a maze. The usual method is based on the Jordan curve theorem which states that any simple closed curve ("simple" means that it does not cross itself, and a sequence of straight edges is just a special kind of curve) divides the plane into two regions—an inside and an outside—where the inside has finite area and the outside infinite area. Furthermore, if a point on the inside is connected to a point on the outside by a straight line, that line must intersect the curve. (Although it is "obvious", this is *not* an easy theorem to prove.)

From the Jordan curve theorem, one can conclude that if you are given a point in the plane not on some simple polygon and you wish to decide if it is inside or outside, draw a ray from that point out to infinity in some arbitrary direction, and count the number of times it crosses the curve (or in your case, segments of the polygon). If the ray crosses an odd number of times, the point was inside; otherwise, it was outside. Even this can be tricky to do in a computational sense. The ray may just touch a vertex of the polygon (this does not count as a crossing), or it may be coincident with an edge for a while. In this second case, whether it crosses or not depends on the relative directions that the segments connected to the endpoints of the coincident segment take.

Convex Polygon A **convex polygon** is a polygon such that given any two points in the interior, the line segment connecting them will also be completely contained in the interior. In other words, a convex polygon does not have any indentations. A star-shaped polygon, for example, is not convex.[7] Every triangle is convex, but if it has more than three sides, a polygon may be non-convex.

Regular Polygon A **regular polygon** is a simple polygon having any number of sides, but all of those sides have the same length, and all the internal angles are the same. We have special names for the smaller ones—an equilateral triangle and a square in the case of three and four sides—but after that, we simply say "a regular pentagon", "a regular hexagon", and so on.

Triangulation A **triangulation** is a subdivision of a polygon into triangles. Any simple polygon can be triangulated using only segments that connect existing vertices of the polygon. Although this may seem obvious, it is not easy to prove— see Section A.8.1.1.

A.8.1 General Polygon Theorems

At first glance, it seems that an unreasonable amount of time is spent studying the properties of a very special type of polygon—the triangle. After all, why not spend an equivalent amount of time studying 4-sided, 5-sided, and 6-sided figures?

A.8.1.1 Polygon Triangulation

The reason is that triangles are relatively simple, and that most other interesting planar figures bounded by straight lines can be decomposed into triangles. If you know everything about triangles, you know a great deal about more complicated geometric figures composed of line segments.

On the left side of figure A.23, for example, we see a polygon that is quite complicated and has 17 sides.

However, on the right side of the same figure, we see that it is possible to break it up into 15 different triangles: $\triangle ABQ$, $\triangle BQE$, $\triangle DBE$, et cetera. If you know all the details of the little triangles composing it, you also know many of the details of the more complex polygon. For example, if you would like to know the measure of $\angle CDE$ in the original polygon, you can look at $\triangle CBD$ and $\triangle BDE$, and find the angles $\angle CDB$ and $\angle BDE$. Those can then be added to get the desired angle in the general polygon.

A.8.1.2 ♦ Triangulation Techniques

Notice that in this example, the subdivision of the complex polygon into triangles was done without adding any vertices, even though the original polygon had all sorts of

---

[7]The word "convex" can be used to describe any figure with this property—not just polygons.
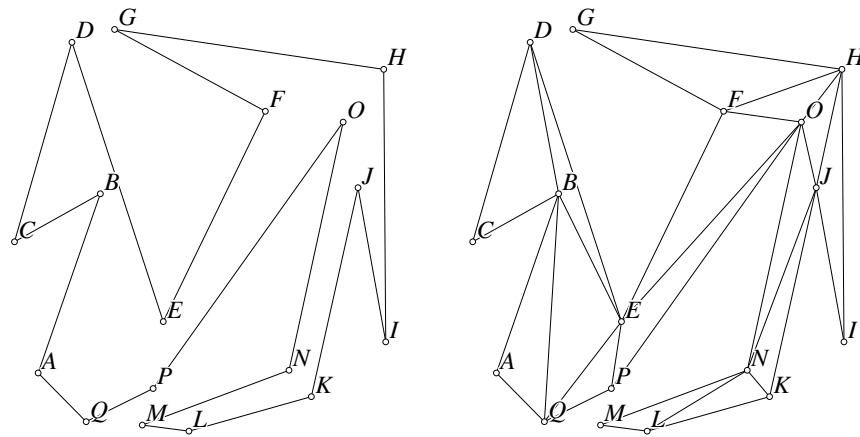
Figure A.23: Triangulation of a Polygon
Review/Polygon.D [M]

weird indentations. It is *always* possible to do this without adding any additional vertices. One proof is beyond the scope of this book, but it involves looking for "ears"—sections of the polygon where a triangle sticks out and can be cut off. If there is always an ear available, clearly any polygon can be so subdivided. Simply find an ear, cut it off, and you are left with one triangle and a polygon with one side fewer. Find the ear on that one, cut it off, and continue the process until all that is left is a triangle.

Another approach to show that an arbitrary polygon can be triangulated finds an interior diagonal (a line connecting two vertices that is completely contained within the polygon). This splits the polygon into two smaller ones, and the triangulation of the polygon can be inferred by an inductive argument.

To prove the existence of an interior diagonal, choose an angle like $\angle ABC$ of polygon $ABCDEFG$ in figure A.24 where the interior of the polygon is on the side of the angle measuring less than $180°$. If the segment AC lies completely within the polygon, then it is the required interior diagonal and we are done. If not, there must be some finite number of vertices in the polygon that lie within $\triangle ABC$. For each such vertex, construct a line parallel to AC and choose a vertex whose line is closest to B. That vertex can be connected to B

Figure A.24: Find an Internal Diagonal
Review/CutPoly.T [M]

completely within the polygon since if it crossed another line, that other line would have to be even nearer B. Note in figure A.24 that we *cannot* use the vertex closest to B—that would be vertex F. It is quite instructive to play with the **Geometer** version

of the figure and experiment with various shapes of the polygon.

A.8.1.3 Total Interior Angle of a Polygon

Since every polygon can be divided into $n - 2$ triangles, it is easy to see that the sum of the interior angles of an n-sided polygon is $(n - 2) \cdot 180°$. Every interior angle will be composed of a sum of angles of the triangles, and all of the angles of all the triangles will be used.

A.8.1.4 Area of a Polygon

As was stated before, sometimes the easiest way to find the area of a polygon is to split it into triangles and then sum the areas of those triangles. If the vertices of a polygon are described by coordinates in a cartesian coordinate system, however, there is a simple formula to calculate the area. See Section A.11.1.

A.9 Circles

A **circle** is the set of all points that are equidistant from a given point. The given point is called the **center** of the circle and the distance from the center to any point on the circle is called the **radius** of the circle. As with polygons, only the points on the boundary are part of the circle; the points inside (including the center) are not. The name **disk** is often used if it is necessary to refer to the points on a circle together with those inside.

Diameter The **diameter** of a circle is twice the radius—it is the distance across the circle through the center.

Circumference The **circumference** of a circle is the distance around the outside of the circle. It is equal to $2\pi r$ or πd, where r is the radius or d is the diameter of the circle. π is just a number, approximately equal to 3.14159265.

Figure A.25: Circle Features
`Review/Circledefs.D` [D]

Tangent A line and a circle can have zero, one, or two points in common. If there is one point in common, we say that the line is **tangent** to the circle and the line is called the **tangent line**. In figure A.25, the line passing through T is tangent to

the circle. In a similar way, two distinct circles can have zero, one, or two points in common, and if there is only one, the two circles are said to be **tangent** at that common point.

Chord If a line intersects a circle at two points, then the segment of the line inside and on the circle is called a **chord**. The entire line that cuts a circle in two places is called a **secant**. The segment AB is a chord, and the line DC is a secant of the circle in figure A.25.

Arc A connected section of a circle is called an **arc**. Note that if points A and B are on a circle, then saying "arc AB" (denoted by $\overset{\frown}{AB}$) is ambiguous—the points determine two arcs, depending on which way around you go.

Except in the case where the two arc endpoints are diametrically opposite, one of the arcs is larger than the other, and $\overset{\frown}{AB}$ usually refers to the smaller of the two. Another reasonable convention is to label the points so the arc goes counterclockwise from the first to the second. But if a theorem simply talks about two arbitrary points on a circle and the arc between them, usually the theorem only makes sense when the shorter one is meant. In this book, we will label arcs using the counterclockwise method when there is an accompanying diagram with a particular configuration.

Sometimes a third point can be given that lies between the endpoints and when the arc is referred to with three letters, it is the part of the circle containing that intermediate point.

If both endpoints of an arc are connected to the center of a circle with a pair of radii, the angle between the radii is the measure of the arc.

Central Angle An angle whose vertex is at the center of a circle is called a **central angle**. An angle inside a circle whose vertex is on the circle is called an **inscribed angle**. $\angle AOD$ is a central angle and $\angle EDC$ is an inscribed angle in the circle in figure A.25.

A.9.1   Circle Theorems

Here are the most important properties of circles:

- The area of a circle is πr^2, where r is the radius of the circle. The symbol π (the Greek letter "pi") is defined to be the ratio of the circumference of a circle to its diameter. $\pi \approx 3.1415926535898$. Figure A.26 shows why this is reasonable to believe. The circle is first divided into a set of pie-shaped segments which are then rearranged to make something that is close to a rectangle. If the circle were divided into more such segments, the approximation to the rectangle would be better. Since the circumference of the circle is $2\pi r$, where r is its radius, it is clear from the figures that the approximated rectangle will have one side approximately equal to r and the other approximately equal to πr (remember that what was the circumference now appears as a bumpy top and bottom of the rectangle approximation), so its area is approximately πr^2, as desired.
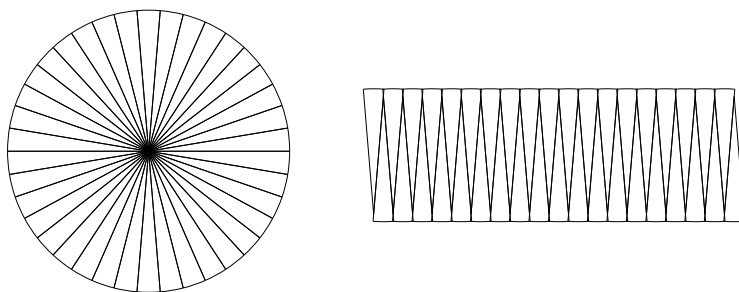
Figure A.26: Circle Segment Rearrangement
`Review/CircAreaFig.T [S]`

- The perpendicular bisector of a chord of a circle passes through the center of the circle.

- Two tangents to a circle from an external point are of equal length.

- A line tangent to a circle at a point T is perpendicular to the radius that connects T with the center of the circle and conversely, if a line is perpendicular to a radius at the point the radius touches the circle, that line is tangent to the circle.
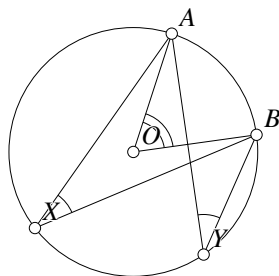
Figure A.27: Inscribed and Central Angles
`Review/Circprops1.D [D]`

- If a central angle and an inscribed angle in a circle subtend the same arc, then the inscribed angle is half the size of the central angle. As a trivial consequence, any two inscribed angles that subtend the same arc have the same measure. In figure A.27, $\angle AOB = 2\angle AXB = 2\angle AYB$. See `Proofs/InscribedAngle.T` for a **Geometer** proof.

- Another trivial, but extremely important consequence of the item above is that any angle inscribed in a half circle (a semicircle) is a right angle. Conversely, if the hypotenuse of a right triangle is the diameter of a circle, then the apex of the angle lies on the circle. In figure A.28 all three of the angles $\angle AXB$, $\angle AYB$, and $\angle AZB$ are right angles since they are inscribed in a semicircle (the segment AB is the diameter). This is known as Thales' theorem.
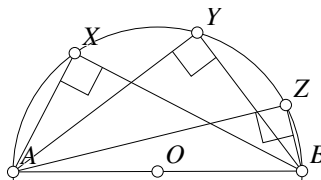
Figure A.28: Inscribed and Central Angles
`Review/Circprops2.D [D]`
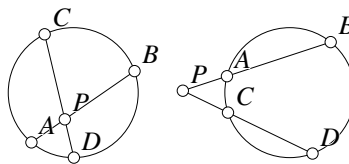
Figure A.29: Intersecting Secants
`Review/Circprops3.D [D]`

- If a secant cuts a circle at points A and B and another secant cuts the same circle at points C and D, and if the two secants intersect at a point P, then $PA \cdot PB = PC \cdot PD$. The point P can be inside or outside the circle. Figure A.29 shows both situations with P, the intersection of secants AB and CD, lying either inside or outside the circle.

- In figure A.29 on the right, $2\angle APC = \overset{\frown}{DB} - \overset{\frown}{AC}$. In other words, the angle is half the difference of the measures of the subtended arcs. A **Geometer** proof of this can be found in `Proofs/CircleProperties.T`.
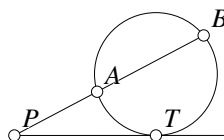
Figure A.30: Tangent and Secant
`Review/Circprops4.D [D]`

- If a secant passing through a circle at points A and B intersects at point P a line that is tangent to the circle at point T, then $PA \cdot PB = PT^2$. (This is about the same as the previous item, where the points C and D happen to be the same point.) See figure A.30.

- If a line is tangent to a circle at T, and TA is any chord of the circle, then the measure of the arc cut off by the chord is twice the measure of the angle made by the two chords. In other words, the central angle $\angle AOT$ in figure A.31 is twice the angle $\angle ATB$ at the tangent: $\angle AOT = 2\angle ATB$. A **Geometer** proof appears in `Proofs/TangAngle.T`.
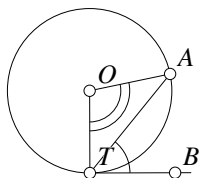
Figure A.31: Tangent and Chord
`Review/Circprops5.D [D]`

A.10 Trigonometric Definitions

The word "trigonometry", of course, means "triangle measurement". Its practical applications are related to making actual measurements—surveyors, for example, can measure some angles and distances and from those compute others. Trigonometry is primarily concerned with a few mathematical functions that are defined by the lengths and angles in right triangles.

A right triangle contains a right angle, so the other two angles add to $90°$. That means any two right triangles that share another angle are similar by AA, giving us all the similarity relationships that relate the lengths of the sides. In a right triangle, if you know one of the other angles and the length of any side, you can determine all the the other angles and lengths.

As a concrete example let us consider all right triangles that also contain an angle α. The ratios of the side opposite the angle α and the hypotenuse will be the same, independent of the size of the triangle. In trigonometry, that particular ratio is given the name $\sin \alpha$, where α is the angle. The particular example of $\alpha = 30°$ will always have a hypotenuse twice as long as the side opposite the $30°$ angle, so $\sin 30° = 1/2$.

The other common trigonometric functions are given by other ratios of lengths of sides in any right triangle that contains the given angle.

At first it seems quite limiting to consider only right triangles but that is not the case. Not only is it possible to convert any triangle to a pair of right triangles by dropping the appropriate altitude, but the standard trigonometric functions appear in all sorts of other formulas. There are many more trigonometric relations than are listed in this section.

A.10.1 Definitions of the Trigonometric Functions

In figure A.32, $\angle C$ is a right angle. The definitions of all the trigonometric ratios for the angle $\angle BAC = \alpha$ in terms of the lengths of the sides a, b, and c appear in Table A.3.

In old texts one sometimes encounters a few other trigonometric functions that are easily expressed in terms of the basic ones shown in Table A.3. These include the **versine** (vers $\alpha = 1 - \cos \alpha = 2 \sin^2(\alpha/2)$), the **coversine** (covers $\alpha = 1 - \sin \alpha$), the **haversine** (hav $\alpha = (\text{vers } \alpha)/2 = (1 - \cos \alpha)/2 = \sin^2(\alpha/2)$), and the **exsecant**
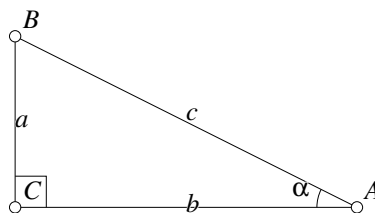
Figure A.32: Trigonometric Ratios
`Review/Trig.D` [D]

sine:	$\sin\alpha = a/c$	cosecant:	$\csc\alpha = c/a = 1/\sin\alpha$
cosine:	$\cos\alpha = b/c$	secant:	$\sec\alpha = c/b = 1/\cos\alpha$
tangent:	$\tan\alpha = a/b$	cotangent:	$\cot\alpha = b/a = 1/\tan\alpha$

Table A.3: Definitions of the Basic Trigonometric Functions

(exsec $\alpha = \sec\alpha - 1$). It is an interesting exercise to try to figure out why these definitions might be important and to determine trigonometric relations among them.

For many practical applications, the definitions in Table A.3 are sufficient, although the drawing only makes sense for the trigonometric functions of angles that are between $0°$ and $90°$. For completeness, here is the official mathematical definition of the trigonometric functions for any angle. The official definition also has the advantage that it is much easier to work with.
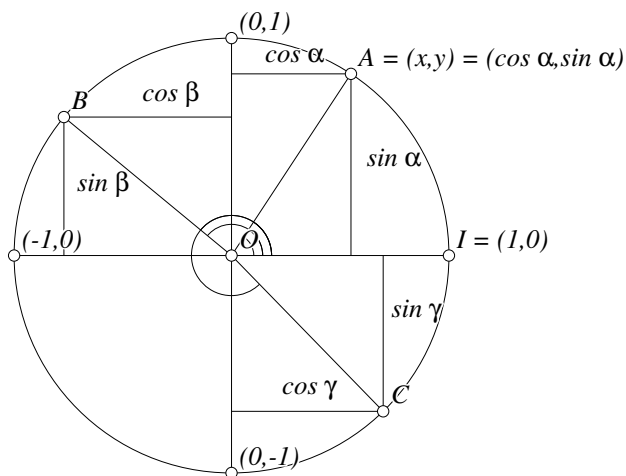
Figure A.33: Trigonometric Function Definitions
`Review/Trig1.D` [D]

In figure A.33, we see the so-called **unit circle** drawn in the standard x-y coordi-

nate system. The unit circle is centered at the origin $(0, 0)$, and has radius 1. Three examples of angles are shown in the figure, each centered at the origin, and each beginning on the positive x-axis. Angle α is the one marked with three concentric rings; angle β includes angle α and in addition the angle marked with two concentric rings, and angle γ includes all of β and the additional angle marked with one ring. In the notation we usually use, angles α, β, and γ are equal to $\angle IOA$, $\angle IOB$, and $\angle IOC$, respectively, where the points A, B, and C are on the unit circle, and the angles are measured counterclockwise from I, so $\angle IOC$ looks as if it is roughly $315°$.

Since the signs of the trigonometric functions change only when an angle changes from one quadrant to another, it is often convenient to give the quadrants names, and Roman numerals are used. The quadrant where $x > 0$ and $y > 0$ is called quadrant I, and the next three, in counter-clockwise order, are quadrants II, III and IV.

Every point on the unit circle determines an angle, and the trigonometric functions are simply defined in terms of the coordinates of those points. For example, $\sin \alpha$ is the y-coordinate of the point A and $\cos \alpha$ is its x-coordinate. For angle α, it is easy to compare these definitions with the ones above; namely, the sine of an angle is the length of the side opposite that angle divided by the hypotenuse. The side opposite angle α is the y-coordinate of point A, and the hypotenuse is 1, since A is on a unit circle that is defined to have radius 1.

In general, if we have an angle like α centered at the origin and whose first edge goes through $I = (1, 0)$ and whose other edge goes through point $A = (x, y)$, then we can define the trigonometric functions using Table A.4.

sine:	$\sin \alpha = y$	cosecant:	$\csc \alpha = 1/y = 1/\sin \alpha$
cosine:	$\cos \alpha = x$	secant:	$\sec \alpha = 1/x = 1/\cos \alpha$
tangent:	$\tan \alpha = y/x$	cotangent:	$\cot \alpha = x/y = 1/\tan \alpha$

Table A.4: Alternative Definitions of the Trigonometric Functions

There are advantages to this definition over the previous one. First, it is much easier to work with coordinates, and second, this definition is valid for all angles—not only those that are bigger than $90°$, but for negative angles, and for angles larger than $360°$. Of course the trigonometric functions for an angle like $390°$ are not terribly interesting since it is the same as going around the circle once completely and then $30°$ more, so every trigonometric function will have the same value for $390°$ as for $30°$: $\sin 390° = \sin 30° = 1/2$; $\cos 390° = \cos 30° = \sqrt{3}/2$, $\sec(\theta + 360°) = \sec \theta$, et cetera.

To better visualize the trigonometric functions, load the file `Review/TrigFuncs.T` and press the *Run Script* button. Each of the six basic trigonometric functions is represented by a segment of a different color, and as the script runs, the angle cycles around the whole circle and the segments change length accordingly. This demonstration, in particular, makes it obvious what is meant when we say that the tangent (or cotangent or secant or cosecant) "goes to infinity", why sine and cosine are always between -1 and 1, et cetera.

Notice that if the angles are between $90°$ and $180°$ (in other words, if the angle lies in quadrant II), the x-coordinates are negative, so the cosines and tangents of those

angles will be negative. If the angle is between $180°$ and $270°$ (in quadrant III), both coordinates are negative, making the sine and cosine negative (but the tangent positive, since it is the quotient of two negative numbers).

There are a lot of other things you can see from the figure, such as $\sin \alpha = \sin(180° - \alpha)$. This is because if you measure an angle counterclockwise from $(1, 0)$ you come to a point that has the same y-coordinate as you get if you measure that same angle clockwise from $(-1, 0)$ (which is how you get to the angle $180° - \alpha$). Similarly, the pythagorean theorem shows instantly that $\sin^2 \alpha + \cos^2 \alpha = 1$ since the sine and cosine are the x- and y-coordinates of a point on a circle of radius 1.

Notice that although the sine and cosine always make sense, there are times when the tangent, cotangent, secant, and cosecant do not; namely, when the denominators in their respective definitions are zero. Sometimes we just write something like $\tan 90° = \infty$ to indicate that as the angle θ gets closer and closer to $90°$, $\tan \theta$ gets larger without bound. Be careful, however. If θ gets close to $90°$ from below, $\tan \theta$ does get larger and larger, but if θ approaches $90°$ from above, then $\tan \theta$ gets to be a larger and larger *negative* number. The same sort of thing holds true for all the other basic trigonometric functions that "go to infinity".

A.10.2 ♦ Key Theorems from Trigonometry

A.10.2.1 Some Concrete Trigonometric Values

Here is a short table that lists some exact values of the trigonometric functions for a few interesting angles. When a trigonometric function is undefined as described in the final paragraph in the previous section, ∞ appears in the table, so be sure to remember what that means.

θ	$\sin \theta$	$\cos \theta$	$\tan \theta$	$\cot \theta$	$\sec \theta$	$\csc \theta$
$0°$	0	1	0	∞	1	∞
$30°$	$1/2$	$\sqrt{3}/2$	$\sqrt{3}/3$	$\sqrt{3}$	$2\sqrt{3}/3$	2
$45°$	$\sqrt{2}/2$	$\sqrt{2}/2$	1	1	$\sqrt{2}$	$\sqrt{2}$
$60°$	$\sqrt{3}/2$	$1/2$	$\sqrt{3}$	$\sqrt{3}/3$	2	$2\sqrt{3}/3$
$90°$	1	0	∞	0	∞	1

♦ Sometimes it is useful to know a few other key values. To save space, only the sine and cosine values are listed; you can calculate the other values easily from these. Note that many other trigonometric values for integer angles can be computed from the tables above and below and using the formulas in Section A.10.2.2. For example:

$$\sin(3°) = \sin(18° - 15°) = \sin(18°) \cos(15°) - \cos(18°) \sin(15°).$$

θ	$\sin\theta$	$\cos\theta$
$15°$	$\frac{1}{4}(\sqrt{6}-\sqrt{2})$	$\frac{1}{4}(\sqrt{6}+\sqrt{2})$
$18°$	$\frac{1}{4}(\sqrt{5}-1)$	$\frac{1}{4}\sqrt{10+2\sqrt{5}}$
$36°$	$\frac{1}{4}\sqrt{10-2\sqrt{5}}$	$\frac{1}{4}(1+\sqrt{5})$
$54°$	$\frac{1}{4}(1+\sqrt{5})$	$\frac{1}{4}\sqrt{10-2\sqrt{5}}$
$72°$	$\frac{1}{4}\sqrt{10+2\sqrt{5}}$	$\frac{1}{4}(\sqrt{5}-1)$
$75°$	$\frac{1}{4}(\sqrt{6}+\sqrt{2})$	$\frac{1}{4}(\sqrt{6}-\sqrt{2})$

A.10.2.2 Basic Trigonometric Identities

Without proof (and it is a great exercise to do the proofs), here is a list of equations satisfied by various of the trigonometric functions:

If $\alpha + \beta = 90°$ (if α and β are complementary[8]):

$$\sin\alpha = \cos\beta \qquad \tan\alpha = \cot\beta \qquad \sec\alpha = \csc\beta.$$

For supplementary angles we have:

$$\sin\alpha = \sin(180° - \alpha) \qquad \csc\alpha = \csc(180° - \alpha)$$
$$\cos\alpha = -\cos(180° - \alpha) \qquad \sec\alpha = -\sec(180° - \alpha)$$
$$\tan\alpha = -\tan(180° - \alpha) \qquad \cot\alpha = -\cot(180° - \alpha)$$

These are direct consequences of the fact that the trigonometric functions are defined in terms of coordinates of points on the unit circle:

$$\sin^2\alpha + \cos^2\alpha = 1 \qquad \sec^2\alpha = \tan^2\alpha + 1 \qquad \csc^2\alpha = \cot^2\alpha + 1$$

Here are the angle addition formulas[9]:

$$\sin(\alpha \pm \beta) = \sin\alpha\cos\beta \pm \cos\alpha\sin\beta \tag{A.5}$$
$$\cos(\alpha \pm \beta) = \cos\alpha\cos\beta \mp \sin\alpha\sin\beta \tag{A.6}$$
$$\tan(\alpha \pm \beta) = \frac{\tan\alpha \pm \tan\beta}{1 \mp \tan\alpha\tan\beta} \tag{A.7}$$

See the diagram `Proofs/CosineSummation.T` for a **Geometer** proof of the sum of angles formula for the cosine.

---

[8]Note the letters "*co*" in the words: sine–*co*sine, tangent–*co*tangent, secant–*co*secant, and *co*mplementary. Coincidence? Not likely.

[9]The symbol $\mp$ is the opposite of $\pm$. It means that if you choose the $+$ option of $\pm$, you must use the $-$ option in $\mp$ and vice-versa. The use of $\pm$ and $\mp$ in equation A.6 is equivalent to:

$$\cos(\alpha + \beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$$
$$\cos(\alpha - \beta) = \cos\alpha\cos\beta + \sin\alpha\sin\beta$$

From these, it is easy to derive the multiple angle formulas since $\sin(2\alpha) = \sin(\alpha + \alpha)$, $\sin(3\alpha) = \sin(2\alpha + \alpha)$, et cetera[10]:

$$\sin(2\alpha) = 2\sin\alpha\cos\alpha \tag{A.8}$$

$$\cos(2\alpha) = \cos^2\alpha - \sin^2\alpha = 2\cos^2\alpha - 1 \tag{A.9}$$

$$\tan(2\alpha) = \frac{2\tan\alpha}{1 - \tan^2\alpha} \tag{A.10}$$

$$\sin(3\alpha) = 3\sin\alpha - 4\sin^3\alpha \tag{A.11}$$

$$\cos(3\alpha) = 4\cos^3\alpha - 3\cos\alpha \tag{A.12}$$

$$\tan(3\alpha) = \frac{3\tan\alpha - \tan^3\alpha}{1 - 3\tan^2\alpha} \tag{A.13}$$

From these equations, it is easy to derive the following useful half-angle formulas[11]:

$$\sin(\alpha/2) = \pm\sqrt{\frac{1 - \cos\alpha}{2}} \tag{A.14}$$

$$\cos(\alpha/2) = \pm\sqrt{\frac{1 + \cos\alpha}{2}} \tag{A.15}$$

$$\tan(\alpha/2) = \frac{\sin\alpha}{1 + \cos\alpha} = \frac{1 - \cos\alpha}{\sin\alpha} \tag{A.16}$$

Here are a few equations that are occasionally very useful:

$$\sin\alpha + \sin\beta = 2\sin\left(\frac{\alpha + \beta}{2}\right)\cos\left(\frac{\alpha - \beta}{2}\right) \tag{A.17}$$

$$\sin\alpha - \sin\beta = 2\cos\left(\frac{\alpha + \beta}{2}\right)\sin\left(\frac{\alpha - \beta}{2}\right) \tag{A.18}$$

$$\cos\alpha + \cos\beta = 2\cos\left(\frac{\alpha + \beta}{2}\right)\cos\left(\frac{\alpha - \beta}{2}\right) \tag{A.19}$$

$$\cos\alpha - \cos\beta = -2\sin\left(\frac{\alpha + \beta}{2}\right)\sin\left(\frac{\alpha - \beta}{2}\right) \tag{A.20}$$

$$\tan\alpha + \tan\beta = \frac{\sin(\alpha + \beta)}{\cos\alpha\cos\beta} \tag{A.21}$$

$$\tan\alpha - \tan\beta = \frac{\sin(\alpha - \beta)}{\cos\alpha\cos\beta} \tag{A.22}$$

From those we can also derive the following related formulas:

$$\sin\alpha\cos\beta = \left(\sin(\alpha - \beta) + \sin(\alpha + \beta)\right)/2 \tag{A.23}$$

$$\cos\alpha\cos\beta = \left(\cos(\alpha - \beta) + \cos(\alpha + \beta)\right)/2 \tag{A.24}$$

$$\sin\alpha\sin\beta = \left(\cos(\alpha - \beta) - \cos(\alpha + \beta)\right)/2 \tag{A.25}$$

---

[10]The formulas for $\cos n\alpha$ are closely related to the Chebychev polynomials: If $x = \cos\alpha$, $T_n(x) = \cos n\alpha$. $T_0(x) = 1$, $T_1(x) = x$, $T_2(x) = 2x^2 - 1$, $T_3(x) = 4x^3 - 3x$, et cetera. It is an interesting exercise for a student to investigate these polynomials.

[11]Depending on the size of α in formulas A.14 and A.15, the positive or negative value of the square root must be used.

A.10.2.3 Area of a Triangle

One trivial application of trigonometry to geometry is a simple formula for the area of a triangle. If a triangle has two sides of lengths a and b, and the included angle is θ, then the area of the triangle is given by $(ab \sin \theta)/2$.

Figure A.34 shows why. Drop a perpendicular line from vertex B to line AC, intersecting AB at point D. By simple trigonometry, the length of BD, the altitude of the triangle, is $a \sin \theta$. The area of the triangle is half the base times the height, or $(ab \sin \theta)/2$.
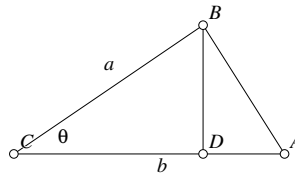
Figure A.34: Area of a Triangle
`Review/Triarea.T` [P]

A.10.2.4 The Law of Sines

The law of sines is one of the most useful theorems that relates the trigonometric functions to arbitrary triangles.

Theorem A.3 (Law of Sines) *In any triangle $\triangle ABC$,*

$$\frac{BC}{\sin A} = \frac{CA}{\sin B} = \frac{AB}{\sin C}.$$

In figure A.35, construct the circle that circumscribes the given $\triangle ABC$ by finding its circumcenter O. Next, construct the diameter AD through O, and connect point D with point B. Since AD is the diameter of the circle, $\angle ABD$ is a right angle, and hence $\sin \angle ADB = AB/AD$. $\angle ADB = \angle ACB = \angle C$ since both of them are inscribed in the same arc $\overset{\frown}{AB}$. Hence $AD = AB/\sin C$. A **Geometer** proof of the law of sines appears in `Proofs/Sinethm.T`.

Figure A.35: Law of Sines
`Review/Sinethm.T` [P]

Exactly the same argument can be made about every other angle in the triangle—that the side opposite that angle divided by the sine of the angle is the diameter of the circumscribed circle. But that diameter is constant, so all three quotients must be equal, and the law of sines is proved.

Actually the law of sines often includes one more equality as follows:

$$\frac{BC}{\sin A} = \frac{CA}{\sin B} = \frac{AB}{\sin C} = 2R,$$

where R is the radius of the circumcircle for $\triangle ABC$. This is easy to see using the same figure. $\angle ACB = \angle ADB$ since they both subtend the same arc, but $\triangle ABD$ is a right triangle, so $\angle ABD = 90°$. Thus

$$\frac{AB}{\sin C} = \frac{AB}{\sin D} = \frac{AD}{\sin 90°} = \frac{AD}{1} = AD = 2R.$$

A.10.2.5 Angle Bisectors in Triangles

The law of sines gives a nice, useful result about angle bisectors in a triangle. In figure A.36, we see $\triangle ABC$, where the line AD bisects $\angle BAC$. Let $\angle DAB = \angle CAD = \alpha$ and $\angle ADB = \beta$, so we know that $\angle ADC = 180° - \beta$.

Two applications of the law of sines give:

$$\begin{aligned}
\frac{\sin \alpha}{DB} &= \frac{\sin \beta}{AB} \\
\frac{\sin \alpha}{DC} &= \frac{\sin(180° - \beta)}{AC} = \frac{\sin \beta}{AC}.
\end{aligned}$$

Figure A.36: Angle Bisector Property
`Review/Anglebisect1.T [P]`

Since $\sin \beta = \sin(180° - \beta)$, we get:

$$\frac{DB}{AB} = \frac{\sin \alpha}{\sin \beta} = \frac{DC}{AC}.$$

There is another easy proof of this result that does not use trigonometry. Extend the line AD and find point E on the extension such that $CD = CE$. Then $\triangle CDE$ is isosceles and from that it is easy to show that $\triangle ADB \sim \triangle AEC$.

A.10.2.6 The Law of Cosines

Theorem A.4 (Law of Cosines) *In any $\triangle ABC$, let a, b and c represent the lengths of the sides opposite angles A, B and C, respectively, as in figure A.37. (In the figure, the side of length b is broken into two parts having lengths l and $b - l$.) Then*

$$c^2 = a^2 + b^2 - 2ab \cos C.$$

Although the law of cosines is true in general, let us just prove it for the special case where $\angle C$ is acute. A tiny modification of the proof will show that the theorem is also true for obtuse angles. If $\angle C$ is a right angle, then $\cos 90° = 0$, so the law of cosines reduces to the pythagorean theorem: $c^2 = a^2 + b^2$.

As shown in figure A.37, drop a perpendicular from B to the line AC meeting it at point D. Let D divide the side AC into two segments of lengths l and $b - l$. Denote the length of BD by h.

Using the pythagorean theorem twice on triangles $\triangle ABD$ and $\triangle BDC$, we obtain:

$$c^2 = h^2 + (b - l)^2$$

and

$$a^2 = h^2 + l^2.$$

Figure A.37: Law of Cosines
`Review/Costhm.D [D]`

If we eliminate h^2 from these two equations, we get:

$$c^2 = a^2 - l^2 + b^2 - 2bl + l^2 = a^2 + b^2 - 2bl.$$

But $l = a \cos C$, so we have:

$$c^2 = a^2 + b^2 - 2ab \cos C,$$

completing the proof for C an acute angle. Be sure to complete the proof for the case where C is obtuse.

The law of cosines looks quite similar to the pythagorean theorem, and in fact, except for the term $-2ab \cos C$, it is identical.

Here are some slightly oddball formulas that, like the laws of sines and cosines, relate the lengths of the sides of a triangle with some trigonometric functions of the angles[12]. Remember that the equations below apply to any sides or angles in a triangle

---

[12]Since they are unusual, if you are teaching a trigonometry class, these may provide some examination or extra credit questions.

so each of the equations below could be rewritten in two additional forms.

$$a \quad = \quad b\cos C + c\cos B \tag{A.26}$$

$$\frac{a+b}{a-b} \quad = \quad \frac{\tan((A+B)/2)}{\tan((A-B)/2)} \tag{A.27}$$

$$\frac{a+b}{c} \quad = \quad \frac{\cos((A-B)/2)}{\sin(C/2)} \tag{A.28}$$

$$\frac{a-b}{c} \quad = \quad \frac{\sin((A-B)/2)}{\cos(C/2)} \tag{A.29}$$

$$\sin(A/2) \quad = \quad \sqrt{\frac{(s-b)(s-c)}{bc}} \tag{A.30}$$

$$\cos(A/2) \quad = \quad \sqrt{\frac{s(s-a)}{bc}} \tag{A.31}$$

$$\tan(A/2) \quad = \quad \sqrt{\frac{(s-b)(s-c)}{s(s-a)}} \tag{A.32}$$

$$\tan((A-B)/2) \quad = \quad \left(\frac{a-b}{a+b}\right)\cot(C/2). \tag{A.33}$$

$\boxed{\text{A.11}}$ Coordinate Geometry

Applying coordinates to points and then simply using algebraic manipulation can be a powerful technique to solve problems that seem totally geometric. The calculations often become quite unwieldy, however. Today that is less of a problem since there are good, commercially available computer algebra programs such as Maple® and Mathematica®. A description of the use of such programs is beyond the scope of this book.

Coordinate geometry (often called analytic geometry) provides a method to convert statements about geometry into statements about algebra, and then to do algebraic manipulations to arrive at results.

The idea is simple—just draw the geometric figure in a standard Cartesian coordinate system where algebraic variables are used for variable points. Then any standard calculations from algebra are allowed.

Coordinate geometry is based on the pythagorean theorem that allows you to calculate the distance between a pair of points if you are given their coordinates.

Figure A.38: Distance: Point to Point

`Review/Coord1.D [D]`

In figure A.38 coordinates have been assigned to the plane, and we can see that the point A has coordinates $(2,2)$ and those of B are $(6,5)$. What is the distance between

points A and B? The segment AB is the hypotenuse of a right triangle whose right angle is at the point $P = (6, 2)$. The horizontal distance from P to A is $6 - 2 = 4$. Similarly, the vertical distance from P to B is $5 - 2 = 3$. In the right triangle $\triangle ABP$, we have $PA^2 + PB^2 = AB^2$, or $3^2 + 4^2 = AB^2$, or $AB = \sqrt{3^2 + 4^2} = 5$.

In general, if the coordinates of any two points A and B are (x_A, y_A) and (x_B, y_B), then the length of the segment AB is $\sqrt{(x_A - x_B)^2 + (y_A - y_B)^2}$.

There is a fair amount of machinery that can be used to do coordinate geometry proofs. To do coordinate geometry effectively you need to know the equations of lines, of circles, how to calculate intersections of lines and circles, the conditions on the slopes of lines for them to be perpendicular, et cetera. All of these topics are covered in a standard high school algebra course so without derivation, a list of the most useful formulas can be found in Table A.5.
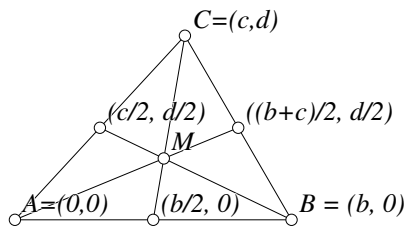
Figure A.39: Analytic Proof of Three Medians Theorem

`Review/Coord2.D [D]`

As an illustration of the technique, we will assume that you remember all of those formulas, and we will prove that the three medians of a triangle meet at a point that is $2/3$ of the distance from any vertex to the opposite side.

A nice thing about using coordinate geometry is that you can choose a coordinate system that is convenient. To prove the theorem, there is no need to orient the triangle in a totally arbitrary way (although the proof could still be done). It is easiest if we twist around our arbitrary triangle until one of its vertices is at the origin and the other is on the x-axis. In figure A.39 we have done just that. Point A is at the origin, B is at $(b, 0)$, and C is at (c, d).

The midpoints of the sides are easy to calculate—their coordinates are the averages of the coordinates of the two endpoints of the segments, so they are as shown in the figure: $(b/2, 0)$, $((b + c)/2, d/2)$, and $(c/2, d/2)$.

The equations of the median lines originating at A, B, and C, respectively are:

$$y = \frac{d}{b+c}x \tag{A.34}$$

$$y = \frac{d}{c - 2b}(x - b) \tag{A.35}$$

$$y = \frac{2d}{2c - b}(x - b/2) \tag{A.36}$$

$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$	Distance between the points (x_1, y_1) and (x_2, y_2).
$\left(\dfrac{x_1 + x_2}{2}, \dfrac{y_1 + y_2}{2}\right)$	Coordinates of the midpoint between the points (x_1, y_1) and (x_2, y_2).
$\dfrac{\|ax_1 + by_1\|}{\sqrt{a^2 + b^2 + c^2}}$	Distance between the point (x_1, y_1) and the line whose equation is $ax+by+c = 0$.
$ax + by + c = 0$	Completely general equation for a line (including vertical lines). Either a or b must be non-zero.
$y = mx + b$	Equation of a line that intersects the y-axis at b and has slope m.
$y - y_1 = m(x - x_1)$	Equation of a line passing through the point (x_1, y_1) and having slope m.
$y - y_1 = \left(\dfrac{y_2 - y_1}{x_2 - x_1}\right)(x - x_1)$	Equation of a line passing through the points (x_1, y_1) and (x_2, y_2).
$y - y_1 = \left(\dfrac{-1}{m}\right)(x - x_1)$	Equation of a line passing through the point (x_1, y_1) and perpendicular to a line having slope m.
$(x - x_1)^2 + (y - y_1)^2 = r^2$	Equation of a circle with center (x_1, y_1) and radius r.
$ax^2 + bxy + cy^2 + dx + ey + f = 0$	Completely general equation of a conic section.

Table A.5: Useful formulas for coordinate geometry

Solving equations A.34 and A.35 simultaneously:

$$
\begin{aligned}
\frac{d}{b+c}x &= \frac{d}{c-2b}(x-b) \\
\frac{1}{b+c}x &= \frac{1}{c-2b}(x-b) \\
(c-2b)x &= (b+c)(x-b) \\
(c-2b-b-c)x &= -b(b+c) \\
-3bx &= -b(b+c) \\
x &= \frac{b+c}{3}
\end{aligned}
$$

plug this result into equation A.34:

$$
y = \left(\frac{d}{b+c}\right)\left(\frac{b+c}{3}\right) = \frac{d}{3}.
$$

If we solve another pair of equations, say equations A.35 and A.36, we find that we get exactly the same result, so the coordinates of the intersections of any pair of lines is the same, and we can give a single name M to that point. $M = ((b+c)/3, d/3)$.

Finally, to show that M is $2/3$ of the distance between any vertex of the triangle and its opposite side, use the pythagorean theorem to write down the distances, and then use algebra to show that the distances are equal. $MA = \sqrt{((b+c)/3)^2 + (d/3)^2}$ and the length of the median from A is $\sqrt{((b+c)/2)^2 + (d/2)^2}$. We need to show that MA is $2/3$ the length of the median, or that

$$
\sqrt{((b+c)/3)^2 + (d/3)^2} = (2/3)\sqrt{((b+c)/2)^2 + (d/2)^2}.
$$

Square both sides and expand:

$$
(b^2 + 2bc + c^2 + d^2)/9 = (4/9)(b^2 + 2bc + c^2 + d^2)/4,
$$

which is clearly true.

Do the same calculation for the other medians, and that will complete the proof.

This seems quite ugly, but at least it is almost completely mechanical. Geometric calculations done by computers are done in exactly this way, and in fact whenever you are dragging around points in **Geometer**, hundreds of equations very similar to those above are being solved by the computer over and over.

A.11.1 Analytic Area of a Polygon

Theorem A.5 (Area of a Polygon) *If the planar coordinates of the vertices of an n-sided simple polygon are given by (x_0, y_0), (x_1, y_1), ..., (x_{n-1}, y_{n-1}), then if we agree to define $(x_n, y_n) = (x_0, y_0)$, the area of the polygon is given by the formula:*

$$
\mathcal{A} = \frac{1}{2}\sum_{i=0}^{n-1}\left(x_i y_{i+1} - x_{i+1} y_i\right).
$$

This formula gives a "signed" area in the sense that the value will be positive if the points trace the polygon in a counterclockwise order and negative otherwise.

The formula holds for non-convex polygons, but does not hold if the edges of the polygon cross each other.

We will complete the proof in a series of steps.

Lemma A.1 (Area of a Triangle with one vertex at the origin) *If a triangle has vertices with coordinates $(0,0)$, (x_0, y_0), and (x_1, y_1) then its (signed) area is given by $(x_0 y_1 - x_1 y_0)/2$.*
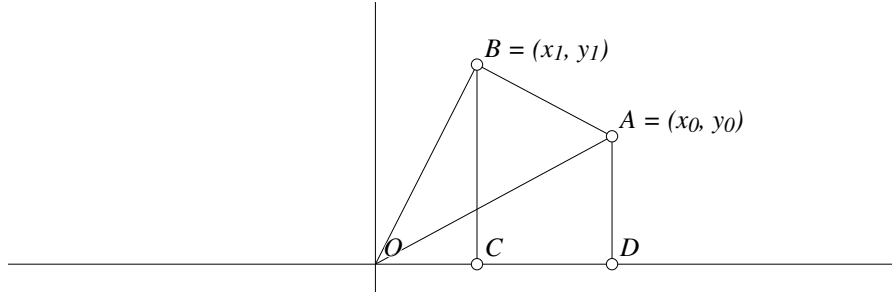
Figure A.40: Area of Triangle at Origin
`Review/Triorigin.D` [D]

See Figure A.40. We would like to find the area of $\triangle OAB$ where A has coordinates (x_0, y_0) and B has coordinates (x_1, y_1). We know that

$$\mathcal{A}(\triangle OAB) = \mathcal{A}(\triangle OCB) + \mathcal{A}(CDAB) - \mathcal{A}(\triangle ODA).$$

It is easy to write down the areas of the two triangles and the trapezoid $CDAB$ on the right side of the equation:

$$
\begin{aligned}
\mathcal{A}(\triangle OAB) &= \frac{x_1 y_1}{2} + \frac{(x_0 - x_1)(y_0 + y_1)}{2} - \frac{(x_0 y_0)}{2} \\
&= \frac{x_1 y_1}{2} + \frac{x_0 y_0 + x_0 y_1 - x_1 y_0 - x_1 y_1}{2} - \frac{x_0 y_0}{2} \\
&= \frac{x_0 y_1 - x_1 y_0}{2}.
\end{aligned}
$$

Notice that this formula is a signed area. Here we visit the points O, A, and B in a counterclockwise order and obtain a positive area. For a clockwise ordering, the area is negative since the area of the trapezoid will have the term $(x_1 - x_0)$ instead of $(x_0 - x_1)$, and the triangle whose area was added is subtracted and vice-versa.

With this formula for the (signed) area of a triangle, we can proceed to a proof of the final theorem.
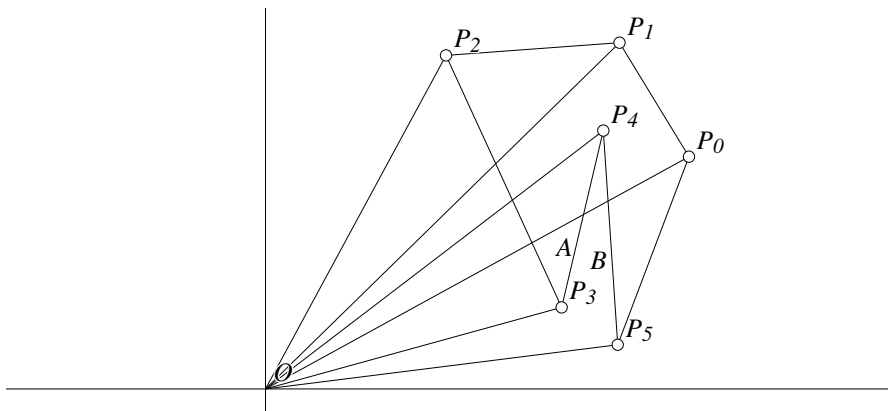
Figure A.41: Area of a General Polygon
`Review/Polyarea.D` [D]

We will not give a completely rigorous proof, but it should be obvious what such a proof would look like from our examination of Figure A.41 which shows a polygon with 6 vertices: $P_0, P_1, \ldots, P_5$. Assume the coordinates are given by $P_i = (x_i, y_i)$.

Draw a segment from each vertex P_i to the origin. From the lemma, we have a formula for the area of $\triangle OP_0P_1$, for $\triangle OP_1P_2$, and so on, up to $\triangle OP_5P_0$. Adding together the formulas for the individual polygons, and keeping in mind the signed areas, it is easy to see that after all the additions and subtractions, the net result will be whatever is contained within the polygon $P_0P_1P_2P_3P_4P_5$.

In this example, the areas of the triangles $\triangle OP_0P_1$, $\triangle OP_1P_2$, $\triangle OP_3P_4$, and $\triangle OP_5P_0$ will be positive, and the areas of the other two will be negative.

The maze of lines heading back to the origin break up the plane into a bunch of regions; let us just look at a couple of them, labeled A and B in the figure. Region A should be inside the polygon. Its area is counted twice positively since it is inside $\triangle OP_3P_4$ and $\triangle OP_5P_0$, but its area is subtracted once since it is inside the clockwise-oriented $\triangle OP_4P_5$. Thus its area is counted once, as it should be.

Region B's area is counted once positively since it is inside $\triangle OP_5P_0$ and once negatively since it is inside $\triangle OP_4P_5$. Its net contribution is zero which is correct since it is outside the polygon. Check some of the other regions to see that the same counting scheme works.

To see why it works in general, imagine one of these regions in the plane. We would like to count how many times its area is included in the grand sum, and see how the positive and negative contributions of various triangles are combined.

Suppose the polygon does not include the origin, as in the figure. Imagine starting at some particular point in the region your are testing and follow a straight line back to the origin. Every time you cross a line it means that you have entered another triangle, and if you were inside the polygon before crossing, you will be outside afterwards and vice-versa. Each time you cross a line that enters the polygon, you must have crossed a line putting you inside a counterclockwise oriented triangle and every time you leave

the polygon, you enter a clockwise oriented polygon. Thus, there is an alternation of adding and subtracting until you get to the origin that you know is outside.

When you get to the origin, if you were initially outside, you will have crossed an even number of lines, so the alternation will guarantee that the grand sum of areas for your point will be zero; otherwise, the grand sum will be one—the area your point was within was counted exactly once.

If the origin happens to be inside the polygon, the only change to the argument is that an even number of crossings implies that you were initially inside.

Now, if we just add up the areas (remembering that $P_6 = P_0$), we obtain:

$$\mathcal{A}(P_0 P_1 P_2 P_3 P_4 P_5) = \frac{1}{2} \sum_{i=0}^{5} \Big(x_i y_{i+1} - x_{i+1} y_i \Big),$$

and of course there is nothing special about the 6 vertex case.

A.12 Vectors

Vectors are usually covered in high school mathematics courses, but only at a very superficial level—students generally have no opportunity to practice solving problems with them. For this reason, the coverage of vectors here will be slightly different than a review. It will consist of a review followed by a few problems solved using vector techniques.

Vectors provide an intermediate method for solving geometric problems that is partly symbolic and partly geometric. Problems approached with vector calculations are often easy to solve and check, but they can also be be translated back to a geometric picture. Vectors can be defined in any number of dimensions, but we will deal here with those in the two-dimensional plane.

There are two types of vectors, "bound" and "unbound". Both indicate a length and direction, but the bound sort also include a starting point. If we consider all the bound vectors with the same direction and length as being the same, the class of all such vectors is like a single unbound vector. Both types have their uses. If you simply need to show that two segments have the same length and are parallel, you need to show that they are the same unbound vector. If you would like to show that a point is at the midpoint of two others, it is easier to use bound vectors.

The same distinction is important in physical problems. Unbound vectors are used to describe relative movements, but if a vector represents a force, it is very different to apply a force to the center of a rod or to its tip. Force on the center will drag the rod; force on the tip will make it spin.

A.12.1 Unbound Vectors as Movements

A nice way to think of an unbound vector is as a description of a linear movement. A complete description requires only a direction and a distance. The starting position is

unimportant, so "go 5 kilometers northeast" is basically a vector. This instruction is independent of where you start, so could be applied to an object anywhere on the plane.
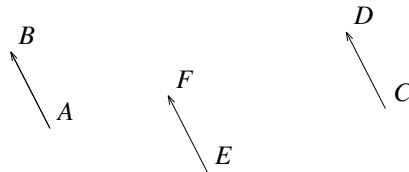
Figure A.42: Equivalent Vectors
`Review/Vectors.D [D]`

If vectors are movement descriptions, then it is easy to represent them as arrows. The direction of the arrow indicates the direction of motion and the length of the arrow indicates the distance to move. If A and B are two points in the plane (as in figure A.42), then the arrow with its "tail" at A and "tip" at B represents the vector whose length is the length of the segment AB and whose direction is from A toward B. We would write that particular vector as $\overrightarrow{AB}$. In figure A.42, $\overrightarrow{AB}$, $\overrightarrow{CD}$ and $\overrightarrow{EF}$ all represent *exactly* the same vector—they go in the same direction for the same distance. The starting point is unimportant.

Equivalently, if an arrow representing a vector is moved parallel to itself anywhere in the plane, it still represents the same vector.

In this book, we will always represent a vector with an arrow drawn over it. If we happen to know the points at the tail and tip, we will write it as $\overrightarrow{AB}$ where A is the tail and B is the tip. If we are just talking about a vector whose tip and tail may be unknown, we can still give it a name like $\overrightarrow{V}$. In some sense, the $\overrightarrow{AB}$ notation is a bit misleading since in reality it stands for that particular length and direction *anywhere* on the plane—not just from the particular point A to the particular point B.

If vectors are instructions for movement, then adding two vectors is the same as adding the instructions—just do the first movement followed by the second. If one vector is "go north 1 kilometer" and a second vector is "go west 1 kilometer", then to add them, first go north 1 kilometer and then turn west and go one more kilometer. The net result is a new vector: "go northwest $\sqrt{2} (= 1.4142135\ldots)$ kilometers." It should be clear that the combined instructions also indicate a distance and a direction, still independent of the starting point.

Adding vectors using the arrow representation is easy: to add vector $\overrightarrow{V_1}$ to vector $\overrightarrow{V_2}$, slide $\overrightarrow{V_2}$ parallel to itself until its tail coincides with the tip of $\overrightarrow{V_1}$. An arrow that connects the tail of $\overrightarrow{V_1}$ to the tip of $\overrightarrow{V_2}$ represents the sum $\overrightarrow{V_1} + \overrightarrow{V_2}$. This is illustrated in figure A.43.

When the vectors are placed tip to tail like this, it is obvious why the addition works since this method clearly combines the two commands for motion.

One possible movement description is "do not move". This is also a vector, the zero vector, sometimes written as $\overrightarrow{0}$. The zero vector is not the same as the number
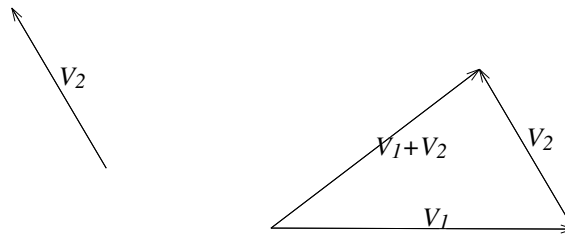
Figure A.43: Adding Vectors
Review/Vectors1.D [M]

zero—it is a vector representing a command not to move. It is clear that for any vector $\overrightarrow{V}$, we have: $\overrightarrow{V} + \overrightarrow{0} = \overrightarrow{0} + \overrightarrow{V} = \overrightarrow{V}$. Doing something then doing nothing is equivalent to just doing the something.

The negative of a vector is a vector going in the opposite direction for the same length. In other words, $\overrightarrow{AB} = -\overrightarrow{BA}$. This makes sense, since negation generally means "do the opposite", and the opposite of going from A to B is going from B to A. As usual, subtracting vectors is the same as adding the negative of a vector, so to calculate $\overrightarrow{V_1} - \overrightarrow{V_2}$ we add the negative of $\overrightarrow{V_2}$ to $\overrightarrow{V_1}$ to obtain $\overrightarrow{V_1} + (-\overrightarrow{V_2})$. Clearly every vector $\overrightarrow{V}$ satisfies $\overrightarrow{V} - \overrightarrow{V} = \overrightarrow{0}$.

It does not matter in what order you add vectors: $\overrightarrow{V_1} + \overrightarrow{V_2} = \overrightarrow{V_2} + \overrightarrow{V_1}$. In other words, it does not matter if you go north a kilometer and then east a kilometer, or if you first go east a kilometer and then north a kilometer.

Vectors can be multiplied (or divided) by real numbers: $3\overrightarrow{AB}$ is a vector in the same direction as $\overrightarrow{AB}$, but 3 times as long. Similarly, $\overrightarrow{AB}/2$ is again in the same direction, but only half as long. Multiplying by negative numbers reverses the direction as well, so $-2\overrightarrow{AB}$ is a vector twice as long as $\overrightarrow{AB}$ and going in the opposite direction.

Since for most people, it is not obvious, at first, how to apply these vector methods to geometric problems, we will begin with three examples worked in complete detail. Read through these, and then try some of the additional examples at the end of the chapter.

A.12.2   A Simple Vector Example

As a first example, we will use vectors to show that in any triangle $\triangle ABC$, the segment connecting the midpoints of AC and CB is parallel to and half the length of segment AB. See figure A.44.

Let M and N be the midpoints of AC and CB, respectively. Then in vector notation:
$$\overrightarrow{AB} = \overrightarrow{AM} + \overrightarrow{MN} + \overrightarrow{NB} .$$

We know that $\overrightarrow{AM}=\overrightarrow{AC}/2$ and $\overrightarrow{NB}=\overrightarrow{CB}/2$, so:

$$\overrightarrow{AB}=\overrightarrow{AC}/2+ \overrightarrow{MN} + \overrightarrow{CB}/2. \tag{A.37}$$

But we also know that $\overrightarrow{AB}=\overrightarrow{AC} + \overrightarrow{CB}$ so if we divide all the vectors by 2, we obtain:

$$\overrightarrow{AB}/2 =\overrightarrow{AC}/2+ \overrightarrow{CB}/2. \tag{A.38}$$

Combining equations A.37 and A.38 yields:

$$\overrightarrow{AB}=\overrightarrow{AB}/2+ \overrightarrow{MN},$$

or

$$\overrightarrow{AB}/2 =\overrightarrow{MN} .$$
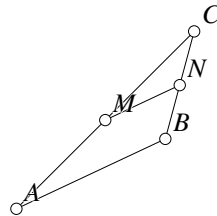
Figure A.44: Triangle Midpoints
`Review/Trianglemids.D` [M]

This proves the statement. Because of the vector equality, not only is the length of MN half of AB, but because they are vectors the directions are the same as well, so MN and AB are parallel.

A.12.3 A More Interesting Vector Example
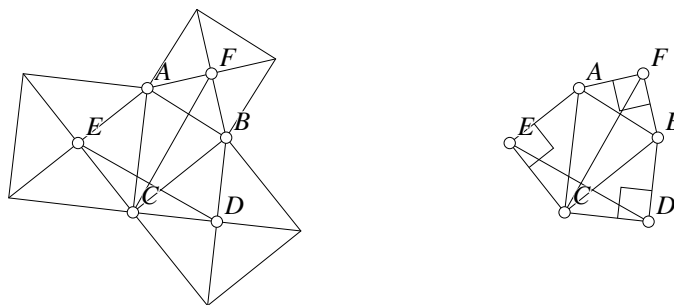
Figure A.45: Triangle Squares
`Review/Trisquares.D` [M]

Given $\triangle ABC$, construct a square on the outside of each side as in the drawing on the left in figure A.45. If the centers of those squares are D, E, and F as in the figure, prove that the segment FC is perpendicular to and the same length as segment ED.

On the right side of the same figure is a diagram with the extraneous lines removed. It contains isosceles right triangles erected on all the edges of $\triangle ABC$.

Here is a proof, using vectors. We begin with a few obvious identities:

$$\overrightarrow{ED} = \overrightarrow{EC} + \overrightarrow{CD} \tag{A.39}$$

$$\overrightarrow{ED} = \overrightarrow{EA} + \overrightarrow{AF} + \overrightarrow{FB} + \overrightarrow{BD} \tag{A.40}$$

$$\overrightarrow{FC} = \overrightarrow{FB} + \overrightarrow{BD} + \overrightarrow{DC} \tag{A.41}$$

$$\overrightarrow{FC} = \overrightarrow{FA} + \overrightarrow{AE} + \overrightarrow{EC} \tag{A.42}$$

There are clearly many routes to choose from E to D or from F to C, but the choices above all use lengths that are known to be equal to other lengths, due to the fact that the outer triangles are isosceles.

Adding equations A.39 to A.40 and A.41 to A.42, we obtain:

$$2\overrightarrow{ED} = \overrightarrow{EC} + \overrightarrow{CD} + \overrightarrow{EA} + \overrightarrow{AF} + \overrightarrow{FB} + \overrightarrow{BD} \tag{A.43}$$

$$2\overrightarrow{FC} = \overrightarrow{FB} + \overrightarrow{BD} + \overrightarrow{DC} + \overrightarrow{FA} + \overrightarrow{AE} + \overrightarrow{EC} \tag{A.44}$$

What we would like to show is that the vector $\overrightarrow{ED}$, when rotated $90°$ in a clockwise direction, is the same as vector $\overrightarrow{FC}$.

We can rotate the vector $2\overrightarrow{ED}$ in equation A.43 by rotating the six vectors on the right each by $90°$ and adding the results. Let $\rho(\overrightarrow{V})$ represent the result of rotating the vector $\overrightarrow{V}$ by $90°$ in a clockwise direction. Since the lengths $FA = FB$, $DB = DC$ and $EC = CA$, it is easy to rotate every term. We just have to be careful to get the directions right: $\rho(\overrightarrow{EC}) = \overrightarrow{AE}$, $\rho(\overrightarrow{CD}) = \overrightarrow{BD}$, et cetera. Continuing like this, we obtain:

$$\rho(2\overrightarrow{ED}) = \overrightarrow{AE} + \overrightarrow{BD} + \overrightarrow{EC} + \overrightarrow{FB} + \overrightarrow{FA} + \overrightarrow{DC}.$$

This is exactly the same as the formula for $2\overrightarrow{FC}$ in equation A.44 (with the vectors rearranged), so we are done.

A.12.4 Three Equilateral Triangles

In figure A.46, let $\triangle OAA'$, $\triangle OBB'$ and $\triangle OCC'$ be three equilateral triangles all described in a clockwise orientation. Let the points K, L and M be the midpoints, respectively, of segments $A'B$, $B'C$ and $C'A$. Show that $\triangle KLM$ is equilateral.

We can use the same general approach that we followed in Section A.12.3. We notice that if we can show that, say, $\overrightarrow{LM}$ when rotated $60°$ clockwise becomes $\overrightarrow{LK}$ we are done, since both with have the same length, and they meet in a $60°$ angle.

Here are some expressions for $\overrightarrow{LM}$ and $\overrightarrow{LK}$. These particular paths were chosen because they all involve only parts of the known equilateral triangles. It is certainly

true that $\overrightarrow{LM}=\overrightarrow{LK}+\overrightarrow{KM}$, but we do not know any of these lengths.

$$\overrightarrow{LM} = \overrightarrow{LC} + \overrightarrow{CC'} + \overrightarrow{C'M} \tag{A.45}$$

$$\overrightarrow{LM} = \overrightarrow{LB'} + \overrightarrow{B'O} + \overrightarrow{OA} + \overrightarrow{AM} \tag{A.46}$$

$$\overrightarrow{LK} = \overrightarrow{LB'} + \overrightarrow{B'B} + \overrightarrow{BK} \tag{A.47}$$

$$\overrightarrow{LK} = \overrightarrow{LC} + \overrightarrow{CO} + \overrightarrow{OA'} + \overrightarrow{A'K} \tag{A.48}$$

Adding equation A.45 to A.46 and A.47 to A.48, we obtain:

$$2\overrightarrow{LM} = \overrightarrow{LC} + \overrightarrow{CC'} + \overrightarrow{C'M} + \overrightarrow{LB'}$$
$$+ \overrightarrow{B'O} + \overrightarrow{OA} + \overrightarrow{AM} \tag{A.49}$$

$$2\overrightarrow{LK} = \overrightarrow{LB'} + \overrightarrow{B'B} + \overrightarrow{BK} + \overrightarrow{LC}$$
$$+ \overrightarrow{CO} + \overrightarrow{OA'} + \overrightarrow{A'K} \tag{A.50}$$

Figure A.46: Three Triangles
`Review/Threetris.D` [M]

Since L and K are the midpoints of the segments $B'C$ and $A'B$, $\overrightarrow{LC}= -\overrightarrow{LB'}$, $\overrightarrow{C'M}= -\overrightarrow{AM}$ and $\overrightarrow{BK}= -\overrightarrow{A'K}$. These equalities simplify equations A.49 and A.50 to:

$$2\overrightarrow{LM} = \overrightarrow{CC'} + \overrightarrow{B'O} + \overrightarrow{OA} \tag{A.51}$$

$$2\overrightarrow{LK} = \overrightarrow{B'B} + \overrightarrow{CO} + \overrightarrow{OA'} \tag{A.52}$$

Let $\rho(\overrightarrow{V})$ represent the vector obtained when $\overrightarrow{V}$ is rotated 60° in the clockwise direction, so from equation A.51 we obtain:

$$\rho(2\overrightarrow{LM}) = \rho(\overrightarrow{CC'}) + \rho(\overrightarrow{B'O}) + \rho(\overrightarrow{OA}). \tag{A.53}$$

It is easy to see that $\rho(\overrightarrow{CC'}) =\overrightarrow{CO}$, $\rho(\overrightarrow{B'O}) =\overrightarrow{B'B}$ and $\rho(\overrightarrow{OA}) =\overrightarrow{OA'}$. Substituting these into equation A.53 we can see that $\rho(2\overrightarrow{LM}) = 2\overrightarrow{LK}$ and we are done.

A.12.5 Touching Triangles

The following example makes use of vectors in the usual way, but uses the fact that the dot product of a pair of vectors is zero if and only if those two vectors are perpendicular.

The problem is this: Suppose two triangles, $\triangle XOZ$ and $\triangle POQ$ share a point O. Suppose further that $OZ = OQ$ and $OX = OP$ and that $OQ \perp OZ$ and $OX \perp OP$ as illustrated in figure A.47. If T is the midpoint of XZ, show that $OT \perp PQ$.

We will show that $OT \perp PQ$ by showing that the vector dot product: $\overrightarrow{OT} \cdot \overrightarrow{PQ} = 0$.

We know that $\overrightarrow{PQ} = \overrightarrow{PO} + \overrightarrow{OQ}$, that $\overrightarrow{OT} = (\overrightarrow{OX} + \overrightarrow{OZ})/2$, and that $\overrightarrow{OQ} \cdot \overrightarrow{OZ} = 0 = \overrightarrow{OP} \cdot \overrightarrow{OX}$ since they are perpendicular pairs.

Since we just want to show that the dot product $\overrightarrow{OT} \cdot \overrightarrow{PQ} = 0$, we can ignore the factor of 2 in the expansion of $\overrightarrow{OT}$ in the paragraph above, or equivalently, we can show that $2 \overrightarrow{OT} \cdot \overrightarrow{PQ} = 0$.

Figure A.47: Touching Triangles
Review/ShowPerp.T [M]

$$\overrightarrow{PQ} \cdot 2\,\overrightarrow{OT} = (\overrightarrow{PO} + \overrightarrow{OQ}) \cdot (\overrightarrow{OX} + \overrightarrow{OZ}).$$

When we multiply ot the expression on the right above, we obtain:

$$\overrightarrow{PO} \cdot \overrightarrow{OX} + \overrightarrow{PO} \cdot \overrightarrow{OZ} + \overrightarrow{OQ} \cdot \overrightarrow{OX} + \overrightarrow{OQ} \cdot \overrightarrow{OZ},$$

the first and last terms of which are zero.

It is also clear from the diagram that $\angle POZ = \angle XOQ$ since each is a right angle added to $\angle POQ$, so the two other dot products have equal values, given that the vectors have equal lengths and make equal angles with each other. The signs of those dot products, however, are opposite since $\overrightarrow{PO} = -\overrightarrow{OP}$. This proves that $OT \perp PQ$.

A.12.6 Bound Vectors and Coordinate Geometry

In this book we will consider all of our bound vectors to have their tail at the same place. If we want to use coordinate geometry, that same place will be the origin of our coordinate system: $(0,0)$.

If we wish to describe our bound vectors with coordinates and all of them have their tail at the origin, then each vector is completely determined by the coordinates of the tip. The point (x, y) thus stands for the bound vector whose tail is at the origin and whose tip is at the point (x, y).

If the vectors $\overrightarrow{V_1}$ and $\overrightarrow{V_2}$ are represented by the points (x_1, y_1) and (x_2, y_2), respectively, the sum or difference: $\overrightarrow{V_1} \pm \overrightarrow{V_2}$ is represented by the point $(x_1 \pm x_2, y_1 \pm y_2)$. Similarly, if α is a real number, then the vector $\alpha \overrightarrow{V_1}$ is represented by $(\alpha x_1, \alpha y_1)$. In particular, $-\overrightarrow{V_1}$ is represented by $(-x_1, -y_1)$. Finally, $(0, 0)$ represents the zero vector.

Following is a list of some properties of bound vectors. Some are important and will be used later; some are not particularly important in the future, but shed some

light on how vectors work. Sometimes the arrow notation is used and sometimes the coordinate notation, depending on which makes the property clearest.

- The average of two vectors $(\overrightarrow{A} + \overrightarrow{B})/2$ is a vector that goes from the origin to a point exactly half way between the tips of the other two vectors. Thus if a parallelogram has a vertex at the origin and two others at $\overrightarrow{A}$ and $\overrightarrow{B}$, then the vector $\overrightarrow{A} + \overrightarrow{B}$ represents the fourth vertex, and $(\overrightarrow{A} + \overrightarrow{B})/2$ is the the center of the parallelogram.

- If α is a real number and $\overrightarrow{A}$ and $\overrightarrow{B}$ are two vectors, then vectors of the form $(1-\alpha)\overrightarrow{A} + \alpha\overrightarrow{B}$ represent all the vectors whose tips lie on the line connecting the tips of $\overrightarrow{A}$ and $\overrightarrow{B}$. If $0 \le \alpha \le 1$ then the points lie between the tips of $\overrightarrow{A}$ and $\overrightarrow{B}$; otherwise, they lie outside. In fact, α represents the fraction of the distance the tip of the new vector is between $\overrightarrow{A}$ and $\overrightarrow{B}$. If $\alpha = 1/3$, then $(1-\alpha)\overrightarrow{A} + \alpha\overrightarrow{B} = (2/3)\overrightarrow{A} + (1/3)\overrightarrow{B}$ is $(1/3)$ of the way from $\overrightarrow{A}$ to $\overrightarrow{B}$ [13]. In the previous item, the average of the two vectors corresponds to $\alpha = 1/2$—half way between the two vectors.

- In $\triangle ABC$, if vectors $\overrightarrow{A}$, $\overrightarrow{B}$ and $\overrightarrow{C}$ represent the three vertices of the triangle, then $(\overrightarrow{A} + \overrightarrow{B} + \overrightarrow{C})/3$ represents the centroid of the triangle. See Section A.12.7.

- The length of the vector $\overrightarrow{A} = (x, y)$ is given by $\sqrt{x^2 + y^2}$. (This is just the pythagorean theorem applied to the triangle that includes the origin $(0, 0)$, the point $(x, 0)$, and the point (x, y). This right triangle has legs of lengths x and y, and the hypotenuse is the length of the vector.) An unbound vector described with tail (x_1, y_1) and head (x_2, y_2) has length $\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$.

- If (x, y) is a vector, then $(-y, x)$ is another vector of the same length and perpendicular to (x, y). In fact, $(-y, x)$ is the vector (x, y) rotated counterclockwise by $90°$.

- If you rotate the vector (x, y) counterclockwise by an angle θ, the resulting vector is $(x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta)$. The previous item is a special case of this, where $\theta = 90°$.

- Vectors can be defined in any number of dimensions. In one dimension, they simply represent directions to the left and to the right; a negative vector goes left and a positive vector goes right. Only a single number is needed to represent a one-dimensional vector. In three dimensions, three numbers are needed, and these represent the motion in the x, y, and z directions, where the z-axis is perpendicular to both the x-axis and to the y-axis.

---

[13]Do not get confused by the fact that the numbers seem to be reversed. $(2/3)\overrightarrow{A} + (1/3)\overrightarrow{B}$ is $2/3$ like $\overrightarrow{A}$ and $1/3$ like $\overrightarrow{B}$. It is more like $\overrightarrow{A}$ than $\overrightarrow{B}$, so it is closer to $\overrightarrow{A}$ than to $\overrightarrow{B}$. If you get confused, think of the extreme conditions where $\alpha = 0$ or $\alpha = 1$.

- If $\overrightarrow{A'}$ is the same as $\overrightarrow{A}$ rotated clockwise by an angle θ and $\overrightarrow{B'}$ is the same as $\overrightarrow{B}$ rotated clockwise by the same angle θ, then $\overrightarrow{A'} + \overrightarrow{B'}$ is the same as $\overrightarrow{A} + \overrightarrow{B}$ rotated clockwise by θ.

A.12.7   Example: Vector form of the Centroid

Suppose we have a triangle $\triangle ABC$, where $\overrightarrow{A}$, $\overrightarrow{B}$ and $\overrightarrow{C}$ are vectors from our origin to the vertices of the triangle. What are the coordinates of the centroid? Notice that in the proof that follows, we never need to make use of the actual coordinates of the vectors. In fact, although we think of our vectors here as being two-dimensional, this proof works in any number of dimensions.

The centroid is at the intersection of any two medians, so let us use the medians originating at $\overrightarrow{C}$ and $\overrightarrow{A}$.

The median from $\overrightarrow{C}$ connects it with the midpoint of the vector $(\overrightarrow{A} + \overrightarrow{B})/2$. We know that all vectors along that line are described by:

$$\alpha \overrightarrow{C} + (1 - \alpha)\left(\frac{\overrightarrow{A} + \overrightarrow{B}}{2} \right), \tag{A.54}$$

for some value of α. Similarly, the values of the vectors along the other median are given by:

$$\beta \overrightarrow{A} + (1 - \beta)\left(\frac{\overrightarrow{B} + \overrightarrow{C}}{2} \right), \tag{A.55}$$

for some value of β.

The intersection will occur when those two vectors are equal, or when:

$$\alpha \overrightarrow{C} + (1 - \alpha)\left(\frac{\overrightarrow{A} + \overrightarrow{B}}{2} \right) = \beta \overrightarrow{A} + (1 - \beta)\left(\frac{\overrightarrow{B} + \overrightarrow{C}}{2} \right).$$

If we multiply everything out and move all the terms to the left side, and gather like terms, we obtain:

$$\overrightarrow{A}\left(\frac{1 - \alpha - 2\beta}{2} \right) + \overrightarrow{B}\left(\frac{\beta - \alpha}{2} \right) + \overrightarrow{C}\left(\frac{1 - \beta - 2\alpha}{2} \right) = \overrightarrow{0}.$$

For this to be identically zero, we need all three coefficients to be zero, and if you solve the equations (at first is looks like there are three equations and two unknowns, but the equations are not independent), we find that the only solution is $\alpha = 1/3$ and $\beta = 1/3$. Putting these values back into either equation A.54 or A.55, we find that the vector describing the centroid is $(\overrightarrow{A} + \overrightarrow{B} + \overrightarrow{C})/3$.

$\boxed{\text{A.13}}$ ♦ **Complex Numbers**

This section provides only a tiny introduction to the huge topic of complex numbers. You can learn much more in virtually any book on complex analysis, but Needham's book (see [Needham, 1997]) is a particularly geometrically oriented book as you can tell from its title, "Visual Complex Analysis".

The reason complex numbers are so useful for geometers is that they provide yet another method to describe the locations of points on a plane.

A complex number has the form $a + b\mathrm{i}$, where i is the imaginary $\sqrt{-1}$ (in other words $\mathrm{i}^2 = -1$), and a and b are arbitrary real numbers. Complex numbers are a lot like two-dimensional vectors but instead of having an x- and a y-coordinate, you can think of them having a real coordinate and an imaginary coordinate.

Thus they are commonly plotted on a plane with a real and an imaginary axis, where the real axis is drawn in place of the x-axis and the imaginary axis in place of the y-axis. The origin represents the point $0 + 0\mathrm{i}$, and the point $a + b\mathrm{i}$ would be plotted a units to the right and b units above the origin (with the obvious reversals if a and/or b is negative). We usually use letters like z or w (or the Greek letter ζ "zeta") to indicate arbitrary complex variables that include both a real and an imaginary part.

All the usual operations exist for complex numbers—addition, subtraction, multiplication, division (except by zero). Here are formulas for the *basic* operations:

$$
\begin{aligned}
(a + b\mathrm{i}) \pm (c + d\mathrm{i}) &= (a \pm c) + (b \pm d)\mathrm{i} \\
(a + b\mathrm{i})(c + d\mathrm{i}) &= (ac - bd) + (ad + bc)\mathrm{i} \\
\frac{(a + b\mathrm{i})}{(c + d\mathrm{i})} &= \left(\frac{ac + bd}{c^2 + d^2}\right) + \left(\frac{bc - ad}{c^2 + d^2}\right)\mathrm{i}
\end{aligned}
$$

To derive the equation for division of complex numbers above, simply multiply numerator and denominator of the expression on the left by $c - d\mathrm{i}$.

In addition to these, many of our favorite functions also make sense, such as: $\sin z$, $\cos z$, e^z, $\sqrt{z}$, $\log z$, et cetera—although some care must be taken to define these properly, especially functions like $\sqrt{z}$ and $\log z$ which are "multiple-valued" functions. Except for the function e^z (see Section A.13.1) we will not need any of these here.

If we think of points on the Euclidean plane as single complex numbers, many important geometric operations have simple representations in terms of mathematical operations on the complex numbers.

For example, if we simply add the number $c + d\mathrm{i}$ to every point on the complex plane, this has the effect of a translation of c units to the right and d units up (where c and d are positive—negative values, of course, translate left and/or down). If we multiply our numbers by $\cos\theta + \mathrm{i}\sin\theta$ this has the effect of rotating them counterclockwise about the origin by an angle θ. There are many similar operations.

Similarly, the midpoint of z and z' is given by $(z + z')/2$. If z_1, z_2, and z_3 are the vertices of a triangle, the centroid is given by $(z_1 + z_2 + z_3)/3$. Multiplication by a pure positive real number scales the points toward or away from the origin uniformly.

For example, multiplication by 2 makes everything twice as big, and multiplication by $1/3$ makes everything one third as big. Multiplication by -1 reflects everything through the origin (or equivalently, rotates it by $180°$), and multiplication by i rotates counterclockwise by $90°$.

If $z = a + b$i is an arbitrary complex number, then we say that $\bar{z} = a - b$i is the complex conjugate of z. Geometrically, it amounts to the reflection of z across the real line (the line containing all the complex numbers whose imaginary parts are zero.) The complex conjugate has the following properties that are not hard to show:

$$\begin{aligned} \overline{(z \pm w)} &= \bar{z} \pm \bar{w} \\ \overline{zw} &= \bar{z}\bar{w} \\ \overline{z/w} &= \bar{z}/\bar{w}. \end{aligned}$$

The complex conjugate allows us to express the distance between a point and the origin $(0 + 0$i$)$, and this distance is written $|z|$, is defined by:

$$|z| = \sqrt{z\bar{z}},$$

and is called the **absolute value** of z.

The formula above works because for any complex z, $z\bar{z}$ is always real and non-negative. To see this, let $z = a + b$i where a and b are real numbers. Then

$$z\bar{z} = (a + b\text{i})(a - b\text{i}) = a^2 + b^2,$$

which is certainly always non-negative. In fact, if you think of a and b as coordinates on the complex plane, then we have

$$|z| = |a + b\text{i}| = \sqrt{a^2 + b^2},$$

which is exactly the same distance that the pythagorean theorem would give us for the distance from a point z to the origin.

The distance between two complex numbers z and w is given by $|z - w|$.

A.13.1 ♦♦ Euler's Formula

If you are not afraid to work with complex numbers, there are some formulas which can simplify many of the trigonometric calculations substantially.

If θ is any real number, then Euler's formula states that:

$$e^{\text{i}\theta} = \cos\theta + \text{i}\sin\theta.$$

We saw the expression $\cos\theta + \text{i}\sin\theta$ in the previous section—multiplication by this number is equivalent to a rotation by an angle θ.

If that is true, then we also have:

$$e^{-\text{i}\theta} = \cos(-\theta) + \text{i}\sin(-\theta) = \cos\theta - \text{i}\sin\theta.$$

The two equations above can be added or subtracted and manipulated in various other ways to yield:

$$\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2} \qquad \sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i} \qquad \tan\theta = \frac{e^{i\theta} - e^{-i\theta}}{i(e^{i\theta} + e^{-i\theta})}$$

We will not spend much time on it, but here is an example of how Euler's formula can be used to show one of the trigonometric equalities above. Let us derive equation A.6—the formula for $\cos(\alpha + \beta)$:

$$\begin{aligned}
\cos(\alpha + \beta) &= \frac{e^{i(\alpha+\beta)} + e^{-i(\alpha+\beta)}}{2} \\
&= \frac{e^{i\alpha}e^{i\beta} + e^{-i\alpha}e^{-i\beta}}{2} \\
&= \frac{(\cos\alpha + i\sin\alpha)(\cos\beta + i\sin\beta) + (\cos\alpha - i\sin\alpha)(\cos\beta - i\sin\beta)}{2} \\
&= \frac{(\cos\alpha\cos\beta - \sin\alpha\sin\beta) + i(\cos\alpha\sin\beta + \sin\alpha\cos\beta)}{2} \\
&\quad + \frac{(\cos\alpha\cos\beta - \sin\alpha\sin\beta) - i(\cos\alpha\sin\beta + \sin\alpha\cos\beta)}{2} \\
&= \cos\alpha\cos\beta - \sin\alpha\sin\beta.
\end{aligned}$$

But if you are a little more lazy, you can get the formulas for both $\sin(\alpha + \beta)$ and $\cos(\alpha + \beta)$ with less work:

$$\begin{aligned}
e^{i(\alpha+\beta)} = e^{i\alpha}e^{i\beta} &= \cos(\alpha + \beta) + i\sin(\alpha + \beta) \\
&= (\cos\alpha + i\sin\alpha)(\cos\beta + i\sin\beta) \\
&= (\cos\alpha\cos\beta - \sin\alpha\sin\beta) + i(\cos\alpha\sin\beta + \sin\alpha\cos\beta).
\end{aligned}$$

Setting the real part of the right side of the first line equal to the real part of the final line and similarly for the imaginary parts, we get both formulas:

$$\begin{aligned}
\cos(\alpha + \beta) &= \cos\alpha\cos\beta - \sin\alpha\sin\beta \\
\sin(\alpha + \beta) &= \cos\alpha\sin\beta + \sin\alpha\cos\beta.
\end{aligned}$$
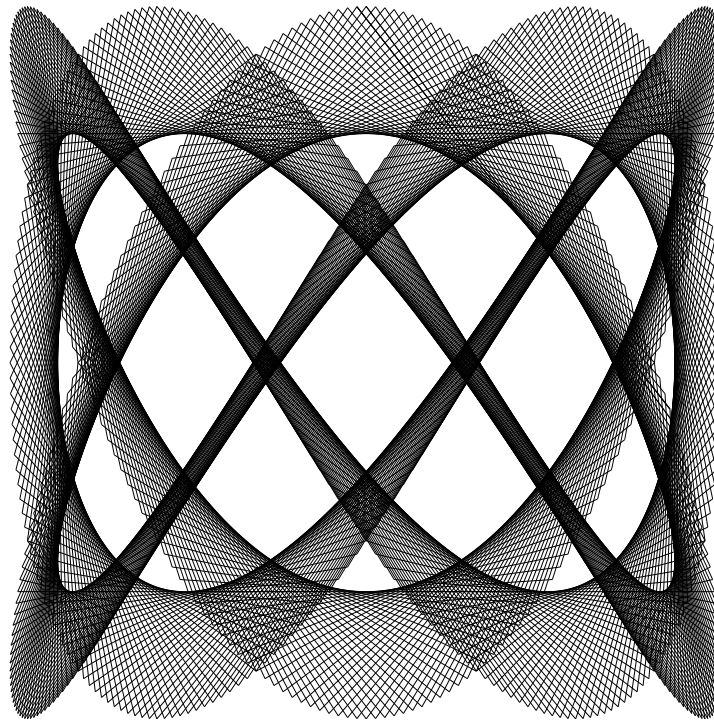
B

Geometer Art

At the top of the first page of each chapter of this book is a figure produced by **Geometer** that is primarily artistic. This appendix contains a short description of each. The **Geometer** files and any C-language files that were used to generate the figures can be found in the `Art` subdirectory of the **Geometer** installation.

Many of the files are parameterized so by editing the files and changing a number or two different versions of the artwork can be generated. Here are the descriptions:

- **Chapter 0: Preface**

  Geometer file: `Art/Koch5.T`
  C file: `Art/koch.c`

  The Koch snowflake curve is the limiting case of a sequence of shapes that begins with an equilateral triangle. To obtain the next shape in the sequence from the previous, erect an outward-pointing equilateral triangle on the middle third of

each segment of the boundary. Thus the second shape is like a star of David, et cetera. The illustration here is the fifth step in this sequence.

The curve is interesting because in its limiting form, it encloses a finite area, but has a boundary whose length is infinite.

- **Chapter 1: Introduction**

  Geometer file: `Art/Pretty1.T`

  This pattern is the result of a **Geometer** programming bug. A cosine function was applied instead of a sine function in an attempt to generate the Maurer rose that appears as the art at the beginning of Chapter 9.

- **Chapter 2: Computer-Assisted Geometry**

  Geometer file: `Art/dragon.T`
  C file: `Art/dragon.c`

  This is the "dragon curve". As you click on the *Next* button, successive approximations to the curve appear. The end result is a space-filling curve in the sense that in the limit, every point of the region is approached arbitrarily closely by the curve. The limiting length of the curve is, of course, infinite.

- **Chapter 3: Mathematics Review**

  Geometer file: `Art/Caustic.T`

  This illustration shows how caustic curves are generated in optics. Light rays from a point source are internally reflected within a sphere, and the resulting pattern of sparse and dense concentrations appears. Light through a wine glass generates similar patterns on a tablecloth. Press the *Run Script* button to generate the caustic.

- **Chapter 4: Computer-Aided Proof** This image is generated by beginning with an array of 9 circles arranged in three rows and columns, and inverting each of the circles in the other 8. Then those circles are again inverted in the original 9 circles and the process is repeated for a few steps. Different initial arrangements of circles and different depths can be achieved by modifying the "C" file and rerunning it.

- **Chapter 5: More Useful Theorems**

  Geometer file: `Art/fourier.T`

  Press *Run Script* to show a sequence of trogonometric Fourier series that approximate a square wave. Eight steps of the approximation are superimposed.

- **Chapter 6: Locus of Points**

  Geometer file: `Art/golden.T`

  A geometric illustration of one property of the golden section is illustrated here. Each time the *Next* command is pressed, a square is chopped off the end of a rectangle whose sides are in the golden ratio: $\tau = (1 + \sqrt{5})/2 \,:\, 1$. Each time a

square is chopped off, the resulting rectangle has exactly the same shape as the previous.

- **Chapter 7: Triangle Centers**

  Geometer file: `Art/hilbert.T`
  C file: `Art/hilbert.c`

  Each press of the *Next* button displays another in a sequence of curves whose limit is the Hilbert space-filling curve.

- **Chapter 8: Inversion in a Circle**

  Geometer file: `Art/Maurer.T`

  This figure is called the Maurer rose and is generated with the *Run Script* button. It is generated by connecting two points that move around an n-leaved rose with a fixed angular offset between them. The **Geometer** diagram can be modified to change both n and the angular displacement.

- **Chapter 9: Projective Geometry**

  Geometer file: `Art/peano.T`
  C file: `Art/peano.c`

  The Peano curve is another space-filling curve. The **Geometer** file is generated by a "C" computer program. Press *Next* to view successive approximations to the true Peano curve.

- **Chapter 10: Harmonic Point Sets**

  Geometer file: `Art/Poly/Poly.T`
  C file: `Art/Poly/poly.c`

  This is the projection of the $\{3, 3, 5\}$ 4-dimensional polytope onto a two dimensional surface. The **Geometer** diagram is generated by a "C" program that can be modified to change the direction of the projection.

- **Chapter 11: Geometric Presentations**

  Geometer file: `Art/square.T`
  C file: `Art/square.c`

  This is another space-filling curve whose approximations are quite interesting.

- **Chapter 12: Geometer Proofs**

  Geometer file: `Art/Whirl/Whirl.T`
  C file: `Art/Whirl/Whirl.c`

  This is a simple illustration created by drawing a regular polygon (in this case a regular heptagon) repeatedly drawn with a slight increment in the rotation angle and a slight decrement in size.

- **Appendix A: Geometric Construction**

  Geometer file: `Art/epicycloid.T`

  One definition of an epicycloid is as the envelope of the straight lines connecting points on the circumference of a circle with points n times as far along the surface. Press the *Run Script* button to view an example. If you edit the **Geometer** code, you can change the value of n to generate different epicycloids. The multiple n need not be an integer, but the results are a bit less interesting.

- **Appendix B: Geometer Art**

  Geometer file: `Art/Liss.T`

  In this pattern a pair of points follow each other with a fixed angular difference as they trace out a lissijous curve. Those points are connected with a line segment. Both the frequencies of the curve and the angular difference can be changed by editing the **Geometer** file.

- **Appendix C: Geometric Problem Solving Strategies**

  Geometer file: `Art/Spirograph.T`

  This is similar to the Spirograph® example in the text (see Section 2.15.3) except with different radii and numbers of teeth, and in addition, it is drawn with connected line segments instead of as a series of dots.

C

Geometric Problem Solving Strategies

It is impossible to give a cookbook method that will take an arbitrary proposed geometric theorem and either prove it or show that it is false. That this cannot be done was proved in 1931 by Kurt Gödel ([Gödel, 1931]). In spite of that, it is certainly possible to learn and organize geometric problem solving techniques. The more problems you work, the more you will see certain techniques used.

There are many excellent books on general problem-solving techniques that can be applied to any area of mathematics, but most do not concentrate on geometric problems. Many of the best were written by George Pólya, such as [Pólya, 1957], [Pólya, 1990a], [Pólya, 1990b], and [Pólya, 1962]. There are many others, but a good recent one is [Zeitz, 1999]. Section C.13 lists some of these more general mathematical approaches to problems.

This appendix is organized a bit like a thesaurus—look up a problem category to obtain a list of possible approaches to problems in that category. Obviously no such listing (either of categories or strategies) could ever be complete, so it may not contain exactly what you need.

To use this appendix, first have a clear idea of what you are trying to do, like: "Show that points A, B and C all lie on a line." or "Find the length of segment AB." or "Construct a circle tangent to two circles and a line." Then look through the list of general categories below to see which ones might apply. Go to the referenced sections and continue from there.

- **Congruence:** Methods to prove that geometric figures are congruent. Section C.1.

- **Similarity:** Methods to prove that geometric figures are similar. Section C.2.

- **Special Figures:** Showing that figures are special in some way. Is the figure a square? A rectangle? An equilateral triangle? An isosceles triangle? Section C.3.

- **Concurrence:** Methods to prove facts about concurrence—are two points in fact the same point? Do three points lie on a line? Do four points lie on a circle? Do three lines go through the same point? Are two circles tangent? Section C.4.

- **Measures:** Methods to measure things. How long is a line segment? What is the measure of an angle? What is the area of a triangle? Section C.5.

- **Equality and Inequality:** Methods to show line segments, angles, areas, et cetera, are of equal or different lengths. If they are different, which one is larger? How to show triangles are equilateral or isosceles, how to show that a quadrilateral is a parallelogram, rhombus, or trapezoid. Section C.6.

- **Ratios:** Methods to work with ratios: how to find them, how to calculate with them, what to use them for. Section C.7.

- **Inversion:** Methods to use inversion to simplify a proof or construction.

- **When to Draw the Line:** When and how to construct auxiliary lines or circles in a geometric figure to complete a proof.

- **Construction Techniques:** Making straightedge and compass constructions.

- **Relabeling** How to work out the details once, and then be able to say, "et cetera".

Each section contains short descriptions of techniques, possibly followed by references to sections where that particular technique was used or described. Techniques that are advanced or somehow difficult to use are preceded by ♦.

$\boxed{\text{C.1}}$ Congruence

Technically[1], congruence is an undefined term. In spite of that, one can think of congruent figures as having exactly the same size and shape, but they may be flipped upside-down or moved around or rotated.

1. **Two segments are congruent (have the same length).**

   a. *They are sums or differences of congruent segments.*

   b. *They are corresponding parts (where "part" means any linear part—edges, altitudes, perimeters, et cetera) of congruent triangles.*

   c. *They are opposite sides of a parallelogram.*

   d. *They are the two parts of a line bisected by another line or point. (For example, the diagonals of a parallelogram bisect each other.)*

   e. *They are opposite equal angles in a triangle.*

   f. *They are radii or diameters of the same circle.*

   g. *They are the two external tangents to a circle from a point outside, or they are a pair of common external tangents (or a pair of internal tangents) of two circles.*

   h. *Any point on an angle bisector is equidistant from the two sides of the angle.*

   i. *Segments that subtend equal arcs are congruent.*

2. **Two angles are congruent.**

   a. *They are vertical angles.*

   b. *They are cut off by a line that is a transversal of two parallel lines.*

   c. *They are corresponding angles in congruent triangles.*

   d. *They are sums or differences of equal angles.*

   e. *They are corresponding angles in similar triangles.*

   f. *They are opposite equal line segments in a triangle.*

   g. *They are both interior angles and are inscribed in the same arc (or equal arcs) in a circle, or they are both central angles and subtend equal arcs in a circle. (This is especially common for (right) angles inscribed in a semicircle.)*

   h. *They are complementary (or supplementary) equal angles.*

   i. *They are opposite angles in a parallelogram.*

3. **Two triangles are congruent.**

   a. *SAS (side-angle-side)—show that two sides and the included angle in both triangles are congruent.*

   b. *SSS (side-side-side)—show that all three sides of the two triangles are congruent in pairs.*

   c. *ASA (angle-side-angle) or AAS (angle-angle-side)—show that one side is congruent and that two of the corresponding angles are also congruent. Be*

---

[1] In a pure axiomatic description of geometry, there are certain undefined terms, and all we can technically know about them is what we can derive from the axioms and logic. The relation "congruent to" is one of these undefined terms.

sure that the angles are in the corresponding positions—in other words, if the side is between the two angles in one, is also between them in the other, and if it is not between them, it must form part of the same angle in both triangles. (See

d. *Hypotenuse-leg—show that they are both right triangles and that the two hypotenuses and two corresponding legs are congruent. Note: This is a special case of the SSA theorem that states that we can conclude that the triangles are congruent only if the angle is $90°$ or greater.*

4. **Two arcs are equal.**

   a. *They have equal central angles.*
   b. *They have equal inscribed angles.*
   c. *They are sums or differences of equal arcs.*
   d. *They are cut from a circle between a pair of parallel lines.*

C.2   Similarity

Two figures are similar if they are the same shape, although they may be different sizes.

1. **Two triangles are similar.**

   a. *AA (angle-angle)—show that two angles in one are the same as two angles in the other. Notice that since any two equal angles imply the similarity of a pair of triangles, there is no need for rules like ASA or AAS for similarity.*
   b. *A line parallel to the base of a triangle through the other two sides cuts off a triangle similar to the original. (This is basically the same as AA.)*
   c. *All three pairs of corresponding edges in the two triangles are parallel.*
   d. *SAS (side-angle-side)—show that two sides in a triangle are in the same ratio with the corresponding sides in another, and that the angles included between those sides are equal.*
   e. *SSS (side-side-side)—show that all three sides in one triangle are in the same ratio with the corresponding sides in another.*
   f. *If two chords AB and CD of a circle intersect at a point P then $\triangle APC \sim \triangle DPB$.*

C.3   Special Figures

1. **A triangle is equilateral.**

   a. *Show all three sides are equal.*
   b. *Show all three angles are equal.*
   c. *Show two angles are equal to $60°$.*

2. **A triangle is isosceles.**

   a. *Show that two sides are equal.*
   b. *Show that two angles are equal.*

3. A quadrilateral is a rectangle.

   a. *Show three angles are* 90°.
   b. *Show that it is a parallelogram with one* 90° *angle.*
   c. *Show that it is a parallelogram and that the diagonals are of equal length.*

4. A quadrilateral is a square.

   a. *Show that it is a rectangle with adjacent equal sides.*
   b. *Show that all sides are equal and one angle is* 90°.

5. A quadrilateral is a trapezoid.

   a. *Show that two sides are parallel.*

6. A quadrilateral is a rhombus.

   a. *Show that it is a parallelogram and that the diagonals are perpendicular.*
   b. *Show that all four sides are equal.*

7. A quadrilateral is a parallelogram.

   a. *Show that both pairs of opposite sides are parallel.*
   b. *Show that both pairs of opposite sides are equal.*
   c. *Show that its vertices are the midpoints of any quadrilateral.*
   d. *Show one pair of sides is parallel and the same length.*
   e. *Show that the opposite angles are congruent.*
   f. *Show that adjacent angles are supplementary.*
   g. *Show that the diagonals bisect each other.*

$\boxed{\text{C.4}}$ Concurrence

1. Two points are concurrent.

   a. *Show that they both divide a line segment in the same ratio.*
   b. *Show that each lies on two non-parallel lines.*

2. Three points lie on the same line.

   a. *Points A, B, and C lie on the same straight line (with B between A and C) if AB and BC make the same angle with some third line. (In a sense they share a point and would normally be parallel, so they are the same line.)*
   b. ♦ *Use Menelaus' theorem (see Section 5.2).*

3. Three lines pass through the same point.

   a. *Show that they pass through a known point. For example, if the three lines can be shown to be altitudes, or angle bisectors, or medians, et cetera, of some triangle, they are concurrent.*
   b. *Show that two of the lines pass through the same point of the third.*
   c. ♦ *Use Ceva's theorem (see Section 5.1).*
   d. ♦ *Use Brianchon's theorem.*

4. Three circles intersect at a point.

    a. *Show that the point of intersection of two of them lies on the third, perhaps by showing it is on an inscribed triangle or quadrilateral.*

    b. *Find a point whose distance to each circle is the radius of that circle.*

5. A point lies on a circle.

    a. *The point is the same distance from the center as another point known to be on the circle.*

    b. *The point makes an equal angle with an arc of the circle as another point known to be on the circle.*

    c. *The point lies on a vertex of a right triangle whose hypotenuse is the diameter of the circle.*

6. A line and a circle are tangent.

    a. *Show the line is perpendicular to a radius and passes through the endpoint of that radius.*

7. Two circles are tangent.

    a. *Show that the distance between their centers is the sum of their radii.*

8. Four points lie on the same circle. (Equivalently, show that the quadrilateral whose vertices are those points can be inscribed in a circle.)

    a. *An isosceles trapezoid can be inscribed in a circle.*

    b. *If the points can be divided into two right triangles that share a hypotenuse, they lie on the same circle.*

    c. *More generally, if the four points are A, B, C, and D, show that $\angle ACB = \angle ADB$.*

    d. *If either pair of opposite angles in a quadrilateral are supplementary, the four points lie on a circle. In other words, if $ABCD$ is a convex quadrilateral and $\angle ABC + \angle CDA = 180°$, then the four points lie on a circle.*

    e. ♦ *(Ptolemy's theorem) If $ABCD$ is a convex quadrilateral and $AB \cdot CD + BC \cdot DA = AC \cdot BD$ then the four points lie on a circle. See Section 5.5.*

9. A quadrilateral circumscribes a circle.

    a. *If the four sides, in order, have lengths a, b, c, and d, the quadrilateral circumscribes a circle if and only if $a + c = b + d$.*

    b. *Show that each side of the quadrilateral is tangent to the circle.*

$\boxed{\text{C.5}}$ Measures

Many problems require that you find a measure of a length, angle, area, or something else in terms of some given measurements. This is a huge category of problems, but there are some useful general approaches:

- Draw a picture—either by hand or with **Geometer**. Label the parts you know and look at the parts you do not.

- What relationships do you see in the figure. Look for things like right angles, parallel lines, centers of circles, equal included angles, angles that sum to a triangle or to a straight line, et cetera.

1. **To find the length of a segment.**

    a. *Use the pythagorean theorem (stated and proved in Section A.6.5) on a right triangle.*
    b. *Use the fact that the segment is the sum or difference of other segments that you can measure.*
    c. *Use the law of cosines (see Section A.10.2.6).*
    d. *Use the law of sines (see Section A.10.2.4).*

2. **An angle is a right angle. (Equivalently, show that the two lines making up the angle are perpendicular.)**

    a. *Show that the angle is inscribed in a semicircle.*
    b. *Use the pythagorean theorem: If the angle is opposite a side of length c in a triangle whose other sides are a and b, show $a^2 + b^2 = c^2$.*
    c. *A radius (or diameter) that bisects a chord of a circle is perpendicular to it. Similarly, a tangent to a circle is perpendicular to the radius at the point of tangency.*
    d. *The diagonals of a rhombus or kite are perpendicular.*
    e. *The line connecting the centers of two intersecting circles is perpendicular to the line connecting the points of intersection.*
    f. *The two angle bisectors of two intersecting lines are perpendicular.*
    g. *Two lines that make the same angle with perpendicular lines are themselves perpendicular.*

3. **Lines are perpendicular. (This is equivalent to showing they form a right angle. See above.)**
4. **Lines are parallel.**

    a. *Show that they have no points in common.*
    b. *Show that a transversal line makes equal angles with both the given lines. (The most common special case of this is to show that they are both perpendicular to the same line.)*
    c. *Show that the lines are opposite sides in a parallelogram.*
    d. *Show that both are perpendicular to the same line.*
    e. *Show both are parallel to the same line.*
    f. *A line connecting the midpoints of two sides of a triangle is parallel to the third side.*
    g. *Two lines that make the same angle with parallel lines are themselves parallel.*

5. **To find the measure of an angle.**

    a. *The three angles in a triangle add to $180°$.*
    b. *The n internal angles in an n-sided polygon add to $(n-2) \cdot 180°$ degrees.*
    c. *Complementary angles add to $90°$.*

   d. *Supplementary angles add to $180°$.*

   e. *Angles that go completely around a point add to $360°$.*

   f. *Angles that meet to form a straight angle add to $180°$.*

   g. *The central angle of an arc is twice an angle inscribed angle inscribed on that arc.*

   h. *A tangent to a circle with a secant at its point of tangency form an angle of half the measure of the arc it cuts.*

   i. *The exterior angle of a triangle is equal to the sum of the two interior angles at the other vertices.*

   j. *Opposite angles in a cyclic quadrilateral add to $180°$.*

6. **To find the area of a polygon.**

   a. *The area of a triangle is half the base times the height. Remember that there are three different base-height combinations and any of them works.*

   b. *Use Heron's Formula to find the area of a triangle given the lengths of the sides (see Section 5.6).*

   c. *The area of a parallelogram (or rectangle, in particular) is the base times the height.*

   d. *If a complex figure can be divided into smaller figures, its area is the sum of the areas of the smaller figures.*

   e. *The area of a triangle that has sides of lengths a and b and included angle θ is $(ab \sin\theta)/2$.*

   f. *Use Brahmagupta's Formula (see Section 5.6.3) to find the area of a cyclic quadrilateral.*

C.6 Equality and Inequality

1. **One segment is larger than another.**

   a. *Show that the segments are opposite unequal angles in a triangle. The larger segment is opposite the larger angle.*

   b. *The hypotenuse of a right triangle is longer than either of the legs.*

   c. *The sum of any two legs of a triangle is longer than the third. The sum is equal only for degenerate triangles. (This is called the triangle inequality.)*

2. **A line is an angle bisector.**

   a. *Show that both the angles formed are equal.*

   b. ♦ *In $\triangle ABC$, if the line through A hits the segment BC at a point D, and if $AD : BD = AC : CD$ then the line AD bisects $\angle BAC$.*

C.7 Ratios

1. **Equality of products or ratios of lengths.**

    a. *Usually when products of lengths are to be proved equal, or ratios are to be proved equal, it is done by finding one or a series of similar triangles whose ratios are then multiplied or divided together to get the desired product or ratio.*

    b. *Three parallel lines cut equal ratios of any transversal line.*

    c. *If two chords AB and CD of a circle intersect at a point P then $PA \cdot PB = PC \cdot PD$.*

    d. *In a triangle, the angle bisector divides the opposite side in the same ratio as the lengths of the adjacent sides.*

2. **Finding Ratios – where to get ratios of lengths.**

    a. *Similar triangles (or any similar polygons) have equal ratios of all the corresponding linear parts. The "parts" may be sides, medians, altitudes, perimeters, et cetera.*

    b. *Fixed ratios. For example, the diameter of a circle is always twice the radius, or the altitude of an equilateral triangle is always $\sqrt{3}/2$ times the length of the side.*

    c. *Chords AB and CD of a circle that cross at P yield $AP \cdot PB = CP \cdot PD$. Note that P need not lie inside the circle for this to be true.*

    d. ♦ *Menelaus' theorem yields products of ratios that are equal to 1 (see Section 5.2).*

    e. ♦ *Ceva's theorem yields products of ratios equal to 1 (see Section 5.1).*

    f. *Points that are inverted in the same circle yield ratios of lengths (ee Section 8.8).*

C.8 ♦ **Inversion in a Circle**

Inversion is so important that an entire chapter (Chapter 8)is devoted to it. The examples in this section will not make much sense unless you are familiar with the contents of that chapter.

This powerful technique converts circles and lines to other circles and lines, but which are converted to which depends on the location of the center of inversion, and on the radius of the circle of inversion. The technique is almost always the same— choose a center of inversion and a radius (sometimes the radius is unimportant), invert the lines, circles, and points in the problem, solve the problem on the inverted figures, and then re-invert that solution to get a solution of the original problem.

For reference, here is a short list of the more important properties of inversion through a circle K centered at O:

1. **Points are inverted to points.**
2. **Circles and lines are inverted to circles and lines, but a circle may become a line and a line may become a circle.**
3. **Only lines or circles passing through O are inverted to lines.**
4. **Angles between lines and/or circles are preserved by inversion.**
5. **Inversion sends O to the point at infinity and the point at infinity to point O.**

6. Points on K are inverted to themselves.

C.9   Algebraic Manipulation

If you are good at algebra, it is often not too hard to assign a coordinate system of some kind to your figure and then to state the various relationships algebraically and using those, to prove the desired result using solely algebraic manipulations.

There are many coordinate systems, and almost all are amenable to algebraic manipulation. But there are coordinate-free representations (vectors) that can also be manipulated algebraically.

Often, one can prove properties with a combination of geometric and algebraic techniques—purely geometric theorems are used to obtain various ratios and then those ratios are manipulated algebraically to obtain the desired results.

C.9.1   Calculating with Angles

One of the most common uses of algebra in geometric proofs is for manipulation of formulas involving angles. It is easy to write down many angle equalities in most figures—the angles in a triangle add to $180°$ or in a quadrilateral to $360°$, opposite angles in a cyclic quadrilateral add to $180°$, et cetera.

C.9.2   Calculating with Ratios

It is also easy to find ratios and quite often a bit of algebra is needed to combine the various formulas in a manner that gives the desired result.

C.9.3   Working with Areas

Often the area of a figure can be calculated in different ways, and those values can be set equal. For example, figures can be divided into triangles in many ways, and every sum of those triangle areas must be the same. Here are some places where areas are used in algebraic manipulations:

C.9.4   Primarily Algebraic Theorems

The following theorems are geometric, but their results are typically applied algebraically:

1. **The pythagorean theorem relates the lengths of the three sides in a right triangle. See Section A.6.5.**
2. **Stewart's theorem relates the length of a cevian (a line connecting a vertex to the opposite side) of a triangle to the other lengths. See Section A.6.6**

3. **Heron's formula gives the area of any triangle in terms of the lengths of its sides. See Section 5.6.**
4. **Brahmagupta's formula allows you to calculate the area of a cyclic quadrilateral given the lengths of its sides. See Section 5.6.3.**
5. **Ptolemy's theorem relates the lengths of the sides and diagonals of a cyclic quadrilateral. See Section 5.5.**
6. **The law of sines relates the sines of the angles of a triangle with the lengths of the sides opposite them and to the diameter of the circumcircle. See Section A.10.2.4.**
7. **The law of cosines is a generalization of the pythagorean theorem to arbitrary triangles. See Section A.10.2.6.**
8. **Various trigonometric formulas. See Section A.10**

C.9.5 Using Trigonometry

See Section A.11 for a discussion of the technique and an example where we show that the three medians of a triangle meet at a point.

C.9.6 Vector Techniques

Section A.12 is a nice introduction to this technique, together with a demonstration that the centroid of a triangle is the average of the three vectors to its vertices. Here are a couple of situations where vector techniques might be used:

1. **If there is rotational symmetry in the problem or solution. For example, if you wish to prove that a triangle is equilateral, show that any edge can be rotated $60°$ to become another edge.**
2. **If an entire figure is rotated or translated and relationships are sought between the figure and the rotated or translated version.**

C.10 When to Draw the Line

One of the hardest techniques for beginners to learn is when it might be useful to draw an auxiliary line or circle to complete a proof. Here are some rules of thumb:

1. **If you want to apply some theorem about triangles, but your figure contains no triangles, it may be useful to draw a line that makes a triangle. Obviously, the word "triangle" in the previous sentence can be replaced by any other geometric figure. A very common case is the construction of a line that produces a pair of similar triangles.**
2. **If you need to calculate lengths or relative lengths, it is often useful to use theorems like the pythagorean theorem. To do so requires a right triangle, so to create one, you often need to drop a perpendicular or do something else to make one or more right triangles in your figure.**

3. **If you need to know something about a sum of lengths or angles, try to find a construction that puts the lines end-to-end, or the angles next to each other so that addition simply amounts to looking at a combined length or angle.**

4. **If there is some apparent concurrency that is not obvious from the statement of the problem, it is often useful to make a construction that somehow helps you prove that the concurrency is in fact true. For example, if in your drawings (with Geometer or drawn by hand) it seems like three points lie on a line, or that four points lie on a circle, try drawing that line through two of the points or the circle through three of the points and see if you can prove that the other point lies on the line or circle.**

   You can sometimes discover concurrencies using Geometer by drawing a suspicious point in a smearing color and then manipulating the rest of the figure to see if that point moves in a regular way—along a line or around a circle, for example.

5. **Remember that there is a whole class of problems that can be solved by inverting your figures through a circle. A discussion of when to use inversion can be found in this chapter in Section C.8.**

6. **The list above does not cover all cases. Here is a list of other "miscellaneous" constructions that have been used in this book, together with a short description of how and why each was used.**

   a. *In Section A.10.2.3, a perpendicular line is dropped to convert an arbitrary triangle into one where we can use the definition of* $\sin\theta$ *on a right triangle.*

   b. *In Section A.10.2.4, we construct a circle and a right triangle inscribed in it to be able to use our definition of* $\sin\theta$ *on a right triangle.*

   c. *In Section 12.1.3, a series of reflections of a pair of triangles is constructed. The legs of one series of reflections lie in a line which is clearly the shortest distance between two points.*

   d. *In Section 8.17 in addition to all the inversion-related constructions, points on a line are projected to use the fact that harmonic sets of points are preserved under projection. Also, a triangle connecting the midpoints of the original is constructed since those points are known to lie on the nine-point circle, and their positions are well-known relative to the original triangle.*

   e. *In Section 12.1.12, a triangle is constructed so that an earlier theorem could be applied to a small part of the figure. In addition, a few other lines required by the previous theorem were constructed.*

C.11 | Construction Techniques

To do general geometric constructions using a straightedge and compass, you should have a solid idea of how to do the basic constructions presented in Section 3.5.

Most construction problems are solved by figuring out some properties that the solution has to satisfy, and then to apply some combination of the primitive constructions referred to above.

It is a bit difficult to divide construction problems into nice categories, but here at least are a few general approaches that might be useful:

1. **Draw a picture of a completed construction and work with it to see what relations hold. It is often quite easy to draw a solution and work back to the problem. For example, if you are required to draw a circle tangent to three lines, draw a circle and then draw three lines tangent to it. From that you will see that the three lines form a triangle, and then you will have something to analyze.**
2. **Sometimes you can use brute force. If you need to construct a $36°$, recall that $\cos(36°) = (1 + \sqrt{5})/4$. Draw an arbitrary segment and call its length 1. Based on that, construct another segment of length $(1 + \sqrt{5})/4$. The right triangle having legs lengths 1 and $(1 + \sqrt{5})/4$ will contain a $36°$ angle.**
3. **♦ If the problem requires the construction of a line or circle tangent to combinations of lines and circles, consider inverting the problem through the proper circle. See Section C.8.**

$\boxed{\text{C.12}}$ Relabeling

Quite often if you wish to prove something for a figure that is completely non-specific, you can work out the properties for one side or angle, and then just change the labels in a consistent way to get the result for the others.

As an example, the law of sines states that

$$\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C} = 2R$$

for any triangle $\triangle ABC$, where R is the radius of the circumscribed circle (See Section A.10.2.4). If you can show that $a/(\sin A) = 2R$ you are done, since the triangle is completely arbitrary, so the labeling does not matter at all. In other words, if you had taken the same triangle, but had originally exchanged the vertices where you wrote down "A" and "B", there would be no change in your proof, but your proof would have shown that the ratio $b/\sin B$ is also equal to $2R$.

An additional example can be found in Section 4.10 were the existence of the Simson line is shown. The theorem concerns an arbitrary triangle $\triangle ABC$ and a point M on the circumcircle. The proof in Section 4.10 depends on the fact that M lies on the arc $\overparen{CB}$, but had it been on a different arc, the same argument would hold after we properly relabeled the vertices of the triangle.

$\boxed{\text{C.13}}$ General Problem Solving Approaches

Here is a list of strategies that can be applied to any mathematical problem, including geometric ones. The list here is by no means complete, but it may be helpful to review it if you are stuck on a particular problem.

- **Do you understand the problem?** Make certain you know what you're looking for. Name the unknowns. List the known facts.

- **Did you draw a picture?** This is almost always absolutely critical in geometry. In fact, more than one picture is even better. Try to draw pictures where the result fails. Often your failed attempts to draw a counter-example make it clear why the result must be true.

- **Have you seen or solved a similar problem?** If so, what techniques worked there? Will they work again? If not, why not?

- **Can you invent a simpler version of the problem that you can solve?** If the problem is for a general polygon, can you do it for a triangle? For a quadrilateral? If it is about a general triangle, can you solve it for an equilateral triangle? For an isosceles or right triangle?

- **Can you decompose the problem?** Can you break the problem into parts so that if you solve all the parts, you have solved the whole problem?

- **Can you use contradiction?** In other words, if you assume that the result you are trying to prove is false, does that lead to a nonsensical result? If so, you are done.

- **Did you use all the information?** Sometimes there is extra data in a mathematics problem, but it is rare. If you solve a problem without using some of the data, it is likely you have made a mistake.

- **Can you guess, or at least estimate, what the correct answer might be?** This gives a check of your final answer, and may help you to solve the problem.

- **Can you work backwards?** If you assume your final result and work toward the problem statement, often those steps can be reversed to obtain a solution.

- **Have you discussed the problem with others?** If you can work with others, do so. Two heads are better than one. Surprisingly, even trying to explain what is going on to someone totally ignorant of mathematics often helps.

- **Can you think of different approaches?** Maybe vectors will work. Maybe a symmetry argument. Maybe you can assign coordinates and just grind out the solution algebraically. Maybe you can use trigonometry.

- **When you get a solution, can you find another approach?** If so, you may have a way to check your first solution.

- **Get some sleep.** This can be surprisingly effective. You struggle late into the night on a problem and make no progress, but with a clear head in the morning, the solution is obvious.

Bibliography

[Abbott, 1952] Abbott, E. A. (1952). *Flatland: A Romance of Many Dimensions*. Dover Publications, Inc., New York, seventh edition.

[Altshiller-Court, 1952] Altshiller-Court, N. (1952). *College Geometry*. Barnes & Noble, Inc., New York.

[Beman and Smith, 1900] Beman, W. W. and Smith, D. E. (1900). *New Plane and Solid Geomtery*. Ginn & Company, Boston.

[Berger, 1987a] Berger, M. (1987a). *Geometry I*. Springer-Verlag, Berlin, Heidelberg, New York.

[Berger, 1987b] Berger, M. (1987b). *Geometry II*. Springer-Verlag, Berlin, Heidelberg, New York.

[Berger et al., 1984] Berger, M., Pansu, P., Berry, J.-P., and Saint-Raymond, X. (1984). *Problems in Geometry*. Springer-Verlag, New York, Berlin, Heidelberg, Tokyo.

[Bland, 1819] Bland, M. (1819). *Geometrical Problems Deducible from the First Six Books of Euclid*. J. Smith, Cambridge, England.

[Burger, 1965] Burger, D. (1965). *Sphereland: A Fantasy About Curved Spaces and an Expanding Universe*. Apollo Editions, New York.

[Conway and Guy, 1996] Conway, J. and Guy, R. (1996). *The Book of Numbers*. Copernicus—An Imprint of Springer-Verlag, New York.

[Coxeter, 1989] Coxeter, H. S. M. (1989). *Introduction to Geometry*. John Wiley & Sons, Inc., New York, second edition.

[Coxeter and Greitzer, 1967] Coxeter, H. S. M. and Greitzer, S. L. (1967). *Geometry Revisited*. The Mathematical Association of America, New York.

[Dudley, 1994] Dudley, U. (1994). *The Trisectors*. The Mathematical Association of America, Washington D.C.

[Euclid, 1956] Euclid (1956). *The Thirteen Books of The Elements*. Dover Publications, Inc., New York.

[Eves, 1965] Eves, H. (1965). *A Survey of Geometry*. Allyn and Bacon, London.

[Gödel, 1931] Gödel, K. (1931). Über formal unentscheidbare sätze der principia mathematica und verwandter systeme, i. *Monatsh. Math. u. Phys.*, 38:173–198.

[Greenberg, 1993] Greenberg, M. J. (1993). *Euclidean and Non-Euclidean Geometries*. W. H. Freeman and Company, New York, third edition.

[Hartshorne, 1997] Hartshorne, R. (1997). *Companion to Euclid*. Number 9 in Berkeley Mathematics Lecture Notes. American Mathematical Society, Providence.

[Hilbert and Cohn-Vossen, 1983] Hilbert, D. and Cohn-Vossen, S. (1983). *Geometry and the Imagination*. Chelsea Publishing Company, New York.

[Johnson, 1929] Johnson, R. A. (1929). *Advanced Euclidean Geometry*. The Riverside Press, Cambridge, Mass.

[Johnson, 1917] Johnson, R. A. (Mar. 1917). Directed angles in elementary geometry. *The American Mathematical Monthly*, 24:3:101–105.

[Kazarinoff, 1961] Kazarinoff, N. D. (1961). *Geometric Inequalities*. The Mathematical Association of America, Washington D.C.

[Kimberling, 1994] Kimberling, C. (1994). Central points and central lines in the plane of a triangle. *Math. Mag.*, 67:163–187.

[King and Schattschneider, 1997] King, J. and Schattschneider, D., editors (1997). *Geometry Turned On! Dynamic Software in Learning, Teaching, and Research*. MAA Notes. The Mathematical Association of America, Washington, D.C.

[Klarner, 1981] Klarner, D. A. (1981). *Mathematical Recreations, A Collection in Honor of Martin Gardner*. Dover Publications, Inc., New York.

[Lawrence, 1972] Lawrence, J. D. (1972). *A Catalog of Special Plane Curves*. Dover Publications, Inc., New York.

[Martin, 1991] Martin, G. E. (1991). *Geometric Constructions*. Springer, New York.

[Meyers, 1996] Meyers, L. F. (1996). Update on william wernick's "triangle constructions with three located points". *Mathematics Magazine*, 69:46–49.

[Minsky, 1967] Minsky, M. L. (1967). *Computation: Finite and Infinite Machines*. Prentice-Hall, Inc., Englewood Cliffs, N.J.

[Moise, 1963] Moise, E. (1963). *Elementary Geometry From An Advanced Standpoint*. Addison-Wesley, Reading, Mass.

[Needham, 1997] Needham, T. (1997). *Visual Complex Analysis*. Oxford University Press, Oxford.

[Nelson, 1993] Nelson, R. B. (1993). *Proofs Without Words: Exercises in Visual Thinking*. The Mathematical Association of America, Washington, D.C.

[Nelson, 2000] Nelson, R. B. (2000). *Proofs Without Words II: More Exercises in Visual Thinking*. The Mathematical Association of America, Washington, D.C.

[Ogilvy, 1969] Ogilvy, C. S. (1969). *Excursions in Geometry*. Dover Publications, Inc., New York.

[Pedoe, 1988] Pedoe, D. (1988). *Geometry, A Comprehensive Course*. Dover Publications, Inc., New York.

[Pedoe, 1995] Pedoe, D. (1995). *Circles: A Mathematical View*. The Mathematical Association of America, Washington, D.C.

[Pólya, 1957] Pólya, G. (1957). *How to Solve It*. Doubleday, second edition.

[Pólya, 1962] Pólya, G. (1962). *Mathematical Discovery: On Understanding, Learning and Teaching Problem Solving*. John Wiley & Sons.

[Pólya, 1990a] Pólya, G. (1990a). *Mathematics and Plausible Reasoning*. Princeton University Press, reprint edition.

[Pólya, 1990b] Pólya, G. (1990b). *Mathematics and Plausible Reasoning: Volume II Patterns of Plausible Inference*. Princeton University Press, reprint edition.

[Posamentier and Salkind, 1988] Posamentier, A. S. and Salkind, C. T. (1988). *Challenging Problems in Geometry*. Dover Publications, Inc., New York.

[Ryser, 1963] Ryser, J. H. (1963). *Combinatorial Mathematics*. The Mathematical Association of America, Rahway, N.J.

[Sharygin, 1988] Sharygin, I. F. (1988). *Problems in Plane Geometry*. Mir Publishers, Moscow.

[Weeks, 1985] Weeks, J. R. (1985). *The Shape of Space: How to Visualize Surfaces and Three-Dimensional Manifolds*. Marcel Dekker, Inc., New York.

[Weisstein, 1999] Weisstein, E. W. (1999). *CRC Concise Encyclopedia of Mathematics*. CRC Press, Boca Raton.

[Wells, 1991] Wells, D. (1991). *The Penguin Dictionary of Curious and Interesting Geometry*. Penguin Books, London.

[Wernick, 1982] Wernick, W. (1982). Triangle constructions with three located points. *Mathematics Magazine*, 55:227–230.

[Yaglom, 1962a] Yaglom, I. M. (1962a). *Geometric Transformations I*. The Mathematical Association of America, New York.

[Yaglom, 1962b] Yaglom, I. M. (1962b). *Geometric Transformations II*. The Mathematical Association of America, New York.

[Yaglom, 1962c] Yaglom, I. M. (1962c). *Geometric Transformations III*. The Mathe-
    matical Association of America, New York.

[Zeitz, 1999] Zeitz, P. (1999). *The Art and Craft of Problem Solving*. John Wiley &
    Sons, Inc., New York.

Index