Geometry Unbound

Kiran S. Kedlaya
version of 18 Jan 2006

Contents

Introduction

Origins, goals, and outcome

The original text underlying this book was a set of notes[1] I compiled, originally as a participant and later as an instructor, for the Math Olympiad Program (MOP),[2] the annual summer program to prepare U.S. high school students for the International Mathematical Olympiad (IMO). Given the overt mission of the MOP, the notes as originally compiled were intended to bridge the gap between the knowledge of Euclidean geometry of American IMO prospects and that of their counterparts from other countries. To that end, they included a large number of challenging problems culled from Olympiad-level competitions from around the world.

However, the resulting book you are now reading shares with the MOP a second mission, which is more covert and even a bit subversive. In revising it, I have attempted to usher the reader from the comfortable world of Euclidean geometry to the gates of "geometry" as the term is defined (in multiple ways) by modern mathematicians, using the solving of routine and nonroutine problems as the vehicle for discovery. In particular, I have aimed to deliver something more than "just another problems book".

In the end, I became unconvinced that I would succeed in this mission through my own efforts alone; as a result, the manuscript remains in some ways unfinished. For one, it still does not include figures (though some of these do exist online; see the chapter "About the license"); for another, I would ideally like to include some additional material in Part III (examples: combinatorial geometry, constructibility).

[1]The original notes have been circulating on the Internet since 1999, under the pedestrian title "Notes on Euclidean Geometry".

[2]The program has actually been called the Math Olympiad Summer Program (MOSP) since 1996, but in accordance to common custom, we refer to the original acronym.

Rather than continue endlessly to "finish" the manuscript, I have instead decided to carry the spirit of the distribution of the notes to a new level, by deliberately releasing an incomplete manuscript as an "open source" document using the GNU Free Documentation License; (for more on which see the chapter "About the license"). My hope is that this will encourage readers to make use of this still unpolished material in ways I have not foreseen.

Methodology

This book is not written in the manner of a typical textbook. (Indeed, it is not really designed to serve as a textbook at all, though it could certainly be used as one with highly motivated students.) That is, we do not present full developments of key theorems up front, leaving only routine exercises for the reader to consider. For one, we leave strategic gaps in the exposition for the reader to fill in. For another, we include a number of nonroutine problems, of the sort found on the IMO or related national competitions. The reader may or may not succeed in solving these, but attempting them should provide a solid test of one's understanding. In any case, solutions to the exercises and problems are included in the back; we have kept these brief, and they are only intended to make sense once you have already thought a bit about the corresponding exercises/problems on your own.

In addition to the MOP (and in some sense the Socratic method), inspirations for this approach include the famous Moore method of learning through problems, and the number theory curriculum of the late Arnold Ross's renowned summer mathematics program[3] for high school students. We also take inspiration from the slender classic *Geometry Revisited* by H.S.M. Coxeter and S. Greitzer, among whose pages this author discovered the beauty of Euclidean geometry so carefully hidden by many textbook writers. Indeed, we originally considered titling this book *"Geometry Revisited" Revisited* in homage to the masters; we ultimately chose instead to follow Aeschylus and Percy Bysshe Shelley in depicting geometry as a titanic subject released from the shackles of school curricula.

[3] Arnold Ross may no longer be with us, but fortunately his program is: its web site is `http://www.math.ohio-state.edu/ross/`.

Structure of the book

Aside from this introduction, the book is divided into four parts. The first part, "Rudiments", is devoted to the foundations of Euclidean geometry and to some of the most pervasive ideas within the subject. The second part, "Special situations", treats some common environments of classical synthetic geometry; it is here where one encounters many of the challenging Olympiad problems which helped inspire this book. The third part, "The roads to modern geometry", consists of two[4] chapters which treat slightly more advanced topics (inversive and projective geometry). The fourth part, "Odds and ends", is the back matter of the book, to be consulted as the need arises; it includes hints for the exercises and problems (for more on the difference, see below), plus bibliographic references, suggestions for further reading, information about the open source license, and an index.

Some words about terminology are in order at this point. For the purposes of this book, a *theorem* is an important result which either is given with its proof, or is given without its proof because inclusion of a proof would lead too far afield. In the latter case, a reference is provided. A *corollary* is a result which is important in its own right, but is easily deduced from a nearby theorem. A *fact* is a result which is important but easy enough to deduce that its proof is left to the reader.

Most sections of the text are accompanied by a section labeled "Problems", which are additional assertions which the reader is challenged to verify. Some of these are actually what we would call *exercises*, i.e., results which the reader should not have any trouble proving on his/her own, given what has come before. By contrast, a true *problem* is a result that can be obtained using the available tools, but which also requires some additional insight. In part to avoid deterring the reader from trying the more challenging problems (but also to forestall some awkwardness in cross-referencing), we have used the term "problem" in both cases. Hints have been included in the back matter of the book for selected problems; in order that the hints may also cover facts, some problems take the form "Prove Fact 21.13." In order to keep the book to a manageable size, and also to avoid challenging the reader's willpower, solutions have not been included; they may be instead found online at

[4] We would like to have additional such chapters, perhaps in a subsequent edition of the book, perhaps in a derivative version.

I have attributed my source for each problem to the best of my knowledge. Problems from the USA Mathematical Olympiad (USAMO), International Mathematical Olympiad (IMO), USA Team Selection Test (TST), and William Lowell Putnam competition (Putnam) are listed by year and number; problems from other national or regional contests are listed by country/contest and year. Problems I obtained from MOP are so labeled when I was unable to determine their true origins; most of these probably come from national contests. *Arbelos*[5] refers to Samuel Greitzer's student publication from 1982–1987 [8], and *Monthly* refers to the *American Mathematical Monthly*. Problems listed as "Original" are my own problems which have not before appeared in print (excluding prior versions of this book). Attributions to other people or web sites should be self-explanatory.

Acknowledgments

The acknowledgments for a book such as this cannot help but be at once tediously voluminous and hopelessly inadequate. That being so, there is nothing to done other than to proceed forthwith.

Let me start with those most directly involved. Thanks to Reid Barton for assembling a partial set of solutions to the included problems. Thanks to Marcelo Alvisio for expanding this solution set, for reporting numerous typos in the 1999 manuscript, and for rendering the missing diagrams from the 1999 manuscript using *The Geometer's Sketchpad*®. Thanks to Arthur Baragar for helpful (though not yet carried out) advice concerning the rendering of diagrams.

Let me next turn to those whose contributions are more diffuse. I first learned Euclidean geometry in the manner of this book from my instructors and later colleagues at the MOP, including Titu Andreescu, Răzvan Gelca, Anne Hudson, Gregg Patruno, and Dan Ullman. The participants of the 1997, 1998, and 1999 MOPs also deserve thanks for working through the notes that formed the basis for this book.

I owe a tremendous expository debt to Bjorn Poonen and Ravi Vakil, my collaborators on the 1985-2000 Putnam compilation [12]. In that volume, we embarked on a grand experiment: to forge a strong expository link between challenging "elementary" problems and "deep" mathematics. The warm re-

[5]For the origin of the name "arbelos", see Section 10.2.

ception received by that volume has emboldened me to apply to the present book some of what we learned from this experiment.

Thanks to the compilers of the wonderfully comprehensive MacTutor History of Mathematics, available online at

`http://www-gap.dcs.st-and.ac.uk/~history/index.html`.

We have used MacTutor as our reference for historical comments, English spellings of names, and birth and death dates. (All dates are A.D. unless denoted B.C.E.[6])

[6]The latter stands for "Before the Common Era", while the former might be puckishly deciphered as "Arbitrary Demarcation".

Part I

Rudiments

Chapter 1

Construction of the Euclidean plane

The traditional axiomatic development of Euclidean geometry originates with the treatment by Euclid of Alexandria (325?–265? B.C.E.) in the classic *Elements*, and was modernized by David Hilbert (1862–1943) in his 1899 *Grundlagen der Geometrie (Foundations of Geometry)*. For the purposes of this book, however, it is more convenient to start with the point of view of a coordinate plane, as introduced by René Descartes[1] (1596–1690) and Pierre de Fermat (1601–1665). We will return to the axiomatic point of view in due course, when we discuss hyperbolic geometry in Chapter 10; however, the coordinate-based point of view will also recur when we dabble briefly in algebraic geometry (see Section 11.7).

Of course one must assume *something* in order to get started. What we are assuming are the basic properties of the real numbers, which should not be too much of an imposition. The subtlest of these properties is the *least upper bound property*: every set of real numbers which is bounded above has a least upper bound. More precisely, if S is a set of real numbers and there exists a real number x such that $x \geq y$ for all $y \in S$, then there is a (unique) real number z such that:

(a) $z \geq y$ for all $y \in S$;

(b) if $x \geq y$ for all $y \in S$, then $x \geq z$.

[1]This attribution explains the term "Cartesian coordinates" to refer to this type of geometric description.

One theme we carry through our definitions is that certain numerical quantities (lengths of segments along a line, areas, arc and angle measures) should be treated with special algebraic rules, including systematic sign conventions. In so doing, one can make some statements more uniform, by eliminating some dependencies on the relative positions of points. This uniformity was unavailable to Euclid in the absence of negative numbers, hampering efforts to maintain logical consistency; see Section 1.6 for a tricky example.

In any case, while the strictures of logic dictate that this chapter must occur first, the reader need not be so restricted. We recommend skipping this chapter on first reading and coming back a bit later, once one has a bit of a feel for what is going on. But do make sure to come back at some point: for mathematicians (such as the reader and the author) to communicate, it is always of the utmost importance to agree on the precise definitions of even the simplest of terms.[2]

1.1 The coordinate plane, points and lines

We start by using Cartesian coordinates to define the basic geometric concepts: points, lines, and so on. The reader should not think his/her intelligence is being insulted by our taking space to do this: it is common in a mathematical text to begin by defining very simple objects, if for no other reason than to make sure the author and reader agree on the precise meaning and usage of fundamental words, as well as on the notation to be used to symbolize them (see previous footnote).

For our purposes, the *plane* $\mathbb{R}^2$ is the set of ordered pairs (x, y) of real numbers; we call those pairs the *points* of the plane. A *line* (or to be more precise, a *straight line*[3]) will be any subset of the plane of the form

$$\{(x, y) \in \mathbb{R}^2 : ax + by + c = 0\}$$

for some real numbers a, b, c with a and b not both zero. Then as one expects, any two distinct points $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$ lie on a unique line,

[2] As mathematician/storyteller Lewis Carroll, *né* Charles Lutwidge Dodgson, put it in the voice of *Through the Looking Glass* character Humpty Dumpty: "When I use a word... it means exactly what I choose it to mean—neither more nor less."

[3] The term "straight line" may be redundant in English, but it is not so in other languages. For instance, the Russian term for a curve literally translates as "curved line".

namely

$$\overleftrightarrow{P_1P_2} = \{((1-t)x_1 + tx_2, (1-t)y_1 + ty_2) : t \in \mathbb{R}\}.$$

Similarly, we define the *ray*

$$\overrightarrow{P_1P_2} = \{((1-t)x_1 + tx_2, (1-t)y_1 + ty_2) : t \in [0, \infty)\}$$

and the *segment* (or *line segment*)

$$\overline{P_1P_2} = \{((1-t)x_1 + tx_2, (1-t)y_1 + ty_2) : t \in [0, 1]\}.$$

Any segment lies on a unique line, called the *extension* of the segment. We say points $P_1, \dots, P_n$ are *collinear* if they lie on a single line; if ℓ is that line, we say that $P_1, \dots, P_n$ lie on ℓ *in order* if for any distinct $i, j, k \in \{1, \dots, n\}$ with $i < j$, we have $i < k < j$ if and only if P_k lies on the segment P_iP_j. For $n = 3$, we also articulate this by saying that P_2 lies *between* P_1 and P_3. We say lines $\ell_1, \dots, \ell_n$ are *concurrent* if they contain (or "pass through") a single point.

We will postpone defining angles for the moment, but we may as well define parallels and perpendiculars now. We say two lines $ax + by + c = 0$ and $dx + ey + f = 0$ are *parallel* if $ae - bd = 0$, and *perpendicular* if $ae + bd = 0$. Then the following facts are easily verified.

Fact 1.1.1. *Through any given point, there is a unique line parallel/perpendicular to any given line.*

For ℓ a line and P a point, the intersection of ℓ with the perpendicular to ℓ through P is called the *foot* of the perpendicular through P.

Fact 1.1.2. *Given three lines ℓ_1, ℓ_2, ℓ_3, the following relations hold.*

If ℓ_1 and ℓ_2 are:	*and ℓ_2 and ℓ_3 are:*	*then ℓ_1 and ℓ_3 are:*
parallel	*parallel*	*parallel*
parallel	*perpendicular*	*perpendicular*
perpendicular	*parallel*	*perpendicular*
perpendicular	*perpendicular*	*parallel.*

Given a segment $\overline{P_1P_2}$, there is a unique point M on $\overline{P_1P_2}$ with $P_1M = MP_2$, called the *midpoint* of $\overline{P_1P_2}$. There is a unique line through M perpendicular to $\overleftrightarrow{P_1P_2}$, called the *perpendicular bisector* of $\overline{P_1P_2}$.

Of course there is nothing special about having only two dimensions; one can construct an n-dimensional Euclidean space for any n. In particular, it is not unusual to do this for $n = 3$, resulting in what we call *space geometry*[4] as opposed to *plane geometry.* Although we prefer for simplicity not to discuss space geometry, we will make occasional reference to it in problems.

1.2 Distances and circles

The *distance* between two points $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$ is defined by

$$P_1P_2 = d(P_1, P_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2};$$

we also describe this quantity as the *length* of the segment $\overline{P_1P_2}$. If $\overline{P_1P_2}$ and $\overline{P_3P_4}$ are collinear segments (i.e., the points P_1, P_2, P_3, P_4 are collinear), we define the *signed ratio of lengths* of the segments $\overline{P_1P_2}$ and $\overline{P_3P_4}$ to be the ratio P_1P_2/P_3P_4 if the intersection of the rays $\overrightarrow{P_1P_2}$ and $\overrightarrow{P_3P_4}$ is a ray, and $-P_1P_2/P_3P_4$ otherwise. (In the latter case, the intersection of the two rays may be a segment, a point, or the empty set.)

From the distance formula, one can verify the *triangle inequality.*

Fact 1.2.1 (Triangle inequality)**.** *Given points P_1, P_2, P_3, we have*

$$P_1P_2 + P_2P_3 \geq P_1P_3,$$

with equality if and only if P_3 lies on the segment $\overline{P_1P_2}$.

We can also define the distance from a point to a line.

Fact 1.2.2. *Let P be a point and $\overleftrightarrow{QR}$ a line. Let S be the intersection of $\overleftrightarrow{QR}$ with the line through P perpendicular to $\overleftrightarrow{QR}$. Then the minimum distance from P to any point on $\overleftrightarrow{QR}$ is equal to PS.*

We call this minimum the *distance* from P to $\overleftrightarrow{QR}$, and denote it $d(P, \overleftrightarrow{QR})$.

With a notion of distance in hand, we may define a *circle* (resp. a *disc* or *closed disc*, an *open disc*) as the set of points P in the plane with the property that $OP = r$ (resp. $OP \leq r$) for some point O (the *center*) and some positive real number r (the *radius*); note that both O and r are

[4]The term "solid geometry" is more common, but less consistent.

uniquely determined by the circle. We call the quantity $2r$ the *diameter* of the circle/disc. Given a closed disc with center O and radius r, we call the circle of the same center and radius the *boundary* of the disc; we call the open disc of the same center and radius the *interior* of the circle or of the closed disc.

Fact 1.2.3. *Any three distinct points which do not lie on a straight line lie on a unique circle.*

Any segment joining the center of a circle to a point on the circle is called a *radius*[5] of the circle. Any segment joining two points on a circle is called a *chord* of the circle. A chord passing through the center is called a *diameter*; clearly its length is twice the radius of the circle.

Fact 1.2.4. *Any line and circle intersect at either zero, one, or two points. Any two distinct circles intersect at either zero, one, or two points.*

A line and circle that meet at exactly one point are said to be *tangent*. If ω is a circle, A is a point on ω, and ℓ is a line through A, we will speak frequently of the "second intersection of ℓ and ω"; when the two are tangent, we mean this to be A itself.

Two distinct circles which meet in exactly one point are also said to be *tangent*. In this case, either one circle lies inside the other, in which case the two are said to be *internally tangent*, or neither circle contains the other, in which case the two are said to be *externally tangent*.

Two or more circles with the same center are said to be *concentric*; concentric circles which do not coincide also do not intersect.

Fact 1.2.5. *Through any point P on a circle ω, there is a unique line tangent to ω: it is the line perpendicular to the radius of ω ending at P.*

It will be useful later (in the classification of rigid motions; see Theorem 3.1.2) to have in hand the "triangulation principle"; this fact was used once upon a time for navigation at sea, and nowadays figures in the satellite-based navigation technology known as the Global Positioning System.

[5]This is the first of numerous occasions on which we use the same word to denote both a segment and its length. This practice stems from the fact that Euclid did not have an independent concept of "length", and instead viewed segments themselves as "numbers" to be manipulated arithmetically.

Fact 1.2.6. *Let A, B, C be distinct points. Then any point P in the plane is uniquely determined by the three distances PA, PB, PC; that is, if P, Q are points in the plane with $PA = QA, PB = QB, PC = QC$, then $P = Q$.*

For $n \geq 4$, if $P_1, \dots, P_n$ are distinct points and ω is a circle, we say that $P_1, \dots, P_n$ *lie on ω in that order* if $P_1, \dots, P_n$ lie on ω and the polygon $P_1 \cdots P_n$ is simple (hence convex).

Problems for Section 1.2

1. Prove Fact 1.2.6.

2. Let ω_1 and ω_2 be circles with respective centers O_1 and O_2 and respective radii r_1 and r_2, and let k be a real number not equal to 1. Prove that the set of points P such that

$$PO_1^2 - r_1^2 = k(PO_2^2 - r_2^2)$$

is a circle. (This statement will be reinterpreted later in terms of the power of a point with respect to a circle; see Section 6.1.)

3. (IMO 1988/1) Consider two circles of radii R and r $(R > r)$ with the same center. Let P be a fixed point on the smaller circle and B a variable point on the larger circle. The line $\overleftrightarrow{BP}$ meets the larger circle again at C. The perpendicular ℓ to $\overleftrightarrow{BP}$ at P meets the smaller circle again at A. (As per our convention, if ℓ is tangent to the circle at P, then we take $A = P$.)

   (i) Find the set of values of $BC^2 + CA^2 + AB^2$.

   (ii) Find the locus of the midpoint of $\overline{AB}$.

1.3 Triangles and other polygons

The word "polygon" can mean many slightly different things, depending on whether one allows self-intersections, repeated vertices, degeneracies, etc. So one has to be a bit careful when defining it, to make sure that everyone agrees on what is to be allowed.

Let $P_1, \dots, P_n$ be a sequence of at least three points in the plane. The *polygon* (or *closed polygon*) with vertices $P_1, \dots, P_n$ is the $(n + 1)$-tuple

$(P_1, \dots, P_n, U)$, where U is the union of the segments $\overline{P_1P_2}, \dots, \overline{P_{n-1}P_n}, \overline{P_nP_1}$. We typically refer to this polygon as $P_1 \cdots P_n$; each of the P_i is called a *vertex*[6] of the polygon, and each of the segments making up U is called a *side*.

The *perimeter* of a polygon is the sum of the lengths of its underlying segments. It is often convenient to speak of the *semiperimeter* of a polygon, which is simply half of the perimeter.

A polygon is *nondegenerate* if no two of its vertices are equal and no vertex lies on a segment of the polygon other than the two of which it is an endpoint. Note that for a nondegenerate polygon, the union of segments uniquely determines the vertices up to cyclic shift and reversal of the list.

A polygon is *simple* (or *non-self-intersecting*) if it is nondegenerate and no two segments of the polygon intersect except at a shared endpoint.

For $P_1 \cdots P_n$ a polygon, a *diagonal* of $P_1 \cdots P_n$ is any segment joining two nonconsecutive vertices. A simple polygon is *convex* if any two diagonals intersect (possibly at an endpoint).

Fact 1.3.1. *If the points $P_1, \dots, P_n$ lie on a circle, then the polygon $P_1 \cdots P_n$ is simple if and only if it is convex.*

If $P_1 \cdots P_n$ is a convex polygon, we define the *interior* of $P_1 \cdots P_n$ to be the set of points Q such that for each i, the segment P_iQ intersects the polygon only at its endpoint P_i.

A nondegenerate polygon of three, four, five or six sides is called a *triangle*, *quadrilateral*, *pentagon*, or *hexagon*, respectively. Since triangles will occur quite often in our discussions, we adopt some special conventions to deal with them. We will often refer to the triangle with vertices A, B, C as $\triangle ABC$, and we will list its sides in the order $\overline{BC}, \overline{CA}, \overline{AB}$. We will often refer to its side lengths as $a = BC, b = CA, c = AB$.

Let ABC be a triangle. If two of the sides AB, BC, CA have equal lengths, we say $\triangle ABC$ is *isosceles*; if all three sides have equal lengths, we say $\triangle ABC$ is *equilateral*. If the angles of ABC are all acute, we say ABC is *acute* .

Let $ABCD$ be a convex quadrilateral. If lines $\overleftrightarrow{AB}$ and $\overleftrightarrow{CD}$ are parallel, we say $ABCD$ is a *trapezoid*. If in addition lines $\overleftrightarrow{BC}$ and $\overleftrightarrow{DA}$ are parallel, we say $ABCD$ is a *parallelogram*. If in addition $\overleftrightarrow{AB} \perp \overleftrightarrow{BC}$, we say $ABCD$

[6] The standard plural of "vertex" is "vertices", although "vertexes" is also acceptable. What is not standard and should be avoided is the back-formation "vertice" as a synonym of "vertex".

is a *rectangle*. If in addition $AB = BC = CD = DA$, we say $ABCD$ is a *square*.

1.4 Areas of polygons

If $P_1 \cdots P_n$ is a polygon and $P_i = (x_i, y_i)$, we define the *directed/signed area* of $P_1 \cdots P_n$, denoted $[P_1 \cdots P_n]_\pm$, by the formula

$$[P_1 \cdots P_n]_\pm = \frac{1}{2}(x_1y_2 - x_2y_1 + x_2y_3 - x_3y_2 + \cdots + x_{n-1}y_n - x_ny_{n-1} + x_ny_1 - x_1y_n).$$

This formula is sometimes called the *surveyor's formula* or the *shoelace formula*; the latter name serves as a mnemonic in the following fashion. If one draws the $2 \times (n+1)$ matrix

$$\begin{pmatrix} x_1 & x_2 & \cdots & x_n & x_1 \\ y_1 & y_2 & \cdots & y_n & y_1 \end{pmatrix},$$

the terms of the shoelace formula are obtained by multiplying terms along the diagonals and attaching signs as follows:

We define the *area* of the polygon $P_1 \cdots P_n$, denoted $[P_1 \cdots P_n]$, to be the absolute value of its directed area.

Fact 1.4.1.

- *For any polygon $P_1 \cdots P_n$,*

$$[P_1 \cdots P_n]_\pm = [P_2 \cdots P_nP_1]_\pm = -[P_n \cdots P_2P_1]_\pm.$$

- *For any polygons $P_1 \cdots P_nXY$ and $YXQ_1 \cdots Q_m$,*

$$[P_1 \cdots P_nXY]_\pm + [YXQ_1 \cdots Q_m]_\pm = [P_1 \cdots P_nQ_1 \cdots Q_m]_\pm.$$

- *For any triangle ABC,*

$$[ABC] = \frac{1}{2}BC \times d(A, \overleftrightarrow{BC}).$$

  In particular, $[ABC] \neq 0$.

- *For any convex quadrilateral $ABCD$, $[ABC]_\pm$ and $[ABD]_\pm$ have the same (nonzero) sign.*

- *For any simple polygon $P_1 \cdots P_n$, the directed areas $[P_i P_j P_k]_\pm$ all have the same sign, and it is the same as the sign of $[P_1 \cdots P_n]_\pm$. (This follows from the previous parts of this Fact; do you see how?)*

For $P_1 \cdots P_n$ a convex polygon, we call the sign of $[P_1 \cdots P_n]_\pm$ the *orientation* of $P_1 \cdots P_n$; we refer to positive and negative orientations also as "counterclockwise" and "clockwise", respectively.

1.5 Areas of circles and measures of arcs

Everything we have discussed so far was described purely in terms of basic algebraic operations on the real numbers: addition, subtraction, multiplication, division and square roots. The area of a circle and the measure of an arc cannot be described quite so simply; one must use the least upper bound property.

Given a circle ω, the *area* of the circle is defined to be the least upper bound of the set of areas of convex polygons $P_1 \cdots P_n$ with vertices on ω; note that this set is indeed bounded, for instance by the area of any square containing ω in its interior.

Fact 1.5.1.
- *There exists a constant π such that the area of a circle of radius r is equal to πr^2.*
- *The area of a circle is also equal to the greatest lower bound of the set of areas of convex polygons $P_1 \cdots P_n$ containing ω in its interior.*

Next we consider arcs and their measures. Given three distinct points A, B, C on a circle ω, we define the *arc* $\widehat{ABC}$ as the set of points $D \in \omega$ such that the quadrilateral $ABCD$ is not simple (including the points A, B, C). Since ω is uniquely determined by any arc, we may unambiguously speak of the *center* and *radius* of an arc. There is a unique point M on $\widehat{ABC}$ with $AM = MC$ (namely the intersection of $\widehat{ABC}$ with the perpendicular bisector of AC), called the *midpoint* of $\widehat{ABC}$. If the line $\overleftrightarrow{AB}$ passes through the center of ω, we call $\widehat{ABC}$ a *semicircle*.

The polygon $P_1 \cdots P_n$ is said to be *inscribed* in $\widehat{ABC}$ or circle ω if $P_1, \ldots, P_n$ all lie on $\widehat{ABC}$. We also say that the polygon is *circumscribed* by the arc/circle. A polygon which can be circumscribed by some circle is said to be *cyclic*, and the unique circle which circumscribes it is called its *circumcircle* (or *circumscribed circle*) of the polygon; points which form the vertices of a cyclic polygon are said to be *concyclic*. The center and radius of the circumcircle of a cyclic polygon are referred to as the *circumcenter* and *circumradius*, respectively, of the polygon. Note that any triangle is cyclic, so we may speak of the circumcenter and circumradius of a triangle without any further assumptions.

To define the measure of an arc, we use the following fact.

Fact 1.5.2. *Let A and B be points on a circle ω with center O and radius r, and let the lines tangent to ω at A and B meet at C. Then*

$$\frac{2[OAB]}{r^2} \leq \frac{AB}{r} \leq \frac{2[OACB]}{r^2}.$$

It follows that the least upper bound of the perimeters of polygons inscribed in an arc/circle of radius r exists, and is equal to $\frac{1}{r}$ times the least upper bound of the areas of polygons inscribed in the arc/circle. We call this quantity the *circumference* of the arc/circle and call $\frac{1}{r}$ times the circumference the *measure* of the arc/circle, denoted $m(\widehat{ABC})$. In particular, by the previous fact, the measure of any circle is equal to 2π.

Fact 1.5.3. *If $ABCDE$ is a convex polygon inscribed in a circle ω, then the measure of $\widehat{ACE}$ is equal to the sum of the measures of $\widehat{ABC}$ and $\widehat{CDE}$.*

As with areas, it is sometimes convenient to give arcs a signed measure. For $\widehat{ABC}$ not a semicircle, we define the *signed measure* of the arc $\widehat{ABC}$, denote $m_\pm(\widehat{ABC})$, to be $m(\widehat{ABC})$ times the sign of $[ABC]_\pm$. We regard this as a quantity "modulo 2π", i.e., only well-defined up to adding multiples of 2π. (If $\widehat{ABC}$ is a semicircle, then π and $-\pi$ differ by a multiple of 2π, so we may declare either one to be $m_\pm(\widehat{ABC})$.) Despite the ambiguity thus introduced, it still makes sense to add, subtract and test for equality signed measures, and one has the following nice properties which are most definitely false for the ordinary measure.

- The signed measure $m_\pm(\widehat{ABC})$ depends only on A, C and the circle on which $\widehat{ABC}$ lies.
- If A, B, C, D, E lie on a circle, then $m_\pm(\widehat{ABC}) + m_\pm(\widehat{CDE}) = m_\pm(\widehat{ACE})$.

1.6 Angles and the danger of configuration dependence

Given distinct points A, B, C in the plane, choose a circle ω centered at B, let A' and C' be the intersections of the rays $\overrightarrow{BA}$ and $\overrightarrow{BC}$, respectively, with ω, and let B' be any point on ω such that the quadrilateral $BA'B'C'$ is convex. Then $m(\widehat{A'B'C'})$ is independent of the choices of ω and B'; we call it the *angle measure*, or simply the *angle*, between the rays $\overrightarrow{BA}$ and $\overrightarrow{BC}$ and denote it $\angle ABC$. We also use "angle" to describe the set consisting of the union of the two rays $\overrightarrow{BA}$ and $\overrightarrow{BC}$, and use the symbol $\angle ABC$ for this set as well.

The implied unit of angle measure above is called the *radian.* By tradition, we also measure angles in units of *degrees* , where $180°$ equal π radians.

We say that the angle $\angle ABC$ is *acute, right, obtuse* according to whether its measure is less than, equal to, or greater than $\pi/2$ (or in degrees, $90°$).

The *interior* of the angle $\angle ABC$ consists of all points D such that the quadrilateral $ABCD$ is convex. The set of points D in the interior of $\angle ABC$ such that $\angle ABD = \angle DBC$, together with B itself, form a ray; that ray is called the *internal angle bisector* of $\angle ABC$. The same term is applied to the line containing that ray. The line perpendicular to the internal angle bisector is called the *external angle bisector* of $\angle ABC$. We refer to the internal and external angle bisectors of $\angle ABC$ also as the internal and external angle bisectors, respectively, of the triangle $\triangle ABC$ at the vertex B.

Fact 1.6.1.

- *If ABC is a triangle, then $\angle ABC + \angle BCA + \angle CAB = \pi$.*
- *If $ABCD$ is a convex quadrilateral, then $ABCD$ is cyclic if and only if $\angle ABC = \pi - \angle CDA$ if and only if $\angle ABD = \angle ACD$.*
- *If $ABCD$ is a convex quadrilateral and ω is the circumcircle of $\triangle ABC$, then $\overleftrightarrow{AD}$ is tangent to ω at A if and only if $\angle ABC = \angle DAC$.*

In the *Elements*, all quantities are treated as unsigned, as was necessary at the time in the absence of negative numbers. (Indeed, Euclid did not attach any numbers to geometrical figures; rather, he would conflate a segment with its length, a polygon with its area, and so on.) As we have seen, this approach makes certain statements dependent on relative positions of points. It is common to not worry much about such issues, but one actually has to be careful. As an example, we offer the following "pseudotheorem" and corresponding "pseudoproof" from [14]; we encourage the reader to concoct other plausible pseudotheorems and pseudoproofs!

Pseudotheorem. *All triangles are isosceles.*

Pseudoproof. Let $\triangle ABC$ be a triangle, and let O be the intersection of the internal angle bisector of A with the perpendicular bisector of $\overline{BC}$, as in Figure 1.6.1.

Figure 1.6.1: The Pseudotheorem.

Let D, Q, R be the feet of perpendiculars from O to $\overleftrightarrow{BC}, \overleftrightarrow{CA}, \overleftrightarrow{AB}$, respectively. By symmetry across OD, $OB = OC$, while by symmetry across AO, $AQ = AR$ and $OQ = OR$. Now the right triangles $\triangle ORB$ and $\triangle OQC$ have equal legs $OR = OQ$ and equal hypotenuses $OB = OC$, so they are congruent, giving $RB = QC$. Finally, we conclude

$$AB = AR + RB = AQ + QC = AC,$$

and hence the triangle $\triangle ABC$ is isosceles. $\square$

Problems for Section 1.6

1. Where is the error in the proof of the Pseudotheorem?

1.7 Directed angle measures

To avoid the problems that led to the Pseudotheorem, it will be useful to have a sign convention for angles.[7] The obvious choice would be to regard angle measures as being defined modulo 2π, but it turns out better to regard them modulo π, as we shall see.

Given three distinct points A, B, C, define the *directed angle measure*, or simply the *directed angle*, $\measuredangle ABC$ as the signed arc measure $m_\pm(\widehat{A'B'C'})$ as a quantity modulo π.

[7]This sign convention would seem to be rather old, but we do not have precise information about its origins.

Yes, you read correctly: although the signed arc measure is well-defined up to multiples of 2π, we regard the directed angle measure as only well-defined up to adding multiples of π. One consequence is that we can unambiguously define the directed angle (modulo π) between two lines ℓ_1 and ℓ_2 as follows. If ℓ_1 and ℓ_2 are parallel, declare the directed angle $\measuredangle(\ell_1, \ell_2)$ to be zero. Otherwise, declare $\measuredangle(\ell_1, \ell_2) = \measuredangle ABC$ for A a point on ℓ_1 but not on ℓ_2, B the intersection of ℓ_1 and ℓ_2, and C a point on ℓ_2 not on ℓ_1.

Note that to avert some confusion, we will systemically distinguish between the words "signed", referring to arc measures modulo 2π, and "directed", referring to angle measures modulo π, even though in common usage these two terms might be interchanged. Such an interchange would be dangerous for us!

One can now verify the following rules of "directed angle arithmetic", all of which are independent of configuration.

Fact 1.7.1. *Let A, B, C, D, O, P denote distinct points in the plane.*

1. $\measuredangle ABC = -\measuredangle CBA$.
2. $\measuredangle APB + \measuredangle BPC = \measuredangle APC$.
3. $\measuredangle ABC = \measuredangle ABD$ *if and only if* B, C, D *are collinear. In particular,* $\measuredangle ABC = 0$ *if and only if* A, B, C *are collinear.*
4. $\measuredangle ABD = \measuredangle ACD$ *if and only if* A, B, C, D *are concyclic.*
5. $\measuredangle ABC = \measuredangle ACD$ *if and only if* CD *is tangent to the circle passing through* A, B, C.
6. $\measuredangle ABC + \measuredangle BCA + \measuredangle CAB = 0$.
7. $2\measuredangle ABC = \measuredangle AOC$ *if* A, B, C *lie on a circle centered at* O.
8. $\measuredangle ABC$ *equals* $\frac{1}{2}$ *of the measure of the arc* $\widehat{AC}$ *of the circumcircle of* ABC.

For example, if A, B, C, D lie on a circle in that order, then we have $\angle ABD = \angle ACD$ as undirected angles. On the other hand, if they lie on a circle in the order A, B, D, C, then we have $\angle ABD = \pi - \angle DCA$, so in terms of directed angles

$$\measuredangle ABD = \pi - \measuredangle DCA = -\measuredangle DCA = \measuredangle ADC.$$

It should be noted that this coincidence is a principal reason why one works modulo π and not 2π! (The other principal reason is of course so that collinear points always make an angle of 0.)

The last two assertions in Fact 1.7.1 ought to raise some eyebrows, because division by 2 is a dangerous thing when working modulo π. To be precise, the equation $2\angle A = 2\angle B$ of directed angles does not imply that $\angle A = \angle B$, for the possibility also exists that $\angle A = \angle B + \pi/2$. (Those familiar with elementary number theory will recognize an analogous situation: one cannot divide by 2 in the congruence $2a \equiv 2b \pmod{c}$ when c is even.) This explains why we do not write $\measuredangle ABC = \frac{1}{2}\measuredangle AOC$: the latter expression is not well-defined. On the other hand, directed arcs can be unambiguously measured mod 2π, so dividing a signed arc measure by 2 gives a directed angle measure mod π.

If all of this seems too much to worry about, do not lose hope; the conventions are easily learned with a little practice. We will illustrate this in Section 4.2.

Chapter 2

Algebraic methods

Since our very construction of the Euclidean plane was rooted in the algebra of the coordinate plane, it is clear that algebraic techniques have something to say about Euclidean geometry; indeed, we have already encountered a few problems that are naturally treated in terms of coordinates, and we will encounter more later (e.g., Theorem 6.5.1). However, coordinatizing a typical problem in Euclidean geometry leads to a complicated mess; one typically succeeds more easily by adopting more high-level techniques.[1]

We will spend much of the second part of the book introducing so-called "synthetic" techniques; for now, we introduce some techniques which, while still rooted in algebra, offer some advantages over blind coordinate manipulations.

2.1 Trigonometry

Define the points $O = (0, 0)$ and $P = (1, 0)$. Given a signed angle measure θ (modulo 2π), let Q be the point on the circle of radius 1 centered at O such that the signed angle measure of $\angle POQ$ is equal to θ. Let $\cos\theta$ and $\sin\theta$ denote the x-coordinate and y-coordinate, respectively, of Q; these define the *cosine* and *sine*. By the distance formula, we have the identity

$$\cos^2\theta + \sin^2\theta = 1.$$

[1] An analogous relationship in computer science is that between a processor-level machine language and a high-level programming language like C++, Java, or Perl.

Figure 2.1.1: Definition of the trigonometric functions.

Also define the *tangent*, *secant*, *cosecant*, and *cotangent* functions as follows (when these expressions are well-defined):

$$\tan\theta = \frac{\sin\theta}{\cos\theta}, \qquad \sec\theta = \frac{1}{\cos\theta}, \qquad \csc\theta = \frac{1}{\sin\theta}, \qquad \cot\theta = \frac{\cos\theta}{\sin\theta}.$$

We do not intend to conduct here a full course in trigonometry; we will content ourselves to summarizing the important facts and provide a few problems where trigonometry can or must be employed. Throughout the following discussion, let ABC denote a triangle, write $a = BC, b = CA, c = AB$, and write A, B, C for the measures of (undirected) angles $\angle CAB$, $\angle ABC$, $\angle BCA$, respectively. Let $s = (a+b+c)/2$ denote the semiperimeter of $\triangle ABC$.

Fact 2.1.1 (Law of Sines)**.** *The area of $\triangle ABC$ equals $\frac{1}{2}ab\sin C$. In particular,*

$$\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C}.$$

Fact 2.1.2 (Extended Law of Sines)**.** *If R is the circumradius of $\triangle ABC$, then $BC = 2R\sin A$.*

Fact 2.1.3 (Law of Cosines)**.** *In $\triangle ABC$,*

$$c^2 = a^2 + b^2 - 2ab\cos C.$$

Fact 2.1.4 (Addition formulae)**.** *For any real numbers α and β,*

$$\begin{aligned}
\cos(\alpha+\beta) &= \cos\alpha\cos\beta - \sin\alpha\sin\beta \\
\cos(\alpha-\beta) &= \cos\alpha\cos\beta + \sin\alpha\sin\beta \\
\sin(\alpha+\beta) &= \sin\alpha\cos\beta + \cos\alpha\sin\beta \\
\sin(\alpha-\beta) &= \sin\alpha\cos\beta - \cos\alpha\sin\beta.
\end{aligned}$$

Using the addition formulae, one can convert products of sines and cosines to sums, and vice versa.

Fact 2.1.5 (Sum-to-product formulae)**.**

$$
\begin{aligned}
\sin\alpha + \sin\beta &= 2\sin\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2} \\
\sin\alpha - \sin\beta &= 2\cos\frac{\alpha+\beta}{2}\sin\frac{\alpha-\beta}{2} \\
\cos\alpha + \cos\beta &= 2\sin\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2} \\
\cos\alpha - \cos\beta &= -2\sin\frac{\alpha+\beta}{2}\sin\frac{\alpha-\beta}{2}.
\end{aligned}
$$

In particular, one has the double and half-angle formulae.

Fact 2.1.6 (Double-angle formulae)**.**

$$
\begin{aligned}
\sin 2\alpha &= 2\sin\alpha\cos\alpha \\
\cos 2\alpha &= 2\cos^2\alpha - 1 = 1 - 2\sin^2\alpha \\
\tan 2\alpha &= \frac{2\tan\alpha}{1-\tan^2\alpha}.
\end{aligned}
$$

Fact 2.1.7 (Half-angle formulae)**.**

$$
\begin{aligned}
\sin\frac{\alpha}{2} &= \pm\sqrt{\frac{1-\cos\alpha}{2}} \\
\cos\frac{\alpha}{2} &= \pm\sqrt{\frac{1+\cos\alpha}{2}} \\
\tan\frac{\alpha}{2} &= \csc\alpha - \cot\alpha.
\end{aligned}
$$

The half-angle formulae take a convenient form for triangles.

Fact 2.1.8. *In $\triangle ABC$,*

$$
\begin{aligned}
\sin\frac{C}{2} &= \sqrt{\frac{(s-a)(s-b)}{ab}} \\
\cos\frac{C}{2} &= \sqrt{\frac{s(s-c)}{ab}}.
\end{aligned}
$$

It may be helpful at times to express certain other quantities associated with a triangle in terms of the angles.

Fact 2.1.9. *If $\triangle ABC$ has inradius r and circumradius R, then*

$$r = 4R \sin \frac{A}{2} \sin \frac{B}{2} \sin \frac{C}{2}.$$

We leave the construction of other such formulae to the reader.

Problems for Section 2.1

1. For any triangle ABC, prove that $\tan A + \tan B + \tan C = \tan A \tan B \tan C$ and that $\cot A/2 + \cot B/2 + \cot C/2 = \cot A/2 \cot B/2 \cot C/2$.

2. Show that if none of the angles of a convex quadrilateral $ABCD$ is a right angle, then
$$\frac{\tan A + \tan B + \tan C + \tan D}{\tan A \tan B \tan C \tan D} = \cot A + \cot B + \cot C + \cot D.$$

3. Find a formula for the area of a triangle in terms of two angles and the side opposite the third angle. More generally, given any data that uniquely determines a triangle, one can find an area formula in terms of that data. Some of these can be found in Fact 7.9.1; can you come up with some others?

4. (USAMO 1996/5) Triangle ABC has the following property: there is an interior point P such that $\angle PAB = 10^\circ$, $\angle PBA = 20^\circ$, $\angle PCA = 30^\circ$ and $\angle PAC = 40^\circ$. Prove that triangle ABC is isosceles. (For an added challenge, find a non-trigonometric solution!)

5. (IMO 1985/1) A circle has center on the side AB of a cyclic quadrilateral $ABCD$. The other three sides are tangent to the circle. Prove that $AD + DC = AB$.

2.2 Vectors

A *vector* in the plane can be defined either as an arrow, where addition of arrows proceeds by the "tip-to-tail" rule illustrated in Figure 2.2.1, or as an ordered pair (x, y) recording the difference in the x and y coordinates between the tip and the tail. Vectors in a Euclidean space of three or more dimensions may be defined similarly.

Figure 2.2.1: Vector addition.

It is important to remember that a vector is not a point, but rather the "difference of two points"; it encodes relative, not absolute, position. In practice, however, one chooses a point as the origin and identifies a point with the vector from the origin to that point. (In effect, one puts the tails of all of the arrows in one place.)

The standard operations on vectors include addition and subtraction, multiplication by real numbers (positive, negative or zero), and the *dot product*, defined geometrically as

$$\vec{A} \cdot \vec{B} = \|\vec{A}\| \cdot \|\vec{B}\| \cos \angle AOB,$$

where O is the origin, and in coordinates as

$$(a_x, a_y) \cdot (b_x, b_y) = a_x b_x + a_y b_y.$$

The key fact here is that $\vec{A} \cdot \vec{B} = 0$ if and only if $\vec{A}$ and $\vec{B}$ are perpendicular.

A more exotic operation is the *cross product*, which is defined for a pair of vectors in three-dimensional space as follows:

$$(a_x, a_y, a_z) \times (b_x, b_y, b_z) = (a_y b_z - a_z b_y, a_z b_x - a_x b_z, a_x b_y - a_y b_x).$$

Geometrically speaking, $\vec{A} \times \vec{B}$ is perpendicular to both $\vec{A}$ and $\vec{B}$ and has length

$$\|\vec{A} \times \vec{B}\| = \|\vec{A}\| \cdot \|\vec{B}\| \sin \angle AOB.$$

This length equals the area of the parallelogram with vertices $0, \vec{A}, \vec{A}+\vec{B}, \vec{B}$, or twice the area of the triangle with vertices $0, \vec{A}, \vec{B}$. The sign ambiguity can be resolved by the *right-hand rule* (see Figure 2.2.2): if you point the fingers of your right hand along $\vec{A}$, then swing them toward $\vec{B}$, your thumb points in the direction of $\vec{A} \times \vec{B}$.

Figure 2.2.2: The right-hand rule.

Fact 2.2.1. *The following identities hold:*

1. *(Triple scalar product identity)* $\vec{A} \cdot (\vec{B} \times \vec{C}) = \vec{B} \cdot (\vec{C} \times \vec{A}) = \vec{C} \cdot (\vec{A} \times \vec{B})$. *(Moreover, this quantity equals the volume of a parallelepiped with edges $\vec{A}, \vec{B}, \vec{C}$, although we did not rigorously define either volumes or parallelepipeds.)*

2. *(Triple cross product identity)* $\vec{A} \times (\vec{B} \times \vec{C}) = (\vec{C} \cdot \vec{A})\vec{B} - (\vec{B} \cdot \vec{A})\vec{C}$.

Problems for Section 2.2

1. (Romania, 1997) Let $ABCDEF$ be a convex hexagon, and let $P = \overleftrightarrow{AB} \cap \overleftrightarrow{CD}, Q = \overleftrightarrow{CD} \cap \overleftrightarrow{EF}, R = \overleftrightarrow{EF} \cap \overleftrightarrow{AB}$, $S = \overleftrightarrow{BC} \cap \overleftrightarrow{DE}, T = \overleftrightarrow{DE} \cap \overleftrightarrow{FA}, U = \overleftrightarrow{FA} \cap \overleftrightarrow{BC}$. Prove that
$$\frac{PQ}{CD} = \frac{QR}{EF} = \frac{RP}{AB} \quad \text{if and only if} \quad \frac{ST}{DE} = \frac{TU}{FA} = \frac{US}{BC}.$$

2. (Răzvan Gelca) Let $ABCD$ be a convex quadrilateral and $O = \overline{AC} \cap \overline{BD}$. Let M, N be points on $\overline{AB}$ so that $AM = MN = NB$, and let P, Q be points on $\overline{CD}$ so that $CP = PQ = QD$. Show that triangles $\triangle MOP$ and $\triangle NOQ$ have the same area.

3. (MOP 1996) Let $ABCDE$ be a convex pentagon, and let F, G, H, I, J be the respective midpoints of $\overline{CD}, \overline{DE}, \overline{EA}, \overline{AB}, \overline{BC}$. If $\overline{AF}, \overline{BG}, \overline{CH}, \overline{DI}$ pass through a common point, show that $\overline{EJ}$ also passes through this point.

4. (Austria-Poland, 1979) Let A, B, C, D be points in space, let M be the midpoint of $\overline{AC}$, and let N be the midpoint of $\overline{BD}$. Prove that
$$4MN^2 = AB^2 + BC^2 + CD^2 + DA^2 - AC^2 - BD^2.$$

2.3 Complex numbers

The set of *complex numbers* consists of the expressions of the form $a + bi$ for a, b real numbers, added and multiplied according to the rules
$$\begin{aligned}(a + bi) + (c + di) &= (a + c) + (b + d)i \\ (a + bi) \times (c + di) &= (ac - bd) + (ad + bc)i.\end{aligned}$$

We identify the real number a with the complex number $a+0i$, and we write bi for the complex number $0+bi$. In particular, the complex number $i = 1i$ is a square root of -1.

One uses complex numbers in Euclidean geometry by identifying a point with Cartesian coordinates (x, y) with the complex number $x + yi$. This represents an extension of vector techniques, incorporating a convenient interpretation of angles (and of similarity transformations; see Section 3.4).

Another interesting use of complex numbers is to prove inequalities. This use exploits the fact that the *magnitude*

$$|a+bi| = \sqrt{a^2+b^2}$$

is multiplicative:

$$|(a+bi)\times(c+di)| = |a+bi|\times|c+di|.$$

Consider the following example (compare with Problem 10.3.9).

Theorem 2.3.1 (Ptolemy's inequality)**.** *Let A, B, C, D be four points in the plane. Then*

$$AC\cdot BD \le AB\cdot CD + BC\cdot DA,$$

with equality if and only if the quadrilateral $ABCD$ is convex (or degenerate) and cyclic.

Proof. Regard A, B, C, D as complex numbers; then we have an identity

$$(A-C)(B-D) = (A-B)(C-D) + (B-C)(A-D).$$

However, the magnitude of $(A-C)(B-D)$ is precisely the product of the lengths of the segments AC and BD, and likewise for the other terms. Thus the desired inequality is simply the triangle inequality applied to these three quantities! (The equality condition is left as an exercise.) $\square$

Although we will not have use of it for a while (not until Section 11.7), we mention now an important, highly nontrivial fact about complex numbers.

Theorem 2.3.2 (Fundamental theorem of algebra)**.** *For every polynomial $P(z) = a_nz^n + \cdots + a_0$ with $a_n, \ldots, a_0$ complex numbers, and $a_n, \ldots, a_1$ not all zero, there exists a complex number z such that $P(z) = 0$.*

This theorem is traditionally attributed to (Johann) Carl Friedrich Gauss[2] (1777-1855), who gave the first proof (and several in addition) that would pass modern standards of rigor. However, a correct proof (which relies on the existence of splitting fields, a concept unavailable at the time) had already been sketched by Leonhard Euler (1707–1783) .

Problems for Section 2.3

1. Prove that x, y, z lie at the corners of an equilateral triangle if and only if either $x + \omega y + \omega^2 z = 0$ or $x + \omega z + \omega^2 y = 0$, where $\omega = e^{2\pi i/3}$.

2. Construct equilateral triangles externally (internally) on the sides of an arbitrary triangle $\triangle ABC$. Prove that the circumcenters of these three triangles form another equilateral triangle. This triangle is known as the *inner (outer) Napoleon*[3] *triangle* of $\triangle ABC$.

3. Let P, Q, R, S be the circumcenters of squares constructed externally on sides $\overline{AB}$, $\overline{BC}$, $\overline{CD}$, $\overline{DA}$, respectively, of a convex quadrilateral $ABCD$. Show that the segments $\overline{PR}$ and $\overline{QS}$ are perpendicular to each other and equal in length.

4. Let $ABCD$ be a convex quadrilateral. Construct squares $CDKL$ and $ABMN$ externally on sides $\overline{AB}$ and $\overline{CD}$. Show that if the midpoints of $\overline{AC}, \overline{BD}, \overline{KM}, \overline{NL}$ do not coincide, then they form a square.

5. (IMO 1977/1) Equilateral triangles $\triangle ABK$, $\triangle BCL$, $\triangle CDM$, $\triangle DAN$ are constructed inside the square $ABCD$. Prove that the midpoints of the four segments $\overline{KL}$, $\overline{LM}$, $\overline{MN}$, $\overline{NK}$ and the midpoints of the eight segments $\overline{AK}, \overline{BK}$, $\overline{BL}, \overline{CL}$, $\overline{CM}, \overline{DM}$, $\overline{DN}, \overline{AN}$ are the twelve vertices of a regular dodecagon. (Nowadays the IMO tends to avoid geometry problems such as this one, which have no free parameters, but they are relatively common in single-answer contests such as ARML.)

6. Given a point P on a circle and the vertices A_1, $A_2, \ldots, A_n$ of an inscribed regular n-gon, prove that:

[2] Note that in German, "Gauss" is sometimes spelled "Gauß". That last character is not a beta, but rather an "s-zed", a ligature of the letters "s" and "z".

[3] Yes, that's French emperor Napoleon Bonaparte (1769–1821), though it's not clear how his name came to be attached to this result.

1. $PA_1^2 + PA_2^2 + \cdots + PA_n^2$ is a constant (independent of P).
2. $PA_1^4 + PA_2^4 + \cdots + PA_n^4$ is a constant (independent of P).

7. (China, 1998) Let P be an arbitrary point in the plane of triangle $\triangle ABC$ with side lengths $BC = a$, $CA = b$, $AB = c$, and put $PA = x$, $PB = y$, $PC = z$. Prove that

$$ayz + bzx + cxy \geq abc,$$

with equality if and only if P is the circumcenter of $\triangle ABC$.

2.4 Barycentric coordinates and mass points

Given a triangle $\triangle ABC$ and a point P, we define the *barycentric coordinates*[4] of P with respect to ABC as the triple

$$\left(\frac{[PBC]_\pm}{[ABC]_\pm}, \frac{[APC]_\pm}{[ABC]_\pm}, \frac{[ABP]_\pm}{[ABC]_\pm}\right)$$

of real numbers. Note that these numbers always add up to 1, and that they are all positive if and only if P lies in the interior of $\triangle ABC$.

A related concept is that of "mass points". A *mass point* consists of a pair (P, r), where P is a point and r is a positive real number. These points may be "added" as follows:

$$((x_1, y_1), r_1) + (x_2, y_2), r_2) = \left(\left(\frac{r_1 x_1 + r_2 x_2}{r_1 + r_2}, \frac{r_1 y_1 + r_2 y_2}{r_1 + r_2}\right), r_1 + r_2\right).$$

In terms of vectors, we have

$$(P_1, r_1) + (P_2, r_2) = \frac{r_1}{r_1 + r_2} P_1 + \frac{r_2}{r_1 + r_2} P_2.$$

In terms of physics, the location of the sum of two mass points is the "center of mass" of a pair of appropriate masses at the two points. This addition law is associative, so we may likewise add three or more mass points unambiguously.

The relation of mass points to barycentric coordinates is as follows.

[4] The word "barycentric" comes from the Greek "barys", meaning "heavy"; it is cognate to the term "baryon" in particle physics.

Fact 2.4.1. *Let $\triangle ABC$ be a triangle. Then for any positive real numbers r_1, r_2, r_3, the sum of the mass points (A, r_1), (B, r_2), (C, r_3) is located at the point with barycentric coordinates*

$$\left(\frac{r_1}{r_1+r_2+r_3}, \frac{r_2}{r_1+r_2+r_3}, \frac{r_3}{r_1+r_2+r_3}\right).$$

Chapter 3

Transformations

In geometry, it is often useful to study transformations of the plane (i.e. functions mapping the plane to itself) preserving certain properties. In fact, Felix Klein (1849-1925) went so far as to define "geometry" as the study of properties invariant under a particular set of transformations!

In this section we describe some fundamental transformations and how they interact with properties of "figures" in the plane. Here and throughout, by a *figure* we simply mean a set of points in the plane. Also, we follow standard usage in mathematical English and refer to a function also as a "map".

3.1 Congruence and rigid motions

Let F_1 and F_2 be two figures and suppose $f : F_1 \to F_2$ is a bijection (one-to-one correspondence). We say that F_1 *and* F_2 *are congruent (via* f*)*, and write $F_1 \cong F_2$, if we have an equality of distances $PQ = f(P)f(Q)$ for all $P, Q \in F_1$. When F_1 and F_2 are polygons with the same number of vertices and f is not specified, we assume it is the map that takes the vertices of F_1 to the vertices of F_2 in the order that they are listed. For instance, the fact that $\triangle ABC \cong \triangle DEF$ are congruent means that $AB = DE, BC = EF, CA = FD$.

A *rigid motion* of the Euclidean plane is a map from the plane to itself which preserves distances; that is, if P maps to P' and Q to Q', then we have $PQ = P'Q'$. In other words, a rigid motion maps every figure to a congruent figure. Here is a list of examples of rigid motions (which we will soon find to be exhaustive; see Theorem 3.1.2):

- *Translation*: each point moves a fixed distance in a fixed direction, so that $PQQ'P'$ is always a parallelogram.
- *Rotation* with center O and angle θ: each point P maps to the point P' such that $OP = OP'$ and $\angle POP' = \theta$, where the angle is signed (i.e., measured modulo 2π, not π). We refer to a rotation with angle π also as a *half-turn.*
- *Reflection* through the line ℓ: each point P maps to the point P' such that ℓ is the perpendicular bisector of PP'.
- *Glide reflection* along the line ℓ: reflection through ℓ followed by a translation along ℓ.

Theorem 3.1.1. *Given two congruent figures, each not contained in any line, there is a unique rigid motion that maps one onto the other (matching corresponding points).*

Note that the rigid motion may not be unique if it is not required to match corresponding points between the two figures: for instance, a regular polygon is mapped to itself by more than one rotation (as for that matter is a circle).

Proof. We first address the uniqueness. If there were two rigid motions carrying the first figure to the second, then composing one with the inverse of the other would yield a nontrivial rigid motion leaving one entire figure in place. By assumption, however, this figure contains three noncollinear points A, B, C, and a point P is uniquely determined by its distances to these three points (Fact 1.2.6), so every point is fixed by the rigid motion, a contradiction. Thus the motion is unique if it exists.

Now we address existence. let A, B, C be three noncollinear points of the first figure, and A', B', C' the corresponding points of the second figure. There exists a translation mapping A to A'; following that with a suitable rotation (since $AB = A'B'$), we can ensure that B also maps to B'. Now we claim C maps either to C' or to its reflection across $\overleftrightarrow{A'B'}$; in other words, given two points A, B and a point C not on $\overleftrightarrow{AB}$, C is determined up to reflection across $\overleftrightarrow{AB}$ by the distances AC and BC. This holds because this data fixes C to lie on two distinct circles, which can only intersect in two points (Fact 1.2.4).

Now if P is any point of the first figure, then P is uniquely determined by the distances AP, BP, CP (again by Fact 1.2.6), and so it must map to the corresponding point of the second figure. This completes the proof of existence. □

Note that rigid motions carry convex polygons to convex polygons. It follows that a rigid motion either preserves the orientation of all convex polygons, or reverses the orientation of all convex polygons. We say two congruent figures are *directly congruent* if the unique rigid motion taking one to the other (provided by Theorem 3.1) preserves orientations, and *oppositely congruent* otherwise.

Theorem 3.1.2. *A rigid motion preserves orientations if and only if it is a translation or a rotation. A rigid motion reverses orientations if and only if it is a reflection or a glide reflection.*

Proof. Let ABC be a triangle carried to the triangle $A'B'C'$ under the rigid motion. First suppose the rigid motion preserves orientations; by the uniqueness assertion in Theorem 3.1, it suffices to exhibit either a translation or a rotation carrying $\triangle ABC$ to $\triangle A'B'C'$. If the perpendicular bisectors of $\overline{AA'}$ and $\overline{BB'}$ are parallel, then $ABB'A'$ is a parallelogram, so there is a translation taking A to A' and B to B'. Otherwise, let these perpendicular bisectors meet at O. Draw the circle through B and B' centered at O; there are (at most) two points on this circle whose distance to A' is the length AB. One point is the reflection of B across the perpendicular bisector of $\overline{AA'}$; by our assumption, this cannot equal B'. Thus B' is the other point, which is the image of B under the rotation about O taking B to B'.

Figure 3.1.1: Proof of Theorem 3.1.2.

Next suppose the rigid motion reverses orientations; again by Theorem 3.1, it suffices to exhibit either a reflection or a glide reflection carrying $\triangle ABC$ to $\triangle A'B'C'$. The lines through which $\overleftrightarrow{AB}$ reflects to a line parallel to $\overleftrightarrow{A'B'}$ form two perpendicular families of parallel lines. In each family there is one line passing through the midpoint of $\overline{AA'}$; the glide reflection through this line taking A to A' takes B either to B' or to its half-turn about A'. In the latter case, switching to the other family gives a glide reflection taking B to B'. As in the first case, C automatically goes to C', and we are done. □

In particular, the composition of two rotations is either a rotation or translation. In fact, one can say more.

Fact 3.1.3. *The composition of a rotation of angle θ_1 with a rotation of angle θ_2 is a rotation of angle $\theta_1 + \theta_2$ if this is not a multiple of 2π, and a translation otherwise.*

On the other hand, given two rotations, it is not obvious where the center of their composition is; in particular, it generally depends on the order of the rotations. (In the language of abstract algebra, the group of rigid motions is not commutative.)

Fact 3.1.4. *Let $\triangle ABC$ and $\triangle A'B'C'$ be two triangles. The following conditions are all equivalent to $\triangle ABC \cong \triangle A'B'C'$.*

(a) (SSS criterion) We have $AB = A'B'$, $BC = B'C'$, $CA = C'A'$ (this equivalence is by definition).

(b) (SAS criterion) We have $AB = A'B'$, $BC = B'C'$, and $\angle ABC = \angle A'B'C'$.

(c) (ASA criterion) We have $AB = A'B'$, and all three (or even any two) of $\angle ABC$, $\angle BCA$, $\angle CAB$ are equal to the corresponding angles $\angle A'B'C'$, $\angle B'C'A'$, $\angle C'A'B'$.

Problems for Section 3.1

1. Show that there is no "SSA criterion" for congruence, by exhibiting two noncongruent triangles $\triangle ABC, \triangle A'B'C'$ with $AB = A'B'$, $BC = B'C'$, $\angle BCA = \angle B'C'A'$.

2. (MOP 1997) Consider a triangle ABC with $AB = AC$, and points M and N on $\overline{AB}$ and $\overline{AC}$, respectively. The lines $\overleftrightarrow{BN}$ and $\overleftrightarrow{CM}$ intersect at P. Prove that $\overleftrightarrow{MN}$ and $\overleftrightarrow{BC}$ are parallel if and only if $\angle APM = \angle APN$.

3. (Butterfly theorem) Let A, B, C, D be points occurring in that order on circle ω and put $P = \overline{AC} \cap \overline{BD}$. Let EF be a chord of ω passing through P, and put $Q = \overline{BC} \cap \overline{EF}$ and $R = \overline{DA} \cap \overline{EF}$. Then $PQ = PR$.

4. (IMO 1986/2) A triangle $A_1A_2A_3$ and a point P_0 are given in the plane. We define $A_s = A_{s-3}$ for all $s \geq 4$. We construct a sequence of points $P_1, P_2, P_3, \dots$ such that P_{k+1} is the image of P_k under rotation with center A_{k+1} through angle 120° clockwise (for $k = 0, 1, 2, \dots$). Prove that if $P_{1986} = P_0$, then $\triangle A_1A_2A_3$ is equilateral.

5. (MOP 1996) Let AB_1C_1, AB_2C_2, AB_3C_3 be directly congruent equilateral triangles. Prove that the pairwise intersections of the circumcircles of triangles AB_1C_2, AB_2C_3, AB_3C_1 form an equilateral triangle congruent to the first three.

3.2 Similarity and homotheties

Let F_1 and F_2 be two figures and suppose $f : F_1 \to F_2$ is a bijection. We say that *F_1 and F_2 are similar (via f)*, and write $F_1 \sim F_2$, if there exists a positive real number c such that for all $P, Q \in F_1$, $f(P)f(Q) = cPQ$. Again, when F_1 and F_2 are polygons with the same number of vertices and f is not specified, we assume it is the map that takes the vertices of F_1 to the vertices of F_2 in the order that they are listed. For instance, the fact that $\triangle ABC \sim \triangle DEF$ means that $AB/DE = BC/EF = CA/FD$.

Fact 3.2.1. *Let $\triangle ABC$ and $\triangle A'B'C'$ be two triangles. The following conditions are all equivalent to $\triangle ABC \sim \triangle A'B'C'$.*

(a) (SSS criterion) We have $AB/A'B' = BC/B'C' = CA/C'A'$ (this equivalence is by definition).

(b) (SAS criterion) We have $AB/A'B' = BC/B'C'$, and $\angle ABC = \angle A'B'C'$.

(c) (AA criterion) All three (or even any two) of $\angle ABC$, $\angle BCA$, $\angle CAB$ are equal to the corresponding angles $\angle A'B'C'$, $\angle B'C'A'$, $\angle C'A'B'$.

A *similarity* of the Euclidean plane is a map from the plane to itself for which there exists a positive real number c such that whenever P maps to P' and Q to Q', we have $P'Q' = cPQ$; the constant c is called the *ratio of similitude* of the similarity. Note that a rigid motion is precisely a similarity with ratio of similitude 1. By imitating the proof of Theorem 3.1, we have the following result.

Fact 3.2.2. *Given two similar figures, each not contained in a line, there is a unique similarity that maps one onto the other (matching corresponding points).*

As for rigid motions, a similarity either preserves the orientation of all convex polygons, or reverses the orientation of all convex polygons. We say two similar figures are *directly similar* if the unique similarity taking one to the other (provided by Fact 3.2.2) preserves orientations, and *oppositely similar* otherwise.

In the spirit of Theorem 3.1.2, we will shortly give a classification of similarities (Theorem 3.4.1); before doing so, however, we introduce a special class of similarities which by themselves are already surprisingly useful. Given a point O and a nonzero real number r, the *homothety* (or *dilation* or *dilatation*) with center O and ratio r maps each point P to the point P' on the line $\overleftrightarrow{OP}$ such that the signed ratio of lengths OP'/OP is equal to r. Note that r is allowed to be negative; in particular, a homothety of ratio -1 is simply a half-turn.

Figure 3.2.1: A homothety.

Homotheties have the property that they map every segment of a figure to a parallel segment. Aside from translations (which might be thought of as degenerate homotheties with center "at infinity"), this property characterizes homotheties; the following theorem is often useful as a concurrence criterion.

Fact 3.2.3. *Two directly similar but not congruent figures with corresponding sides parallel are homothetic. In particular, the lines $\overleftrightarrow{AA'}$, where A and A' are corresponding points, all pass through a common point.*

As for rotations, we conclude that the composition of two homotheties is a homothety, though again it is less than obvious where the center is!

Problems for Section 3.2

1. Given a triangle ABC, construct (with straightedge and compass) a square with one vertex on $\overline{AB}$, one vertex on $\overline{CA}$, and two (adjacent) vertices on $\overline{BC}$.

2. (USAMO 1992/4) Chords $\overline{AA'}$, $\overline{BB'}$, $\overline{CC'}$ of a sphere meet at an interior point P but are not contained in a plane. The sphere through A, B, C, P is tangent to the sphere through A', B', C', P. Prove that $AA' = BB' = CC'$.

3. (Putnam 1996/A-2) Let C_1 and C_2 be circles whose centers are 10 units apart, and whose radii are 1 and 3. Find, with proof, the locus of all points M for which there exists points X on C_1 and Y on C_2 such that M is the midpoint of the line segment $\overline{XY}$.

4. Given three nonintersecting circles, draw the intersection of the external tangents to each pair of the circles. Show that these three points are collinear.

5. (Russia, 2003) Let ABC be a triangle with $AB \neq AC$. Point E is such that $AE = BE$ and $\overleftrightarrow{BE} \perp \overleftrightarrow{BC}$. Point F is such that $AF = CF$ and $\overleftrightarrow{CF} \perp \overleftrightarrow{BC}$. Let D be the point on line $\overleftrightarrow{BC}$ such that $\overleftrightarrow{AD}$ is tangent to the circumcircle of triangle ABC. Prove that D, E, F are collinear.

3.3 Spiral similarities

Let P be a point, let θ be a signed angle measure (i.e., measured modulo 2π) and let r be a nonzero real number. The *spiral similarity* of angle θ and ratio r centered at P consists of a homothety of ratio r centered at P followed by a rotation of angle θ centered at P. (The order of these operations do not matter; one easy way to see this is to express both operations in terms of complex numbers.) In particular, a spiral similarity is the composition of a similarity and a rigid motion, and hence is a similarity.

Problems for Section 3.3

1. (USAMO 1978/2) The squares $ABCD$ and $A'B'C'D'$ represent maps of the same region, drawn to different scales and superimposed. Prove that there is only one point O on the small map which lies directly over point O' of the large map such that O and O' represent the same point of the country. Also, give a Euclidean construction (straightedge and compass) for O.

2. (MOP 1998) Let $ABCDEF$ be a cyclic hexagon with $AB = CD = EF$. Prove that the intersections of $\overline{AC}$ with $\overline{BD}$, of $\overline{CE}$ with $\overline{DF}$, and of $\overline{EA}$ with $\overline{FB}$ form a triangle similar to $\triangle BDF$.

3. Let C_1, C_2, C_3 be circles such that C_1 and C_2 meet at distinct points A and B, C_2 and C_3 meet at distinct points C and D, and C_3 and C_1 meet at distinct points E and F. Let P_1 be an arbitrary point on C_1, and define points $P_2, \dots, P_7$ as follows:

$$\begin{array}{ll}
P_2 & \text{is the second intersection of line } \overleftrightarrow{P_1A} \text{ with } C_2; \\
P_3 & \text{is the second intersection of line } \overleftrightarrow{P_2C} \text{ with } C_3; \\
P_4 & \text{is the second intersection of line } \overleftrightarrow{P_3E} \text{ with } C_1; \\
P_5 & \text{is the second intersection of line } \overleftrightarrow{P_1B} \text{ with } C_2; \\
P_6 & \text{is the second intersection of line } \overleftrightarrow{P_2D} \text{ with } C_3; \\
P_7 & \text{is the second intersection of line } \overleftrightarrow{P_3F} \text{ with } C_1.
\end{array}$$

Prove that $P_7 = P_1$.

3.4 Complex numbers and the classification of similarities

One can imagine generating similarities in a rather complicated fashion, e.g., take a homothety about one center followed by a rotation about a different center. It turns out that this does not really yield anything new; this can be seen most easily by interpreting similarities in terms of complex numbers.

Theorem 3.4.1. *Every orientation-preserving similarity is either a translation or a spiral similarity.*

Proof. First we show that every orientation-preserving similarity can be expressed in terms of homothety, translation, and rotation. Let A and B be two points with images C and D. If we perform a homothety about A of ratio CD/AB, then a translation mapping A to C, then a suitable rotation, we get another similarity mapping A and B, respectively, to C and D. On the other hand, if P is any point not on the line $\overleftrightarrow{AB}$, and Q and Q' are its images under the original similarity and the new similarity, then the triangles $\triangle ABP$, $\triangle CDQ$, $\triangle CDQ'$ are all similar. This implies that C, Q, Q' are collinear and that D, Q, Q' are collinear, forcing $Q = Q'$. In other words, the original similarity coincides with the new one.

The basic transformations can be expressed in terms of complex numbers as follows:

Translation by vector v	$z \mapsto z + v$
Homothety of ratio r, center x	$z \mapsto r(z - x) + x$
Rotation by angle θ, center x	$z \mapsto e^{i\theta}(z - x) + x$

The point is that each of these maps has the form $z \mapsto az+b$ for some complex numbers a, b with $a \neq 0$, and hence all orientation-preserving similarities have this form.

If $a = 1$, the map $z \mapsto az+b$ of complex numbers represents a translation by b. If $a \neq 1$, then let $t = b/(1 - a)$ be the unique solution of $t = at + b$. Then our map can be written $z \mapsto a(z - t) + t$, which is clearly a spiral similarity about t. $\square$

Problems for Section 3.4

1. Let A, C, E be three points on a circle. A $60°$ rotation about the center of the circle carries A, C, E to B, D, F, respectively. Prove that the triangle whose vertices are the midpoints of $\overline{BC}, \overline{DE}, \overline{FA}$ is equilateral.

3.5 Affine transformations

The last type of transformation we introduce in this chapter is the most general, at the price of preserving the least structure. However, for sheer strangeness it does not rival either inversion (see Chapter 10) or projective transformations (see Chapter 11).

An *affine transformation* is a map from the plane to itself of the form

$$(x, y) \mapsto (ax + by + c, dx + ey + f)$$

for some real numbers a, b, c, d, e, f with $ae - bd \neq 0$; this last condition ensures that the map is a bijection. From the proof of Theorem 3.4.1, we see that every similarity is an affine transformation. However, there are more exotic affine transformations, including the stretch $(x, y) \mapsto (x, cy)$ and the shear $(x, y) \mapsto (x + y, y)$.

These last examples demonstrate that angles and distances do not behave predictably under affine transformation. However, one does have the following result, as well as a partial converse (Problem 3.5.1).

Fact 3.5.1. *Affine transformations preserve collinearity of points, parallelness of lines, and concurrence of lines. Moreover, the affine transformation*

$$(x, y) \mapsto (ax + by + c, dx + ey + f)$$

multiplies areas by a factor of $|ae - bd|$*, and preserves orientation if and only if* $ae - bd > 0$*.*

Fact 3.5.2. *Any three noncollinear points can be mapped to any three other noncollinear points by a unique affine transformation.*

As an example of the use of the affine transformation, we offer the following theorem.

Theorem 3.5.3. *Let* $ABCDE$ *be a convex pentagon, and let* $F = \overleftrightarrow{BC} \cap \overleftrightarrow{DE}$*,* $G = \overleftrightarrow{CD} \cap \overleftrightarrow{EA}$*,* $H = \overleftrightarrow{DE} \cap \overleftrightarrow{AB}$*,* $I = \overleftrightarrow{EA} \cap \overleftrightarrow{BC}$*,* $J = \overleftrightarrow{AB} \cap \overleftrightarrow{DE}$*. Suppose that the areas of the triangles* $\triangle AHI, \triangle BIJ, \triangle CJF, \triangle DFG, \triangle EGH$ *are all equal. Then the lines* $\overleftrightarrow{AF}, \overleftrightarrow{BG}, \overleftrightarrow{CH}, \overleftrightarrow{DI}, \overleftrightarrow{EJ}$ *are all concurrent.*

Figure 3.5.1: Diagram for Theorem 3.5.3.

Proof. Everything in the theorem is preserved by affine transformations, so we may place three of the points anywhere we want. Let us assume that A, C, D form an isosceles triangle with $AC = AD$ and $\angle CDA = \pi/5$, which is to say that A, C, D are three vertices of a regular pentagon.

Our first observation is that since $\triangle CJF$ and $\triangle DFG$ have equal areas, so do $\triangle CFG$ and $\triangle FDJ$, by adding the area of $\triangle CDH$ to both sides. By the base-height formula, this means GJ is parallel to CD, and similarly for the other sides. In particular, since $\triangle ACD$ was assumed to be isosceles, F lies on the internal angle bisector of $\angle DAC$, and J and C are the reflections of G and D across $\overleftrightarrow{AF}$.

We next want to show that B and E are mirror images across $\overleftrightarrow{AF}$. To that end, let E' and H' be the reflections of E and H, respectively. Since the lines $\overleftrightarrow{FC}$ and $\overleftrightarrow{FD}$ are mirror images across $\overleftrightarrow{AF}$, we know that E' lies on $\overleftrightarrow{BD}$, and similarly H' lies on $\overleftrightarrow{AC}$. Suppose that $E'D < BD$, or equivalently that E is closer than B is to the line $\overleftrightarrow{CD}$. Then we also have $DH' < CI'$; since $CJ = DG$, we deduce $JH' < JI$. Now it is evident that the triangle $E'H'J$, being contained in $\triangle BJI$, has smaller area; on the other hand, it has the same area as $\triangle EHG$, which by assumption has the same area as $\triangle BJI$, a contradiction. So we cannot have $E'D < BD$, or $E'D > BD$ by a similar argument. We conclude $E'D = BD$, i.e. B and E are mirror images across $\overleftrightarrow{AF}$.

Figure 3.5.2: Proof of Theorem 3.5.3.

In particular, this implies that $\overleftrightarrow{BE}$ is parallel to $\overleftrightarrow{CD}$. Since we could just as well have put B, D, E at the vertices of an isosceles triangle, we also may conclude $AC \parallel DE$ and so forth.

Now let ℓ be the line through C parallel to $\overleftrightarrow{AD}$; by the above argument, we know B is the intersection of ℓ with $\overleftrightarrow{DF}$. On the other hand, B is also the intersection of ℓ with the line through A parallel to $\overleftrightarrow{CF}$. If we move F towards A along the internal angle bisector of $\angle DAC$, the intersection of $\overleftrightarrow{DF}$ with ℓ moves away from C, but the intersection of the parallel to $\overleftrightarrow{CF}$ through A with ℓ moves closer to C. Hence these can only coincide for at most one choice of F, and of course they do coincide when $ABCDE$ is a regular pentagon. We conclude that $ABCDE$ is the image of a regular pentagon under an affine transformation, which in particular implies that $\overleftrightarrow{AF}, \overleftrightarrow{BG}, \overleftrightarrow{CH}, \overleftrightarrow{DI}, \overleftrightarrow{EJ}$ are concurrent. □

Figure 3.5.3: Proof of Theorem 3.5.3.

Some authors choose to independently speak of *oblique coordinates*, measured with respect to two coordinate axes which are not necessarily perpendicular. We prefer to think of these as the result of perfoming an affine transformation to a pair of Cartesian coordinate axes.

Problems for Section 3.5

1. Prove that any transformation that takes collinear points to collinear points is an affine transformation.

2. Let $\triangle ABC$ be a triangle, and let X, Y, Z be points on sides $\overline{BC}$, $\overline{CA}$, $\overline{AB}$, respectively, satisfying
$$\frac{BX}{XC} = \frac{CY}{YA} = \frac{AZ}{ZB} = k,$$
where k is a given constant greater than 1. Find, in terms of k, the ratio of the area of the triangle formed by the three segments $\overline{AX}$, $\overline{BY}$, $\overline{CZ}$ to the area of $\triangle ABC$. (Compare Problem 4.1.3, which is the case $k = 2$.)

3. In the hexagon $ABCDEF$, opposite sides are equal and parallel. Prove that triangles $\triangle ACE$ and $\triangle BDF$ have the same area.

4. (Greece, 1996) In a triangle ABC the points D, E, Z, H, Θ are the midpoints of the segments $\overline{BC}, \overline{AD}, \overline{BD}, \overline{ED}, \overline{EZ}$, respectively. If I is the point of intersection of $\overline{BE}$ and $\overline{AC}$, and K is the point of intersection of $\overline{H\Theta}$ and $\overline{AC}$, prove that
   1. $AK = 3CK$;
   2. $HK = 3H\Theta$;
   3. $BE = 3EI$;
   4. the area of $\triangle ABC$ is 32 times that of $\triangle E\Theta H$.

5. (Sweden, 1996) Through a point in the interior of a triangle with area T, draw lines parallel to the three sides, partitioning the triangle into three triangles and three parallelograms. Let T_1, T_2, T_3 be the areas of the three triangles. Prove that
$$\sqrt{T} = \sqrt{T_1} + \sqrt{T_2} + \sqrt{T_3}.$$

6. (France, 1996) Let $\triangle ABC$ be a triangle. For any line ℓ not parallel to any side of $\triangle ABC$, let G_ℓ be the vector average of the intersections of ℓ with $\overleftrightarrow{BC}, \overleftrightarrow{CA}, \overleftrightarrow{AB}$. Determine the locus of G_ℓ as ℓ varies.

Chapter 4

Tricks of the trade

We conclude our presentation of fundamentals with a chapter that highlights a small core of basic techniques that prove useful in a large number of problems. The point is to show how much one can accomplish even with very little advanced knowledge.

4.1 Slicing and dicing

One of the most elegant ways of establishing a geometric result is to dissect the figure into pieces, then rearrange the pieces so that the result becomes obvious. The quintessential example of this technique is the proof of the Pythagorean theorem[1] given by the Indian mathematician Bhaskara (Bhaskaracharya) (1114-1185), which consists of a picture plus only one word.

Theorem 4.1.1 (Pythagoras). *If $\triangle ABC$ is a right triangle with hypotenuse $\overline{BC}$, then $AB^2 + BC^2 = AC^2$.*

Proof. Behold! $\square$

Other useful dissections include a proof of the fact that the area of a triangle is half its base times its height (Figure 4.1.1), a proof that the median to the hypotenuse of a right triangle divides the triangle into two isosceles triangles (Figure 4.1.2), and in three dimensions, an embedding of a tetrahedron in a rectangular parallelepiped (Figure 4.1.3).

[1] The list of other authors who have given proofs of the Pythagorean theorem in this vein is a long one, but surely the oddest name on it is U.S. President James A. Garfield (1831–1881).

Figure 4.1.1: Area of a triangle equals half base times height.

Figure 4.1.2: Median to the hypotenuse of a right triangle.

Problems for Section 4.1

1. (MOP 1997) Let Q be a quadrilateral whose side lengths are a, b, c, d, in that order. Show that the area of Q does not exceed $(ac + bd)/2$.

2. Let $\triangle ABC$ be a triangle and M_A, M_B, M_C the midpoints of the sides $\overline{BC}, \overline{CA}, \overline{AB}$, respectively. Show that the triangle with side lengths AM_A, BM_B, CM_C has area $3/4$ that of $\triangle ABC$.

3. In triangle $\triangle ABC$, points D, E, F are marked on sides $\overline{BC}, \overline{CA}, \overline{AB}$, respectively, so that
$$\frac{BD}{DC} = \frac{CE}{EA} = \frac{AF}{FB} = 2.$$
Show by a dissection argument that the triangle formed by the lines $\overleftrightarrow{AD}, \overleftrightarrow{BE}, \overleftrightarrow{CF}$ has area $1/7$ that of $\triangle ABC$. (Compare Problem 3.5.2.)

4. Give a dissection solution to Problem 3.5.3.

5. (For those familiar with space geometry) The 1982 SAT (an American college entrance exam) included a question asking for the number of faces of the polyhedron obtained by gluing a regular tetrahedron to a square pyramid along one of the triangular faces. The answer expected by the test authors was 7, since the two polyhedra have 9 faces, 2 of which are removed by gluing. However, a student taking the exam pointed out that this is incorrect! What is the correct answer, and why?

6. (For those familiar with space geometry) A regular tetrahedron and a regular octahedron have edges of the same length. What is the ratio between their volumes?

7. Given four segments which form a convex cyclic quadrilateral of a given radius, the same is true no matter what order the segments occur in. Prove that the resulting quadrilaterals all have the same area. (This will be evident later from Brahmagupta's formula; see Fact 8.2.2.)

Figure 4.1.3: A tetrahedron embedded in a box.

4.2 Angle chasing

A surprising number of propositions in Euclidean geometry can be established using nothing more than careful bookkeeping of angles, which allows one to detect similar triangles, cyclic quadrilaterals, and the like. In some problem circles, this technique is known as "angle chasing". As an example of angle chasing in action, we offer a theorem first published in 1838 by one A. Miquel.

Figure 4.2.1: Miquel's theorem.

Theorem 4.2.1. *Let $\triangle ABC$ be a triangle and let P, Q, R be any points on the sides $\overline{BC}, \overline{CA}, \overline{AB}$, respectively. Then the circumcircles of triangles ARQ, BPR, CQP pass through a common point.*

Proof. Let T be the second intersection (other than R) of the circumcircles of ARQ and BPR. By collinearity of points,

$$\angle TQA = \pi - \angle CQT, \quad \angle TRB = \pi - \angle ART, \quad \angle TPC = \pi - \angle BPT.$$

In a convex cyclic quadrilateral, opposite angles are supplementary. Therefore

$$\angle TQA = \pi - \angle ART, \angle TRB = \pi - \angle BPT.$$

We conclude $\angle TPC = \pi - \angle CQT$. Now conversely, a convex quadrilateral whose opposite angles are supplementary is cyclic. Therefore T also lies on the circumcircle of $\triangle CQP$, as desired. $\square$

A defect of the angle chasing technique is that the relevant theorems depend on the configuration of the points involved, particularly on whether certain points fall between certain other points. For example, one might ask whether the above theorem still holds if P, Q, R are allowed to lie on the extensions of the sides of $\triangle ABC$. It does hold, but the above proof breaks down because some of the angles claimed to be equal become supplementary, and vice versa.

The trick to getting around this is to use "mod π directed angles" as described in Section 1.7. To illustrate how "directed angle chasing" works, we give an example which is both simple and important: it intervenes in our proof of Pascal's theorem (Theorem 6.3.1).

Figure 4.2.2: Diagram for Theorem 4.2.2.

Theorem 4.2.2. *Suppose that the circles ω_1 and ω_2 intersect at distinct points A and B. Let $\overline{CD}$ be any chord on ω_1, and let E and F be the second intersections of the lines $\overleftrightarrow{CA}$ and $\overleftrightarrow{BD}$, respectively, with ω_2. Then $\overleftrightarrow{EF}$ is parallel to $\overleftrightarrow{CD}$.*

Proof. We chase directed angles as follows:

$$\begin{aligned}
\measuredangle CDF &= \measuredangle CDB && \text{(collinearity of } B, D, F\text{)} \\
&= \measuredangle CAB && \text{(cyclic quadrilateral } ABCD\text{)} \\
&= \measuredangle EAB && \text{(collinearity of } A, C, E\text{)} \\
&= \measuredangle EFB && \text{(cyclic quadrilateral } ABEF\text{).}
\end{aligned}$$

Hence the lines $\overleftrightarrow{CD}$ and $\overleftrightarrow{EF}$ make the same angle with $\overleftrightarrow{BF}$, and so are parallel. $\square$

Directed angles can be expressed in terms of lines as well as in terms of points: the directed angle $\measuredangle(\ell_1, \ell_2)$ between lines ℓ_1 and ℓ_2 can be interpreted as the angle of any rotation carrying ℓ_1 to a line parallel to ℓ_2. This alternate perspective simplifies some proofs, as in the following example; for a situation where this diagram occurs, see Problem 6.4.3.

Theorem 4.2.3. *Let $\triangle ABC$ be a triangle. Suppose that the lines ℓ_1 and ℓ_2 are perpendicular, and meet each side (or its extension) in a pair of points symmetric across the midpoint of the side. Then the intersection of ℓ_1 and ℓ_2 is concyclic with the midpoints of the three sides.*

Figure 4.2.3: Diagram for Theorem 4.2.3.

Proof. Let M_A, M_B, M_C be the midpoints of the sides $\overline{BC}, \overline{CA}, \overline{AB}$, respectively, and put $P = \ell_1 \cap \ell_2$. Since the lines $\ell_1, \ell_2, \overleftrightarrow{BC}$ form a right triangle and M_A is the midpoint of the hypotenuse of that triangle, the triangle formed by the points $P, M_A, \ell_2 \cap BC$ is isosceles with

$$\measuredangle(\overleftrightarrow{M_AP}, \ell_2) = \measuredangle(\ell_2, \overleftrightarrow{BC}).$$

By a similar argument,

Figure 4.2.4: Proof of Theorem 4.2.3.

$$\measuredangle(\ell_2, \overleftrightarrow{M_BP}) = \measuredangle(\overleftrightarrow{CA}, \ell_2),$$

and adding these gives

$$\measuredangle M_APM_B = \measuredangle ACB = \measuredangle M_AM_CM_B$$

since the sides of the triangle $\triangle M_AM_BM_C$ are parallel to those of $\triangle ABC$. We conclude that M_A, M_B, M_C, P are concyclic, as desired. $\square$

Problems for Section 4.2

1. (USAMO 1994/3) A convex hexagon $ABCDEF$ is inscribed in a circle such that $AB = CD = EF$ and diagonals $\overline{AD}, \overline{BE}, \overline{CF}$ are concurrent. Let P be the intersection of $\overline{AD}$ and $\overline{CE}$. Prove that $CP/PE = (AC/CE)^2$.

2. (IMO 1990/1) Chords $\overline{AB}$ and $\overline{CD}$ of a circle intersect at a point E inside the circle. Let M be an interior point of the segment $\overline{EB}$. The tangent line of E to the circle through D, E, M intersects the lines $\overleftrightarrow{BC}$ and $\overleftrightarrow{AC}$ at F and G, respectively. If $AM/AB = t$, find EG/EF in terms of t.

3. Let $\triangle A_0B_0C_0$ be a triangle and P a point. Define a new triangle whose vertices A_1, B_1, C_1 as the feet of the perpendiculars from P to B_0C_0, C_0A_0, A_0B_0, respectively. Repeat the construction twice, starting with $\triangle A_1B_1C_1$, to produce the triangles $\triangle A_2B_2C_2$ and $\triangle A_3B_3C_3$. Show that $\triangle A_3B_3C_3$ is similar to $\triangle A_0B_0C_0$.

Figure 4.3.1: Diagram for Theorem 4.3.1.

4. (MOP 1991) Two circles intersect at points A and B. An arbitrary line through B intersects the first circle again at C and the second circle again at D. The tangents to the first circle at C and the second at D intersect at M. Through the intersection of $\overleftrightarrow{AM}$ and $\overleftrightarrow{CD}$, there passes a line parallel to CM and intersecting $\overleftrightarrow{AC}$ at K. Prove that $\overleftrightarrow{BK}$ is tangent to the second circle.

5. Let ω_1, ω_2, ω_3, ω_4 be four circles in the plane. Suppose that ω_1 and ω_2 intersect at P_1 and Q_1, ω_2 and ω_3 intersect at P_2 and Q_2, ω_3 and ω_4 intersect at P_3 and Q_3, and ω_4 and ω_1 intersect at P_4 and Q_4. Show that if P_1, P_2, P_3, and P_4 lie on a line or circle, then Q_1, Q_2, Q_3, and Q_4 also lie on a line or circle. (This is tricky; see the proof of Theorem 10.1.2.)

4.3 Working backward

A common stratagem, when trying to prove that a given point has a desired property, is to construct a phantom point with the desired property, then reason backwards to show that it coincides with the original point. We illustrate this point with an example.

Theorem 4.3.1. *Suppose the triangles $\triangle ABC$ and $\triangle AB'C'$ are directly similar. Then the points A, B, C, $\overleftrightarrow{BB'} \cap \overleftrightarrow{CC'}$ lie on a circle.*

Proof. Since we want to show that $\overleftrightarrow{BB'} \cap \overleftrightarrow{CC'}$ lies on the circle through A, B, C, and analogously on the circle through A, B', C', we define the point P to be the intersection of these two circles. Then

$$\measuredangle APB = \measuredangle ACB = \measuredangle AC'B' = \measuredangle APB'$$

and so P lies on the line $\overleftrightarrow{BB'}$, and similarly on the line $\overleftrightarrow{CC'}$. $\square$

Problems for Section 4.3

1. (IMO 1994/2) Let $\triangle ABC$ be an isosceles triangle with $AB = AC$. Suppose that

(i) M is the midpoint of $\overline{BC}$ and O is the point on the line $\overleftrightarrow{AM}$ such that $\overleftrightarrow{OB}$ is perpendicular to $\overleftrightarrow{AB}$;

(ii) Q is an arbitrary point on the segment $\overline{BC}$ different from B and C;

(iii) E lies on the line $\overleftrightarrow{AB}$ and F lies on the line $\overleftrightarrow{AC}$ such that E, Q, F are distinct and collinear.

Prove that $\overleftrightarrow{OQ}$ is perpendicular to $\overleftrightarrow{EF}$ if and only if $QE = QF$.

2. (USAMO 2005/3) Let $\triangle ABC$ be an acute-angled triangle, and let P and Q be two points on side $\overline{BC}$. Construct point C_1 in such a way that convex quadrilateral $APBC_1$ is cyclic, $\overleftrightarrow{QC_1} \parallel \overleftrightarrow{CA}$, and C_1 and Q lie on opposite sides of line $\overleftrightarrow{AB}$. Construct point B_1 in such a way that convex quadrilateral $APCB_1$ is cyclic, $\overleftrightarrow{QB_1} \parallel \overleftrightarrow{BA}$, and B_1 and Q lie on opposite sides of line $\overleftrightarrow{AC}$. Prove that points B_1, C_1, P, and Q lie on a circle.

3. (Morley's theorem) Let $\triangle ABC$ be a triangle, and for each side, draw the intersection of the two angle trisectors closer to that side. (That is, draw the intersection of the trisectors of A and B closer to AB, and so on.) Prove that these three intersections determine an equilateral triangle.

Figure 4.3.2: Morley's theorem (Problem 4.3.3).

Part II

Special situations

Chapter 5

Concurrence and collinearity

This chapter is devoted to the study of two fundamental and reciprocal questions: when are three given points collinear, and when are three given lines concurrent or parallel? (The idea that parallel lines should be considered "concurrent" is an idea from the theory of perspective; it will be fleshed out in our discussion of projective geometry in Chapter 11.) This study only begins here; the themes of collinearity and concurrence recur throughout this text, so it is worth mentioning places where they occur beyond this chapter.

- The Pascal and Brianchon theorems (Section 6.3).
- The radical axis theorem (Section 6.2).

5.1 Concurrent lines: Ceva's theorem

We begin with a simple but useful result, published in 1678 by the Italian engineer Giovanni Ceva (1647-1734). In his honor, a segment joining a vertex of a triangle to a point on the opposite side is called a *cevian*[1].

Figure 5.1.1: Ceva's theorem (Theorem 5.1.1).

[1]Depending on who you ask, this word is pronounced either CHAY-vee-un or CHEH-vee-un. I've heard other pronunciations as well, but I don't recommend them.

Theorem 5.1.1 (Ceva)**.** *Let $\triangle ABC$ be a triangle, and let P, Q, R be points on the lines $\overleftrightarrow{BC}, \overleftrightarrow{CA}, \overleftrightarrow{AB}$, respectively, none equal to any of A, B, C. Then the lines $\overleftrightarrow{AP}, \overleftrightarrow{BQ}, \overleftrightarrow{CR}$ are concurrent or parallel if and only if*

$$\frac{BP}{PC}\frac{CQ}{QA}\frac{AR}{RB} = 1 \tag{5.1.1.1}$$

as an equality of signed ratios of lengths.

Proof. First suppose that $\overleftrightarrow{AP}, \overleftrightarrow{BQ}, \overleftrightarrow{CR}$ concur at a point T. Then the ratio of lengths BP/PC is equal, by similar triangles, to the ratio of the distances from B and C to $\overleftrightarrow{AP}$. On the other hand, that ratio is also equal to the ratio of areas $[ATB]/[CTA]$, since we can calculate these areas as half of base times height, with $\overline{AT}$ as the base. Moreover, the signed ratios BP/PC and $[ATB]_\pm/[CTA]_\pm$ also have the same sign, so are in fact equal.

Figure 5.1.2: Proof of Ceva's theorem (Theorem 5.1.1).

By this argument, we get

$$\frac{BP}{PC}\frac{CQ}{QA}\frac{AR}{RB} = \frac{[ATB]_\pm}{[CTA]_\pm}\frac{[BTC]_\pm}{[ATB]_\pm}\frac{[CTA]_\pm}{[BTC]_\pm} = 1.$$

In case $\overleftrightarrow{AP}, \overleftrightarrow{BQ}, \overleftrightarrow{CR}$ are parallel, we may deduce the same conclusion by continuity, or directly: we leave this to the reader.

Conversely, suppose that (5.1.1.1) holds; we will apply the trick of working backward. The lines $\overleftrightarrow{AP}$ and $\overleftrightarrow{BQ}$ meet at some point T, and the line $\overleftrightarrow{CT}$ meets $\overleftrightarrow{AB}$ at some point R'. (If $\overleftrightarrow{AP}$ and $\overleftrightarrow{BQ}$ are parallel, interpret $\overleftrightarrow{CT}$ as the common parallel to these lines through C, and the previous sentence will still make sense.) By construction, $\overleftrightarrow{AP}, \overleftrightarrow{BQ}, \overleftrightarrow{CR'}$ are concurrent. However, using Ceva in the other direction (which we just proved), we find that

$$\frac{BP}{PC}\frac{CQ}{QA}\frac{AR'}{R'B} = 1.$$

Combining this equation with (5.1.1.1) yields

$$\frac{AR}{RB} = \frac{AR'}{R'B}.$$

Since $AR + RB = AR' + R'B = AB$, adding 1 to both sides gives

$$\frac{AB}{RB} = \frac{AB}{R'B}$$

as a signed ratio of lengths, from which we conclude that $RB = R'B$, and hence $R = R'$. $\square$

In certain cases, Ceva's Theorem is more easily applied in the following form ("trig Ceva").

Fact 5.1.2 (Ceva's theorem, trigonometric form). *Let $\triangle ABC$ be a triangle, and let P, Q, R be any points in the plane distinct from A, B, C, respectively. Then $\overleftrightarrow{AP}, \overleftrightarrow{BQ}, \overleftrightarrow{CR}$ are concurrent if and only if*

$$\frac{\sin \angle CAP}{\sin \angle PAB} \frac{\sin \angle ABQ}{\sin \angle QBC} \frac{\sin \angle BCR}{\sin \angle RCA} = 1.$$

One can either deduce this from Ceva's theorem or prove it directly. Be careful when using trig Ceva with directed angles, as signs matter: the ratio $(\sin \angle CAP)/(\sin \angle PAB)$ must be defined in terms of angles modulo 2π, but the sign of the ratio itself only depends on the line $\overleftrightarrow{AP}$, not on the choice of P on one side or the other of A.

Problems for Section 5.1

1. Suppose that the cevians $\overline{AP}, \overline{BQ}, \overline{CR}$ meet at T. Prove that
$$\frac{TP}{AP} + \frac{TQ}{BQ} + \frac{TR}{CR} = 1.$$

2. Let $\triangle ABC$ be a triangle, and let D, E, F be points on sides $\overline{BC}, \overline{CA}, \overline{AB}$, respectively, such that the cevians $\overline{AD}, \overline{BE}, \overline{CF}$ are concurrent. Show that if M, N, P are points on $\overline{EF}, \overline{FD}, \overline{DE}$, respectively, then the lines $\overleftrightarrow{AM}, \overleftrightarrow{BN}, \overleftrightarrow{CP}$ concur if and only if the lines $\overleftrightarrow{DM}, \overleftrightarrow{EN}, \overleftrightarrow{FP}$ concur. (Many special cases of this question occur in the problem literature.)

3. (Hungary-Israel, 1997) The three squares $ACC_1A'', ABB_1'A', BCDE$ are constructed externally on the sides of a triangle $\triangle ABC$. Let P be the center of $BCDE$. Prove that the lines $\overleftrightarrow{A'C}, \overleftrightarrow{A''B}, \overleftrightarrow{PA}$ are concurrent.

4. (USAMO 1995/3) Given a nonisosceles, nonright triangle $\triangle ABC$ inscribed in a circle with center O, let A_1, B_1, C_1 be the midpoints of sides $\overline{BC}, \overline{CA}, \overline{AB}$ respectively. Point A_2 is located on the ray $\overrightarrow{OA_1}$ so that $\triangle OAA_1$ is similar to $\triangle OA_2A$. Points B_2, C_2 on rays $\overrightarrow{OB_1}, \overrightarrow{OC_1}$ respectively, are defined similarly. Prove that the lines $\overleftrightarrow{AA_2}, \overleftrightarrow{BB_2}, \overleftrightarrow{CC_2}$ are concurrent.

5. Given a triangle $\triangle ABC$ and points X, Y, Z such that $\angle ABZ = \angle XBC$, $\angle BCX = \angle YCA$, $\angle CAY = \angle ZAB$, prove that the lines $\overleftrightarrow{AX}, \overleftrightarrow{BY}, \overleftrightarrow{CZ}$ are concurrent. (Again, many special cases of this problem can be found in the literature.)

6. Let A, B, C, D, E, F, P be seven points on a circle. Show that the lines $\overleftrightarrow{AD}, \overleftrightarrow{BE}, \overleftrightarrow{CF}$ are concurrent if and only if

$$\frac{\sin \angle APB \sin \angle CPD \sin \angle EPF}{\sin \angle BPC \sin \angle DPE \sin \angle FPA} = -1,$$

where the angles are measured modulo 2π. (The only tricky part is the sign.)

5.2 Collinear points: Menelaus's theorem

When he published his theorem, Ceva also revived interest in an ancient theorem attributed to Menelaus[2] (70?-130?).

Figure 5.2.1: Menelaus's theorem (Theorem 5.2.1).

Theorem 5.2.1 (Menelaus)**.** *Let $\triangle ABC$ be a triangle, and let P, Q, R be points on the lines $\overleftrightarrow{BC}, \overleftrightarrow{CA}, \overleftrightarrow{AB}$, respectively, none equal to any of A, B, C. Then P, Q, R are collinear if and only if*

$$\frac{BP}{PC}\frac{CQ}{QA}\frac{AR}{RB} = -1$$

as an equality of signed ratios of lengths.

Figure 5.2.2: Menelaus's theorem (Theorem 5.2.1).

Proof. Assume that P, Q, R are collinear. Let x, y, z be the directed distances from A, B, C, respectively, to the line $\overleftrightarrow{PQR}$. Then $BP/PC = -y/z$ and so forth, so

$$\frac{BP}{PC}\frac{CQ}{QA}\frac{AR}{RB} = (-1)(-1)(-1)\frac{y}{z}\frac{z}{x}\frac{x}{y} = -1.$$

The converse follows by the same argument as for Ceva's theorem. □

An important consequence of Menelaus's theorem is the following result of Desargues (for more on whom see the introduction to Chapter 11), which is most easily stated by introducing two pieces of terminology. Two triangles $\triangle ABC$ and $\triangle DEF$ are said to be *perspective from a point* if the lines $\overleftrightarrow{AD}, \overleftrightarrow{BE}, \overleftrightarrow{CF}$ are concurrent or parallel. The triangles are said to be *perspective from a line* if the points $\overleftrightarrow{AB} \cap \overleftrightarrow{DE}, \overleftrightarrow{BC} \cap \overleftrightarrow{EF}, \overleftrightarrow{CA} \cap \overleftrightarrow{FD}$ are collinear.

Theorem 5.2.2 (Desargues)**.** *Two triangles $\triangle ABC$ and $\triangle DEF$ are perspective from a point if and only if they are perspective from a line.*

Proof. We only prove that if $\triangle ABC$ and $\triangle DEF$ are perpective from a point, then they are perspective from a line. We leave it as an exercise to deduce the reverse implication from this (stare at the diagram); we will do this again later, using duality.

Figure 5.2.3: Proof of Desargues's theorem (Theorem 5.2.2).

Suppose that $\overleftrightarrow{AD}, \overleftrightarrow{BE}, \overleftrightarrow{CF}$ concur at O, and put $P = \overleftrightarrow{BC} \cap \overleftrightarrow{EF}$, $Q = \overleftrightarrow{CA} \cap \overleftrightarrow{FD}$, $R = \overleftrightarrow{AB} \cap \overleftrightarrow{DE}$. To show that P, Q, R are collinear, we want to show that

$$\frac{AR}{RB}\frac{BP}{PC}\frac{CQ}{QA} = -1$$

and then invoke Menelaus's theorem. To get hold of the first term, we apply Menelaus to the points R, D, E on the sides of the triangle $\triangle OAB$, giving

$$\frac{AR}{RB}\frac{BD}{DO}\frac{OE}{EA} = -1.$$

[2]Not to be confused with the brother of Agamemnon in Homer's *Iliad*.

Analogously,

$$\frac{BP}{PC}\frac{CE}{EO}\frac{OF}{FB} = \frac{CQ}{QA}\frac{AF}{FO}\frac{OD}{DC} = -1.$$

When we multiply these three expressions together and cancel equal terms, we get precisely the condition of Menelaus's theorem. $\square$

Another important consequence of Menelaus's theorem is the following result of Pappus of Alexandria (290?-350?).

Theorem 5.2.3 (Pappus)**.** *Let A, C, E be three collinear points, and let B, D, F be three other collinear points. Then the points $\overleftrightarrow{AB} \cap \overleftrightarrow{DE}$, $\overleftrightarrow{BC} \cap \overleftrightarrow{EF}$, $\overleftrightarrow{CD} \cap \overleftrightarrow{FA}$ are collinear.*

The proof is similar, but more complicated; we omit it, save to say that one applies Menelaus repeatedly using the triangle formed by the lines $\overleftrightarrow{AB}, \overleftrightarrow{CD}, \overleftrightarrow{EF}$. If you can't make the cancellation work, see [5].

Note that Desargues's and Pappus's theorems only involve points and lines, with no mention of distances or angles. This makes them "theorems of projective geometry," and we will see later (Section 11.2) how projective transformations often yield simple proofs of such theorems.

Problems for Section 5.2

1. Prove Pappus's theorem (Theorem 5.2.3) directly in terms of Cartesian coordinates; the hope is that you will find this doable but not pleasant!

2. Prove the reverse implication of Desargues' theorem.

3. Let A, B, C be three points on a line. Pick a point D in the plane, and a point E on $\overleftrightarrow{BD}$. Then draw the line through $\overleftrightarrow{AE} \cap \overleftrightarrow{CD}$ and $\overleftrightarrow{CE} \cap \overleftrightarrow{AD}$. Show that this line meets the line $\overleftrightarrow{AC}$ in a point P that depends only on A, B, C. (The points A, B, C, P are in fact harmonic conjugates, for more on which see Section 11.6.)

4. (MOP 1990) Let A, B, C be three collinear points and D, E, F three other collinear points. Put $G = \overleftrightarrow{BE} \cap \overleftrightarrow{CF}$, $H = \overleftrightarrow{AD} \cap \overleftrightarrow{CF}$, $I = \overleftrightarrow{AD} \cap \overleftrightarrow{CE}$. If $AI = HD$ and $CH = GF$, prove that $BI = GE$.

Figure 5.3.1: Fact 5.3.1.

5. (Original) Let $\triangle ABC$ be a triangle and let P be a point in its interior, not lying on any of the medians of $\triangle ABC$. Put $A_1 = \overleftrightarrow{PA} \cap \overline{BC}$, $B_1 = \overleftrightarrow{PB} \cap \overline{CA}$, $C_1 = \overleftrightarrow{CA} \cap \overline{AB}$, $A_2 = \overleftrightarrow{B_1C_1} \cap \overleftrightarrow{BC}$, $B_2 = \overleftrightarrow{C_1A_1} \cap \overleftrightarrow{CA}$, $C_2 = \overleftrightarrow{A_2B_2} \cap \overleftrightarrow{AB}$. Prove that the midpoints of $\overline{A_1A_2}, \overline{B_1B_2}, \overline{C_1C_2}$ are collinear. (See also Problem 6.5.2.)

6. (Aaron Pixton) Let $\triangle ABC$ and $\triangle DEF$ be triangles, and let P be a point. For each nonzero real number r, let T_r be the triangle obtained from $\triangle ABC$ by a homothety about P of ratio r. Suppose that for some three distinct nonzero real numbers r_1, r_2, r_3, each of $T_{r_1}, T_{r_2}, T_{r_3}$ is perspective with $\triangle DEF$. Prove that T_r and $\triangle DEF$ are perspective for any r.

5.3 Concurrent perpendiculars

Some of the special points of a triangle are constructed by drawing perpendiculars to the sides of a triangle. For example, the circumcenter can be constructed by drawing the perpendicular bisectors. It is convenient that a result analogous to Ceva's Theorem holds for perpendiculars; the analogy is so strong that we can safely leave the proof to the reader (see Problem 1).

Fact 5.3.1. *Let $\triangle ABC$ be a triangle, and let P, Q, R be three points in the plane. Then the lines through P, Q, R perpendicular to $\overleftrightarrow{BC}$, $\overleftrightarrow{CA}$, $\overleftrightarrow{AB}$, respectively, are concurrent or parallel if and only if*

$$BP^2 - PC^2 + CQ^2 - QA^2 + AR^2 - RB^2 = 0.$$

A surprising consequence is that the lines through P, Q, R perpendicular to $\overleftrightarrow{BC}$, $\overleftrightarrow{CA}$, $\overleftrightarrow{AB}$, respectively, are concurrent or parallel if and only if the lines through A, B, C perpendicular to $\overleftrightarrow{QR}$, $\overleftrightarrow{RP}$, $\overleftrightarrow{PQ}$, respectively, are concurrent or parallel!

Problems for Section 5.3

1. Prove that the lines $\overleftrightarrow{AB}$ and $\overleftrightarrow{CD}$ are perpendicular if and only if $AC^2 - AD^2 = BC^2 - BD^2$. (Use vectors, coordinates or Pythagoras.) Then prove Fact 5.3.1.

2. (Germany, 1996) Let $\triangle ABC$ be a triangle, and construct squares ABB_1A_2, BCC_1B_2, CAA_1C_2 externally on its sides. Prove that the perpendicular bisectors of the segments $\overline{A_1A_2}$, $\overline{B_1B_2}$, $\overline{C_1C_2}$ are concurrent.

3. Let $\triangle ABC$ be a triangle, ℓ a line and L, M, N the feet of the perpendiculars to ℓ from A, B, C, respectively. Prove that the perpendiculars to $\overleftrightarrow{BC}$, $\overleftrightarrow{CA}$, $\overleftrightarrow{AB}$ through L, M, N, respectively, are concurrent. (Their intersection is called the *orthopole* of the line ℓ and the triangle ABC.)

4. (USAMO 1997/2) Let $\triangle ABC$ be a triangle, and draw isosceles triangles $\triangle DBC$, $\triangle AEC$, $\triangle ABF$ external to $\triangle ABC$ (with $\overline{BC}, \overline{CA}, \overline{AB}$ as their respective bases). Prove that the lines through A, B, C perpendicular to $\overleftrightarrow{EF}$, $\overleftrightarrow{FD}$, $\overleftrightarrow{DE}$, respectively, are concurrent. (Several solutions are possible.)

Chapter 6

Circular reasoning

This chapter is of course devoted not to logical fallacies, but to reasoning about the most fundamental of geometric objects, the circle. Note that we will gain further insight into the geometry of circles after introducing inversion in Chapter 10.

6.1 Power of a point

The following is a theorem of Euclidean geometry in the strictest of senses: it appears in the *Elements* as Propositions III.35–III.37.

Theorem 6.1.1. *Given a fixed circle ω and a fixed point P, draw a line through P intersecting ω at A and B. Then the product $PA \cdot PB$ depends only on P and ω, not on the line.*

Proof. Draw another line through P meeting ω at C and D, labeled as in one of the diagrams. Then

Figure 6.1.1: Proof of the power of a point theorem (Theorem 6.1.1).

$$\measuredangle PAC = \measuredangle BAC = \measuredangle BDC = -\measuredangle PDB$$

as directed angles, so the triangles $\triangle PAC$ and $\triangle PDB$ are (oppositely) similar, giving $PA/PD = PC/PB$, or equivalently $PA \cdot PB = PC \cdot PB$. $\square$

We may view $PA \cdot PB$ as a signed quantity by the same convention as for signed ratios of lengths. This signed quantity is called the *power* of P with respect to ω; note that it is positive if P lies outside ω, zero if P lies on ω, and negative if P lies inside ω. If O is the center of ω and r is the radius, we may choose $\overleftrightarrow{OP}$ as our line and so express the power as

$$(OP + r)(OP - r) = OP^2 - r^2.$$

Note that for P outside ω, the limiting case $A = B$ means that $\overleftrightarrow{PA}$ is tangent to ω at A.

The power of a point theorem has an occasionally useful converse.

Fact 6.1.2. *If the lines $\overleftrightarrow{AB}$ and $\overleftrightarrow{CD}$ meet at P, and there is an equality $PA \cdot PB = PC \cdot PD$ of signed products of lengths, then A, B, C, D are concyclic.*

Problems for Section 6.1

1. If A, B, C, D are concyclic and $\overleftrightarrow{AB} \cap \overleftrightarrow{CD} = E$, prove that $(AC/BC)(AD/BD) = AE/BE$.

2. (*Mathematics Magazine*, Dec. 1992) Let $\triangle ABC$ be an acute triangle, let H be the foot of the altitude from A, and let D, E, Q be the feet of the perpendiculars from an arbitrary point P in the triangle onto $\overline{AB}, \overline{AC}, \overline{AH}$, respectively. Prove that

$$|AB \cdot AD - AC \cdot AE| = BC \cdot PQ.$$

3. Draw tangents $\overline{OA}$ and $\overline{OB}$ from a point O to a given circle. Through A is drawn a chord $\overline{AC}$ parallel to $\overleftrightarrow{OB}$; let E be the second intersection of $\overleftrightarrow{OC}$ with the circle. Prove that the line $\overleftrightarrow{AE}$ bisects the segment $\overline{OB}$.

4. (MOP 1995) Given triangle $\triangle ABC$, let D, E be any points on $\overline{BC}$. A circle through A cuts the lines $\overleftrightarrow{AB}, \overleftrightarrow{AC}, \overleftrightarrow{AD}, \overleftrightarrow{AE}$ at the points P, Q, R, S, respectively. Prove that

$$\frac{AP \cdot AB - AR \cdot AD}{AS \cdot AE - AQ \cdot AC} = \frac{BD}{CE}.$$

5. (IMO 1995/1) Let A, B, C, D be four distinct points on a line, in that order. The circles with diameters $\overline{AC}$ and $\overline{BD}$ intersect at X and Y. The line $\overleftrightarrow{XY}$ meets $\overleftrightarrow{BC}$ at Z. Let P be a point on the line $\overleftrightarrow{XY}$ other than Z. The line $\overleftrightarrow{CP}$ intersects the circle with diameter $\overline{AC}$ at C and M, and the line $\overleftrightarrow{BP}$ intersects the circle with diameter $\overline{BD}$ at B and N. Prove that the lines $\overleftrightarrow{AM}, \overleftrightarrow{DN}, \overleftrightarrow{XY}$ are concurrent.

6. (USAMO 1998/2) Let ω_1 and ω_2 be concentric circles, with ω_2 in the interior of ω_1. From a point A on ω_1 one draws the tangent $\overline{AB}$ to ω_2 ($B \in \omega_2$). Let C be the second point of intersection of $\overleftrightarrow{AB}$ and ω_1, and let D be the midpoint of $\overline{AB}$. A line passing through A intersects ω_2 at E and F in such a way that the perpendicular bisectors of $\overline{DE}$ and $\overline{CF}$ intersect at a point M on $\overline{AB}$. Find, with proof, the ratio AM/MC.

6.2 Radical axis

Given two circles, one with center O_1 and radius r_1, the other with center O_2 and radius r_2, what is the set of points with equal power with respect to the two circles? By our explicit formula for the power of a point, this is simply the set of points P such that $PO_1^2 - r_1^2 = PO_2^2 - r_2^2$, or equivalently such that $PO_1^2 - PO_2^2 = r_1^2 - r_2^2$. By Problem 5.3.1, this set is a straight line perpendicular to $\overleftrightarrow{O_1O_2}$; we call this line the *radical axis* of the two circles.

Theorem 6.2.1 (Radical axis theorem). *Let $\omega_1, \omega_2, \omega_3$ be three circles. Then the radical axes of ω_1 and ω_2, of ω_2 and ω_3, and of ω_3 and ω_1 either all coincide, or are concurrent (or parallel).*

Proof. A point on two of the radical axes has equal power with respect to all three circles. Hence if two of the axes coincide, so does the third, and otherwise if any two of the axes have a common point, this point lies on the third axis as well. $\square$

Corollary 6.2.2. *The common chords of three mutually intersecting circles lie on concurrent lines.*

If the radical axes coincide, the three circles are said to be *coaxial*[1] ; otherwise, the intersection of the three radical axes is called the *radical center*

[1] The word is also spelled "coaxal", as in [5].

of the circles. (As usual, this intersection could be "at infinity", if the three lines are parallel.) There are three types of coaxial families, depending on whether the circles have zero, one, or two intersections with the common radical axis; these three cases are illustrated in Figure 6.2.1. (Note: the zero and two cases become identical in the complex projective plane; see Section 11.7.) A useful criterion for recognizing and applying the coaxial

Figure 6.2.1: Some coaxial families of circles.

property is the following simple observation and partial converse.

Fact 6.2.3. *If three circles are coaxial, their centers are collinear. Conversely, if three circles pass through a common point and have collinear centers, they are coaxial.*

Like the power-of-a-point theorem, the radical axis theorem has an occasionally useful converse.

Fact 6.2.4. *Suppose that $ABCD$ and $CDEF$ are cyclic quadrilaterals, and the lines $\overleftrightarrow{AB}, \overleftrightarrow{CD}, \overleftrightarrow{EF}$ are concurrent. Then $EFAB$ is also cyclic. More generally, if ω_1, ω_2 are two circles with radical axis ℓ, A, B are points on ω_1, C, D are points on ω_2, and $\overleftrightarrow{AB}$ and $\overleftrightarrow{EF}$ meet at a point on ℓ, then A, B, E, F are concyclic. (We may allow $A = B$ by taking $\overleftrightarrow{AB}$ to be the tangent line to ω_1 at that point, and likewise we may allow $C = D$.)*

Problems for Section 6.2

1. Prove Fact 6.2.4. (Hint: draw a third circle and apply the radical axis theorem.)

2. Use the radical axis theorem to give another solution for Problem 5.3.4.

3. (MOP 1995) Let $\overline{BB'}, \overline{CC'}$ be altitudes of triangle $\triangle ABC$, and assume $AB \neq AC$. Let M be the midpoint of $\overline{BC}$, H the orthocenter of $\triangle ABC$, and D the intersection of $\overleftrightarrow{BC}$ and $\overleftrightarrow{B'C'}$. Show that $\overleftrightarrow{DH}$ is perpendicular to $\overleftrightarrow{AM}$.

4. (IMO 1994 proposal) A circle ω is tangent to two parallel lines ℓ_1 and ℓ_2. A second circle ω_1 is tangent to ℓ_1 at A and to ω externally at C. A third circle ω_2 is tangent to ℓ_2 at B, to ω externally at D and to ω_1 externally at E. Let Q be the intersection of $\overleftrightarrow{AD}$ and $\overleftrightarrow{BC}$. Prove that $QC = QD = QE$.

5. (India, 1995) Let $\triangle ABC$ be a triangle. A line parallel to $\overleftrightarrow{BC}$ meets sides $\overline{AB}$ and $\overline{AC}$ at D and E, respectively. Let P be a point inside triangle $\triangle ADE$, and let F and G be the intersection of $\overleftrightarrow{DE}$ with $\overleftrightarrow{BP}$ and $\overleftrightarrow{CP}$, respectively. Show that A lies on the radical axis of the circumcircles of $\triangle PDG$ and $\triangle PFE$.

6. (IMO 1985/5) A circle with center O passes through the vertices A and C of triangle $\triangle ABC$, and intersects the segments $\overline{AB}$ and $\overline{BC}$ again at distinct points K and N, respectively. The circumscribed circles of the triangle $\triangle ABC$ and $\triangle KBN$ intersect at exactly two distinct points B and M. Prove that angle $\angle OMB$ is a right angle.

7. (TST 2004/4) Let $\triangle ABC$ be a triangle, and let D be a point in its interior. Construct a circle ω_1 passing through B and D, and a circle ω_2 passing through C and D, such that the point of intersection of ω_1 and ω_2 other than D lies on the line $\overleftrightarrow{AD}$. Denote by E, F the points where ω_1, ω_2 intersect side $\overline{BC}$, respectively, and by X, Y the intersections $\overleftrightarrow{DF} \cap \overleftrightarrow{AB}, \overleftrightarrow{DE} \cap \overleftrightarrow{AC}$, respectively. Prove that $\overleftrightarrow{XY}$ is parallel to $\overleftrightarrow{BC}$.

6.3 The Pascal-Brianchon theorems

Although Blaise Pascal (1623–1662) is most famous for "Pascal's triangle[2]", he also left behind an amazing theorem about hexagons inscribed in circles. His original proof, which was favorably described by calculus pioneer Gottfried Wilhelm von Leibniz (1646–1716), has unfortunately been lost; we present here an ingenious proof essentially due to Jan van Yzeren (A simple proof of Pascal's hexagon theorem, *Monthly*, December 1993).

Theorem 6.3.1 (Pascal)**.** *Let $ABCDEF$ be a hexagon inscribed in a circle. Then the intersections $\overleftrightarrow{AB} \cap \overleftrightarrow{DE}$, $\overleftrightarrow{BC} \cap \overleftrightarrow{EF}$, $\overleftrightarrow{CD} \cap \overleftrightarrow{FA}$ are collinear.*

[2]The famous triangle was actually known in ancient China. However, Pascal investigated the triangle much more deeply, in his foundational work on probability theory.

Figure 6.3.1: Pascal's theorem (Theorem 6.3.1) and van Yzeren's proof.

Proof. Put $P = \overleftrightarrow{AB} \cap \overleftrightarrow{DE}$, $Q = \overleftrightarrow{BC} \cap \overleftrightarrow{EF}$, $R = \overleftrightarrow{CD} \cap \overleftrightarrow{FA}$. Draw the circle ω through C, F, R, and extend the lines $\overleftrightarrow{BC}$ and $\overleftrightarrow{EF}$ to meet this circle again at G and H, respectively; see Figure 6.3.1. By Theorem 4.2.2, we have $\overleftrightarrow{BE} \parallel \overleftrightarrow{GH}$, $\overleftrightarrow{ED} \parallel \overleftrightarrow{HR}$, $\overleftrightarrow{AB} \parallel \overleftrightarrow{RG}$.

Now notice that the triangles $\triangle RGH$ and $\triangle PBE$ have parallel sides, which means that they are homothetic. In other words, the lines $\overleftrightarrow{BG}, \overleftrightarrow{EH}, \overleftrightarrow{PR}$ are concurrent, which means $\overleftrightarrow{BG} \cap \overleftrightarrow{EH} = Q$ is collinear with P and R, as desired. □

Some time later, Charles Brianchon (1783–1864) discovered a counterpart to Pascal's theorem for a hexagon circumscribed about a circle. We will give Brianchon's proof of his theorem, which uses the polar map to reduce it to Pascal's theorem, in Section 11.5; a direct but somewhat complicated proof can be found in [5].

Theorem 6.3.2 (Brianchon)**.** *Let $ABCDEF$ be a hexagon circumscribed about a circle (i.e., the extension of each side is tangent to the circle). Then the lines $\overleftrightarrow{AD}, \overleftrightarrow{BE}, \overleftrightarrow{CF}$ are concurrent.*

Figure 6.3.2: Brianchon's theorem (Theorem 6.3.2).

It is sometimes useful to apply Pascal's theorem or Brianchon's theorem in certain degenerate cases, in which some of the vertices coincide. For example, in Pascal's theorem, if two adjacent vertices of the hexagon coincide, one should take the line through them to be the tangent to the circle at that point. Thus in Figure 6.3.3, we may conclude that $\overleftrightarrow{AA} \cap \overleftrightarrow{DE}$, $\overleftrightarrow{AC} \cap \overleftrightarrow{EF}$,

Figure 6.3.3: A degenerate case of Pascal's theorem (Theorem 6.3.1).

$\overleftrightarrow{CD} \cap \overleftrightarrow{FA}$ are collinear, where $\overleftrightarrow{AA}$ denotes the tangent to the circle at A.

As for Brianchon's theorem, the analogous argument shows that the "vertex" between two collinear sides belongs at the point of tangency, as in Figure 6.3.4.

Problems for Section 6.3

Figure 6.3.4: A degenerate case of Brianchon's theorem (Theorem 6.3.2).

1. What do we get if we apply Brianchon's theorem with three degenerate vertices? (We will encounter this fact again later.)
2. Let $ABCD$ be a quadrilateral whose sides $\overline{AB}$, $\overline{BC}$, $\overline{CD}$, $\overline{DA}$ are tangent to a single circle at points M, N, P, Q, respectively. Prove that the lines AC, BD, MP, NQ are concurrent.
3. (MOP 1995) With notation as in the previous problem, let lines $\overleftrightarrow{BQ}$, $\overleftrightarrow{BP}$ intersect the circle at E, F, respectively. Prove that $\overleftrightarrow{ME}$, $\overleftrightarrow{NF}$, $\overleftrightarrow{BD}$ are concurrent.
4. (Poland, 1997) Let $ABCDE$ be a convex pentagon with $CD = DE$ and $\angle BCD = \angle DEA = \pi/2$. Let F be the point on side $\overline{AB}$ such that $AF/FB = AE/BC$. Show that
$$\angle FCE = \angle FDE \quad \text{and} \quad \angle FEC = \angle BDC.$$

6.4 Simson lines

The following theorem is often called Simson's theorem in honor of Robert Simson (1687–1768), but it is actually originally due to William Wallace (1768–1843).

Theorem 6.4.1. *Let A, B, C be three points on a circle. Then the feet of the perpendiculars from P to the lines $\overleftrightarrow{AB}, \overleftrightarrow{BC}, \overleftrightarrow{CA}$ are collinear if and only if P also lies on the circle.*

Proof. The proof is by (directed) angle-chasing. Let X, Y, Z be the feet of the respective perpendiculars from P to $\overleftrightarrow{BC}, \overleftrightarrow{CA}, \overleftrightarrow{AB}$; then the quadrilaterals $PXCY, PYAZ, PZBX$ each have two right angles, and are thus cyclic. Therefore

$$\begin{aligned} \measuredangle PXY &= \measuredangle PCY && \text{(cyclic quadrilateral } PXCY) \\ &= \measuredangle PCA && \text{(collinearity of } A, C, Y) \end{aligned}$$

and analogously $\measuredangle PXZ = \measuredangle PBA$. Now X, Y, Z are collinear if and only if $\measuredangle PXY = \measuredangle PXZ$, which by the above equations occurs if and only if $\measuredangle PCA = \measuredangle PBA$; in other words, if and only if A, B, C, P are concyclic. $\square$

For P on the circle, the line described in the theorem is called the *Simson line* of P with respect to the triangle $\triangle ABC$. We note in passing that an alternate proof of the collinearity in this case can be given using Menelaus's theorem.

Problems for Section 6.4

1. Let A, B, C, P, Q be points on a circle. Show that the (directed) angle between the Simson lines of P and Q with respect to the triangle $\triangle ABC$ equals half of the (directed) arc measure $m(\widehat{PQ})$.

2. Let A, B, C, D be four points on a circle. Prove that the intersection of the Simson line of A with respect to $\triangle BCD$ with the Simson line of B with respect to $\triangle ACD$ is collinear with C and the orthocenter of $\triangle ABD$.

3. If A, B, C, P, Q are five points on a circle such that $\overline{PQ}$ is a diameter, show that the Simson lines of P and Q with respect to $\triangle ABC$ intersect at a point concyclic with the midpoints of the sides of $\triangle ABC$.

4. Let I be the incenter of triangle $\triangle ABC$, and D, E, F the projections of I onto $\overline{BC}, \overline{CA}, \overline{AB}$, respectively. The incircle of ABC meets the segments $\overline{AI}, \overline{BI}, \overline{CI}$ at M, N, P, respectively. Show that the Simson lines of any point on the incircle with respect to the triangles $\triangle DEF$ and $\triangle MNP$ are perpendicular.

6.5 Circle of Apollonius

The ancient geometer Apollonius of Perga (262?–190? B.C.E.) is most famous for his early work on conic sections (see Section 11.3). However, his name has also come to be attached to another pretty geometrical construction.

Theorem 6.5.1. *Let A, B be any two points, and let $k \neq 1$ be a positive real number. Then the locus of points P such that $PA/PB = k$ is a circle whose center lies on $\overleftrightarrow{AB}$.*

Proof. One can show this synthetically, but the shortest proof involves introducing Cartesian coordinates such that $A = (a, 0)$ and $B = (b, 0)$. The condition $PA/PB = k$ is equivalent to $PA^2 = k^2 PB^2$, which in coordinates can be written

$$(x-a)^2 + y^2 = k^2[(x-b)^2 + y^2].$$

Combining terms and dividing through by $1 - k^2$, we get

$$x^2 + \frac{2k^2 b - 2a}{1-k^2} x + y^2 = \frac{k^2 b^2 - a^2}{1-k^2},$$

which is easily recognized as the equation of a circle whose center lies on the x-axis. □

This circle is called the *circle of Apollonius* corresponding to the points A, B and the ratio k. (This term usually also includes the degenerate case $k = 1$, where the "circle" becomes the perpendicular bisector of $\overline{AB}$.)

Problems for Section 6.5

1. Use circles of Apollonius to give a synthetic proof of the classification of similarities (Theorem 3.4.1).

2. (Original) Set notation as in Problem 5.2.5. Prove that if some two of the circles with diameters $\overline{A_1A_2}, \overline{B_1B_2}, \overline{C_1C_2}$ intersect, then the three circles are coaxial (and so Problem 5.2.5 follows). Beware that the case where the circles do not meet is trickier, unless you work in the complex projective plane as described in Section 11.7.

6.6 Additional problems

Problems for Section 6.6

1. (Archimedes' "broken-chord" theorem) Point D is the midpoint of arc $\widehat{AC}$ of a circle; point B is on minor arc $\widehat{CD}$; and E is the point on $\overline{AB}$ such that $\overleftrightarrow{DE}$ is perpendicular to $\overleftrightarrow{AB}$. Prove that $AE = BE + BC$.

2. The convex hexagon $ABCDEF$ is such that

$$\angle BCA = \angle DEC = \angle FAE = \angle AFB = \angle CBD = \angle EDF.$$

Prove that $AB = CD = EF$.

3. (Descartes's four circles theorem) Let r_1, r_2, r_3, r_4 be the radii of four mutually externally tangent circles. Prove that

$$\sum_{i=1}^{4} \frac{2}{r_1^2} = \left(\sum_{i=1}^{4} \frac{1}{r_i} \right)^2 .$$

Also verify that the result holds without requiring the tangencies to be external, if one imposes the sign convention that two radii have the same sign if they correspond to externally tangent circles, and have opposite sign otherwise.

4. Deduce from the previous problem the following interesting consequence. Define the *curvature* of a circle to be the reciprocal of its radius. Draw three mutually externally tangent circles with integer curvatures, each internally tangent to a given unit circle. Then repeatedly insert the circle externally tangent to three previously drawn mutually externally tangent circles. Show that all of the resulting circles have integer curvature. These form an example of an *Apollonian gasket* (or *Apollonian circle packing*) .

Figure 6.6.1: An Apollonian gasket.

Chapter 7

Triangle trivia

In this chapter, we study but a few of the most important constructions associated to a triangle. One could pursue this study almost indefinitely; simply restricting to "centers" associated to a triangle leads to literally hundreds[1] of examples.

For convenience, we adopt the following convention throughout this chapter: in triangle $\triangle ABC$, we write a, b, c for the respectively side lengths BC, CA, AB.

7.1 Centroid

For $\triangle ABC$ a triangle, the *median* of $\triangle ABC$ from vertex A is the cevian from A whose endpoint is the midpoint of $\overline{BC}$.

Fact 7.1.1. *The medians of a triangle are concurrent. Moreover, the point of concurrency trisects each median.*

Figure 7.1.1: Fact 7.1.1.

One can easily show this using Ceva and Menelaus, or by performing an affine transformation making the triangle equilateral, or by using vectors.

[1] Indeed, the companion web site `http://faculty.evansville.edu/ck6/encyclopedia/` of the book [13] lists over 1000 special points associated to a triangle!

The concurrency point of the medians is called the *centroid*[2] of $\triangle ABC$. It is also called the *center of mass* for the following reason: if equal masses are placed at each of A, B, C, the center of mass of the resulting system will lie at the centroid of $\triangle ABC$. (Compare the discussion of "mass points" in Section 2.2.)

Problems for Section 7.1

1. (`http://www.cut-the-knot.org`) Let G be the centroid of triangle $\triangle ABC$. Draw a line through G meeting $\overleftrightarrow{AB}$ at M and $\overleftrightarrow{CA}$ at N. Prove that as ratios of signed lengths,
$$\frac{BM}{MA} + \frac{CN}{NA} = 1.$$

2. (Floor van Loemen, *Monthly* April 2002) A triangle is divided into six smaller triangles by its medians. Prove that the circumcenters of these six triangles lie on a circle.

7.2 Incenter and excenters

If the point P lies in the interior of triangle $\triangle ABC$, then then the distances from P to the lines $\overleftrightarrow{AB}$ and $\overleftrightarrow{AC}$ are

$$PA \sin \angle PAB \quad \text{and} \quad PA \sin \angle PAC$$

and these are equal if and only if $\angle PAB = \angle PAC$, in other words, if P lies on the internal angle bisector of $\angle A$.

It follows that the intersection of two internal angle bisectors is equidistant from all three sides, and consequently lies on the third bisector. This intersection is the *incenter* of $\triangle ABC$, and its distance to any side is the *inradius*, usually denoted r. The terminology comes from the fact that the circle of radius r centered at the incenter is tangent to all three sides of $\triangle ABC$, and thus is called the *inscribed circle*, or *incircle*, of ABC.

Do not forget, though, that the angle $\angle A$ in triangle $\triangle ABC$ has *two* angle bisectors, one internal and one external. The locus of points equidistant to

[2]The existence of the centroid seems to be one of the few nontrivial facts proved in standard American geometry courses.

the two lines $\overleftrightarrow{AB}$ and $\overleftrightarrow{AC}$ is the union of both lines, and so one might expect to find other circles tangent to all three sides. Indeed, the internal angle bisector at A concurs with the external bisectors of the other two angles (by the same argument as above); the point of concurrence is the *excenter* of $\triangle ABC$ opposite A, and the circle centered there tangent to all three sides is the *escribed (exscribed) circle*, or *excircle*, of $\triangle ABC$ opposite A.

In studying the incircle and excircles, a fundamental tool is the fact that the two tangents to a circle from an external point have the same length. This fact is equally useful is its own right, and we have included some problems that take advantage of equal tangents. In any case, the key observation we need is that if D, E, F are the points where the incircle touches $\overline{BC}, \overline{CA}, \overline{AB}$, respectively, then $AE = AF$ and so on, so a little algebra gives

$$AE = \frac{1}{2}(AE + EC + AF + FB - CD - DB).$$

This establishes the first half of the following result; the second half is no harder. Recall that $s = (a + b + c)/2$ is called the *semiperimeter* of $\triangle ABC$.

Fact 7.2.1. *Let $s = (a + b + c)/2$. Then the distance from A to the point where the incircle touches $\overline{AB}$ is $s - a$, and the distance from A to the point where the excircle opposite C touches $\overline{AB}$ is $s - b$.*

Figure 7.2.1: Fact 7.2.1.

It will often be helpful to know in what ratio an angle bisector divides the opposite side. The answer can be used to give another proof of the concurrence of the angle bisectors.

Fact 7.2.2 (Angle bisector theorem)**.** *If D is the foot of either angle bisector of A in triangle $\triangle ABC$, then (as unsigned lengths)*

$$\frac{DB}{DC} = \frac{AB}{AC}.$$

Another useful construction for studying incenters is the following.

Fact 7.2.3. *Let $\triangle ABC$ be a triangle inscribed in a circle ω with center O, and let M be the second intersection of ω with the internal angle bisector of A.*

Figure 7.2.2: Fact 7.2.3.

1. *The line $\overleftrightarrow{MO}$ is perpendicular to $\overleftrightarrow{BC}$, i.e., M is the midpoint of arc $\widehat{BC}$.*
2. *The circle centered at M passing through B and C also passes through the incenter I and the excenter I_A opposite A; that is, $MB = MI = MC = MI_A$.*
3. *We have $OI^2 = R^2 - 2Rr$, where R is the circumradius and r is the inradius of $\triangle ABC$.*

Problems for Section 7.2

1. Use the angle bisector theorem to give a synthetic proof of Theorem 6.5.1.
2. (*Arbelos*) The two common external tangent segments between two nonintersecting circles cut off a segment along one of the common internal tangents. Show that all three segments have the same length.
3. The incircle of a triangle is projected onto each of the three sides. Prove that the six endpoints of the resulting segments are concyclic.
4. (Răzvan Gelca) Let $\triangle ABC$ be a triangle, and let D, E, F be the points where the incircle touches the sides $\overline{BC}, \overline{CA}, \overline{AB}$, respectively. Let M, N, P be points on the segments $\overline{EF}, \overline{FD}, \overline{DE}$, respectively. Show that the lines $\overleftrightarrow{AM}, \overleftrightarrow{BN}, \overleftrightarrow{CP}$ intersect if and only if the lines $\overleftrightarrow{DM}, \overleftrightarrow{EN}, \overleftrightarrow{FP}$ intersect.
5. (USAMO 1991/5) Let D be an arbitrary point on side $\overline{AB}$ of a given triangle $\triangle ABC$, and let E be the interior point where $\overline{CD}$ intersects the external common tangent to the incircles of triangles $\triangle ACD$ and $\triangle BCD$. As D assumes all positions between A and B, show that E traces an arc of a circle.
6. (Iran, 1997) Let $\triangle ABC$ be a triangle, and let P a varying point on the arc $\widehat{BC}$ of the circumcircle of $\triangle ABC$. Prove that the circle through P and the incenters of $\triangle PAB$ and $\triangle PAC$ passes through a fixed point independent of P.

7. (USAMO 1999/6) Let $ABCD$ be an isosceles trapezoid with $\overleftrightarrow{AB} \parallel \overleftrightarrow{CD}$. The inscribed circle ω of triangle $\triangle BCD$ meets $\overline{CD}$ at E. Let F be a point on the (internal) angle bisector of $\angle DAC$ such that $\overleftrightarrow{EF} \perp \overleftrightarrow{CD}$. Let the circumscribed circle of triangle $\triangle ACF$ meet line $\overleftrightarrow{CD}$ at C and G. Prove that the triangle $\triangle AFG$ is isosceles.

8. (IMO 1992/4) In the plane let C be a circle, let L be a line tangent to the circle C, and let M be a point on L. Find the locus of all points P with the following property: there exists two points Q, R on L such that M is the midpoint of $\overline{QR}$ and C is the inscribed circle of triangle $\triangle PQR$.

9. (Bulgaria, 1996) The circles k_1 and k_2 with respective centers O_1 and O_2 are externally tangent at the point C, while the circle k with center O is externally tangent to k_1 and k_2. Let ℓ be the common tangent of k_1 and k_2 at the point C and let $\overline{AB}$ be the diameter of k perpendicular to ℓ. Assume that O_2 and A lie on the same side of ℓ. Show that the lines $\overleftrightarrow{AO_1}$, $\overleftrightarrow{BO_2}$, ℓ have a common point.

10. (MOP 1997) Let $\triangle ABC$ be a triangle, and D, E, F the points where the incircle touches sides $\overline{BC}, \overline{CA}, \overline{AB}$, respectively. The parallel to $\overleftrightarrow{AB}$ through E meets $\overleftrightarrow{DF}$ at Q, and the parallel to $\overleftrightarrow{AB}$ through D meets $\overleftrightarrow{EF}$ at T. Prove that the lines $\overleftrightarrow{CF}, \overleftrightarrow{DE}, \overleftrightarrow{QT}$ are concurrent.

11. (Stanley Rabinowitz[3]) The incircle of triangle $\triangle ABC$ touches sides $\overline{BC}, \overline{CA}, \overline{AB}$ at D, E, F, respectively. Let P be any point inside triangle $\triangle ABC$, and let X, Y, Z be the points where the segments $\overline{PA}, \overline{PB}, \overline{PC}$, respectively, meet the incircle. Prove that the lines $\overleftrightarrow{DX}, \overleftrightarrow{EY}, \overleftrightarrow{FZ}$ are concurrent.

7.3 Circumcenter and orthocenter

It was pointed out earlier that any triangle $\triangle ABC$ has a unique circumscribing circle (circumcircle); note that its center, the circumcenter of $\triangle ABC$, is the point of concurrency of the perpendicular bisectors of $\overline{AB}$, $\overline{BC}$, $\overline{CA}$.

[3] Rabinowitz uses the diagram for this problem as the logo of his company Mathpro Press.

A closely related point is the *orthocenter*, defined as the intersection of the altitudes of a triangle. One can apply Fact 5.3.1 to show that these actually concur, or one can modify the proof of the following theorem to include this concurrence as a consequence.

Theorem 7.3.1. *Let $\triangle ABC$ be a triangle, and let O, G, H be its circumcenter, centroid and orthocenter, respectively. Then O, G, H lie on a line in that order, and $2OG = GH$.*

The line OGH is called the *Euler line* of triangle $\triangle ABC$.

Proof. The homothety with center G and ratio $-1/2$ carries $\triangle ABC$ to the *medial triangle* $\triangle A'B'C'$, where A', B', C' are the respective midpoints of $\overline{BC}, \overline{CA}, \overline{AB}$. Moreover, the altitude from A' in the medial triangle coincides with the perpendicular bisector of $\overline{BC}$, since both are perpendicular to $\overleftrightarrow{BC}$ and pass through A'. Hence H maps to O under the homothety, and the claim follows. □

Some of the problems will use the following facts about the orthocenter, which we leave as exercises in angle-chasing.

Figure 7.3.1: The orthic triangle of a triangle (Fact 7.3.2).

Fact 7.3.2. *In triangle $\triangle ABC$, let H be the orthocenter, and let H_A, H_B, H_C be the feet of the respective altitudes from A, B, C. Then the following statements hold.*

1. *The triangles $\triangle AH_BH_C$, $\triangle H_ABH_C$, $\triangle H_AH_BC$ are (oppositely) similar to $\triangle ABC$.*

2. *The altitudes bisect the angles of the triangle $\triangle H_AH_BH_C$ (so H is its incenter).*

3. *The reflections of H across $\overleftrightarrow{BC}, \overleftrightarrow{CA}, \overleftrightarrow{AB}$ lie on the circumcircle of $\triangle ABC$.*

The triangle formed by the feet of the altitudes is called the *orthic triangle*.

Fact 7.3.3. *Let $\triangle ABC$ be a triangle with orthocenter H. Define the following points:*

- *let M_A, M_B, M_C be the midpoints of the sides $\overline{BC}, \overline{CA}, \overline{AB}$;*
- *let H_A, H_B, H_C be the feet of the altitudes from A, B, C;*
- *let A', B', C' be the midpoints of the segments $\overline{HA}, \overline{HB}, \overline{HC}$.*

Then the following statements hold.

(a) The triangle $\triangle A'B'C'$ is the half-turn of $\triangle M_A M_B M_C$ about its circumcenter.

(b) The points $M_A, M_B, M_C, H_A, H_B, H_C, A', B', C'$ lie on a single circle.

(c) The center of the circle in (b) is the midpoint of $\overline{OH}$.

The circle described in Fact 7.3.3 is called the *nine-point circle* of $\triangle ABC$.

Problems for Section 7.3

1. Let $\triangle ABC$ be a triangle. A circle passing through B and C intersects $\overleftrightarrow{AB}$ and $\overleftrightarrow{AC}$ again at C' and B', respectively. Prove that $\overleftrightarrow{BB'}, \overleftrightarrow{CC'}, \overleftrightarrow{HH'}$ are concurrent, where H, H' are the respective orthocenters of $\triangle ABC, \triangle A'B'C'$.

2. (USAMO 1990/5) An acute-angled triangle $\triangle ABC$ is given in the plane. The circle with diameter $\overline{AB}$ intersects altitude $\overline{CC'}$ and its extension at points M and N, and the circle with diameter $\overline{AC}$ intersects altitude $\overline{BB'}$ and its extensions at P and Q. Prove that the points M, N, P, Q lie on a common circle.

3. Let ℓ be a line through the orthocenter H of a triangle $\triangle ABC$. Prove that the reflections of ℓ across $\overleftrightarrow{AB}$, $\overleftrightarrow{BC}$, $\overleftrightarrow{CA}$ pass through a common point lying on the circumcircle of $\triangle ABC$.

4. (Bulgaria, 1997) Let $\triangle ABC$ be a triangle with orthocenter H, and let M and K denote the midpoints of $\overline{AB}$ and $\overline{CH}$. Prove that the internal angle bisectors of $\angle CAH$ and $\angle CBH$ meet at a point on the line $\overleftrightarrow{MK}$.

5. Prove Fact 7.3.3.

7.4 Gergonne and Nagel points

These points are less famous than some of the others, but they make for a few interesting problems, so let's get straight to work.

Fact 7.4.1. *In triangle $\triangle ABC$, the cevians joining each vertex to the point where the incircle touches the opposite side are concurrent.*

The concurrency point in Fact 7.4.1 is called the *Gergonne point* of $\triangle ABC$.

Figure 7.4.1: The Gergonne point (Fact 7.4.1).

Fact 7.4.2. *In triangle $\triangle ABC$, the cevians joining each vertex to the point where the excircle opposite that vertex touches the opposite side are concurrent.*

The concurrency point in Fact 7.4.2 is called the *Nagel point* of $\triangle ABC$.

Figure 7.4.2: The Nagel point (Fact 7.4.2).

Problems for Section 7.4

1. Prove Facts 7.4.1 and 7.4.2.

2. Here is a result analogous to the existence of the Euler line. In triangle $\triangle ABC$, let G, I, N be the centroid, incenter, and Nagel point, respectively. Show that I, G, N lie on a line in that order, and that $NG = 2 \cdot IG$.

3. Continuing the analogy from the previous problem, prove that in triangle $\triangle ABC$, if P, Q, R are the midpoints of sides $\overline{BC}, \overline{CA}, \overline{AB}$, respectively, then the incenter of $\triangle PQR$ is the midpoint of $\overline{IN}$.

7.5 Isogonal conjugates

Two points P and Q inside triangle ABC are said to be *isogonal conjugates* if

$$\angle PAB = \angle QAC, \qquad \angle PBC = \angle QCB, \qquad \angle PCA = \angle QAC.$$

In other words, Q is the reflection of P across each of the internal angle bisectors of $\triangle ABC$.

Fact 7.5.1. *Every point in the interior of $\triangle ABC$ has an isogonal conjugate.*

This instantly gives rise to some new special points of a triangle. For example, the isogonal conjugate of the centroid of $\triangle ABC$ is called the *Lemoine point* ; the cevians through the Lemoine point are called *symmedians*.

Problems for Section 7.5

1. Prove Fact 7.5.1, then formulate and prove a correct version for points not in the interior of the triangle.
2. Prove that in an acute triangle, the orthocenter and the circumcenter are isogonal conjugates. If you completed the previous problem, you should also be able to prove this for a general triangle.
3. Given a triangle, draw through its Lemoine point a line parallel to each side of the triangle, and consider the intersections of that line with the other two sides. Show that the resulting six points are concyclic.
4. Show that the tangents to the circumcircle of a triangle at two vertices intersect on the symmedian of the third vertex.
5. (Dan Moraseski) Let D, E, F be the feet of the symmedians of triangle $\triangle ABC$. Prove that the Lemoine point of $\triangle ABC$ is the centroid of $\triangle DEF$.

7.6 Brocard points

The problems in this section establish the existence and several properties of the Brocard points. Unlike the other special points we have thus far associated to a triangle, the Brocard points are only defined in a cyclically symmetric fashion. Consequently, there are two of them which are interchanged by reversal of the order of the vertices.

Fact 7.6.1. *For any triangle $\triangle ABC$, there exists a unique point P in the interior of $\triangle ABC$ such that*

$$\angle PAB = \angle PBC = \angle PCA.$$

The point P in Fact 7.6.1 is called a *Brocard point* of $\triangle ABC$; there is a second Brocard point obtained by reversing the order of the vertices.

Fact 7.6.2. *The two Brocard points of a triangle are isogonal conjugates.*

This is equivalent to the fact that the common angle in Fact 7.6.1 is the same for the two Brocard points. It is called the *Brocard angle* of $\triangle ABC$; see Problem 7.6.2 for an explicit formula for the Brocard angle.

Problems for Section 7.6

1. Prove Fact 7.6.1.

2. Let ω be the angle such that

   $$\cot \omega = \cot A + \cot B + \cot C.$$

   Prove that the common angle in Fact 7.6.1 is equal to ω; deduce Fact 7.6.2.

3. (IMO **FIXME! (get reference)**) In triangle $\triangle ABC$, put $K = [ABC]$. Prove that

   $$a^2 + b^2 + c^2 \geq 4\sqrt{3}K$$

   by expressing the Brocard angle in terms of a, b, c, K.

4. (IMO 1991/5) Prove that inside any triangle $\triangle ABC$, there exists a point P such that one of the angles $\angle PAB, \angle PBC, \angle PCA$ has measure at most $30°$.

7.7 Frame shift

Once one has gathered up a lot of triangle trivia, it becomes necessary to use it effectively. Often this is accomplished by what I call a "frame shift": you are originally given some points in reference to a given triangle, but you then view them in reference to a different triangle. For instance, given triangle $\triangle ABC$:

1. The orthocenter is the incenter of the orthic triangle.

2. The circumcenter is the orthocenter of the medial triangle.

Problems for Section 7.7

1. (Russia, 2003) Let O and I be the circumcenter and incenter of triangle $\triangle ABC$, respectively. Let ω_A be the excircle of triangle $\triangle ABC$ opposite to A; let it be tangent to $\overleftrightarrow{AB}$, $\overleftrightarrow{AC}$, $\overleftrightarrow{BC}$ at K, M, N, respectively. Assume that the midpoint of segment $\overline{KM}$ lies on the circumcircle of triangle $\triangle ABC$. Prove that O, N, I are collinear.

7.8 Vectors for special points

The vector equations for some of the special points of a triangle $\triangle ABC$ are summarized in the following table. The asterisked expressions assume the circumcenter of the triangle has been chosen as the origin; the origin-independent expressions are not nearly so pleasant to work with!

Circumcenter*	0
Centroid	$\frac{1}{3}(\vec{A} + \vec{B} + \vec{C})$
Orthocenter*	$\vec{A} + \vec{B} + \vec{C}$
Incenter	$\frac{1}{a+b+c}(a\vec{A} + b\vec{B} + c\vec{C})$

Problems for Section 7.8

1. Let P, Q, R be the feet of concurrent cevians in triangle $\triangle ABC$. Determine the vector expression for the point of concurrence in terms of the ratios BP/PC, CQ/QA, AR/RB. Use this formula to extend the above table to other special points. In particular, do so for the Nagel point and obtain an alternate solution to Problem 7.4.2.

2. Let A, B, C, D be four points on a circle. Use the result of Problem 6.4.2 to show that the Simson line of each point with respect to the triangle formed by the other three passes through the midpoint of the segment joining the center of the circle to the centroid of $ABCD$. In particular, the four Simson lines are concurrent.

3. (MOP 1995) Five points are given on a circle. A perpendicular is drawn through the centroid of the triangle formed by three of them, to the chord connecting the remaining two. Similar perpendiculars are drawn for each of the remaining nine triplets of points. Prove that the ten lines obtained in this way have a common point.

4. Compute the distance between the circumcenter and orthocenter of a triangle in terms of the side lengths a, b, c.

5. Show that the distance between the incenter and the nine-point center (see Problem 7.3.5) of a triangle is equal to $R/2 - r$, where r and R are inradius and circumradius, respectively. Deduce *Feuerbach's theorem*, that the incircle and nine-point circle are tangent. (Similarly, one can show the nine-point circle is also tangent to each of the excircles.)

7.9 Additional problems

Here are a few additional problems concerning triangle trivia. Before proceeding to the problems, we state as facts a few standard formulae for the area of a triangle.

Fact 7.9.1. *Let $\triangle ABC$ be a triangle with side lengths $a = BC, b = CA, c = AB$, inradius r, circumradius R, exradius opposite A r_A, semiperimeter s, and area K. Then*

$$\begin{aligned} K &= \frac{1}{2}ab\sin C \quad \textit{(Law of Sines)} \\ &= \frac{abc}{4R} \quad \textit{(by Extended Law of Sines)} \\ &= rs = r_A(s-a) \\ &= \sqrt{s(s-a)(s-b)(s-c)}. \quad \textit{(Heron's formula)} \end{aligned}$$

Problems for Section 7.9

1. Let D be a point on side BC, and let $m = BD, n = CD$ and $d = AD$. Prove *Stewart's formula*:[4]

$$mb^2 + nc^2 = a(d^2 + mn).$$

Figure 7.9.1: Stewart's formula (Problem 7.9.1).

2. Use Stewart's formula to prove the *Steiner-Lehmus theorem*: a triangle with two equal angle bisectors must be isosceles. (A synthetic proof is possible but not easy to find.)

3. (United Kingdom, 1997) In acute triangle $\triangle ABC$, $\overline{CF}$ is an altitude, with F on $\overline{AB}$, and $\overline{BM}$ is a median, with M on $\overline{CA}$. Given that $BM = CF$ and $\angle MBC = \angle FCA$, prove that the triangle $\triangle ABC$ is equilateral. Also, what happens if $\triangle ABC$ is not acute?

4. The point D lies inside the acute triangle $\triangle ABC$. Three of the circumscribed circles of the triangles $\triangle ABC, \triangle BCD, \triangle CDA, \triangle DAB$ have equal radii. Prove that the fourth circle has the same radius, and characterize all such sets of four points. Also, what happens if $\triangle ABC$ need not be acute, or D need not lie in its interior?

5. (Bulgaria, 1997) Let $\triangle ABC$ be a triangle and let M, N be the feet of the angle bisectors of B, C, respectively. Let D be the intersection of the ray $\overrightarrow{MN}$ with the circumcircle of $\triangle ABC$. Prove that

$$\frac{1}{BD} = \frac{1}{AD} + \frac{1}{CD}.$$

6. Let $ABCDE$ be a cyclic pentagon such that $r_{ABC} = r_{AED}$ and $r_{ABD} = r_{ACE}$, where r_{XYZ} denotes the inradius of triangle $\triangle XYZ$. Prove that $AB = AE$ and $BC = DE$.

[4] If written "man + dad = bmb + cnc", this admits the mnemonic "A man and his dad build a bomb in the sink," in case you can recall what the letters stand for before the U.S. Department of Homeland Security pays a visit.

7. (MOP 1990) Let $\overline{AA_1}, \overline{BB_1}, \overline{CC_1}$ be the altitudes in an acute triangle $\triangle ABC$, and let K and M be points on the line segments $\overline{A_1C_1}$ and $\overline{B_1C_1}$, respectively. Prove that if the angles $\angle MAK$ and $\angle CAA_1$ are equal, then the angle $\angle C_1KM$ is bisected by $\overrightarrow{KA}$.

Chapter 8

Quadrilaterals

8.1 General quadrilaterals

There's not a great deal that can be said about an arbitrary quadrilateral: the extra freedom in placing an additional vertex disrupts much of the structure we found in triangles. What little there is to say we offer in the form of a few problems.

Problems for Section 8.1

1. Prove that the midpoints of the sides of any quadrilateral form a parallelogram (known as the *Varignon parallelogram*).

2. Let $ABCD$ be a convex quadrilateral, and let θ be the angle between the diagonals $\overline{AC}$ and $\overline{BD}$. Prove that
$$[ABCD] = \frac{1}{2} AC \cdot BD \sin\theta.$$

3. Derive a formula for the area of a convex quadrilateral in terms of its four sides and a pair of opposite angles.

8.2 Cyclic quadrilaterals

The condition that the four vertices of a quadrilateral lie on a circle gives rise to a wealth of interesting structures, which we investigate in this section.

We start with a classical result of Claudius Ptolemy (85?-165?), who is more famous for his geocentric model of planetary motion.

Theorem 8.2.1 (Ptolemy). *Let $ABCD$ be a convex cyclic quadrilateral. Then*

$$AB \cdot CD + BC \cdot DA = AC \cdot BD.$$

Proof. Mark the point P on $\overline{BD}$ such that $BP = (AB \cdot CD)/AC$, or equivalently $BP/AB = CD/AC$. Since $\angle ABP = \angle ACD$, the triangles ABP and ACD are similar.

Figure 8.2.1: Proof of Ptolemy's theorem (Theorem 8.2.1).

On the other hand, we now have

$$\angle DPA = \pi - \angle APB = \pi - \angle ADC = \angle CBA.$$

Thus the triangles $\triangle APD$ and $\triangle ABC$ are also similar, yielding $DP/BC = AD/AC$. Consequently

$$BD = BP + PD = \frac{AB \cdot CD}{AC} + \frac{AD \cdot BC}{AC}$$

and the theorem follows. □

This proof is elegant, but one cannot help wondering, "How could anyone think of that?" (I wonder that myself; the proof appears in an issue of Samuel Greitzer's *Arbelos*, but he gives no attribution.) The reader might enjoy attempting a proof using trigonometry or complex numbers.

Another important result about cyclic quadrilaterals is an area formula attributed to the ancient Indian mathematician Brahmagupta (598-670).[1]

Fact 8.2.2 (Brahmagupta). *If a cyclic quadrilateral has sides a, b, c, d and area K, then*

$$K = \sqrt{(s-a)(s-b)(s-c)(s-d)},$$

where $s = (a + b + c + d)/2$ is the semiperimeter.

[1] This is a rare case where an Eastern mathematical discovery is reflected by Western naming customs. Compare the situation for Pascal's triangle; see Section 6.3.

Heron's formula for the area of a triangle follows from Brahmagupta's formula by regarding a triangle as a cyclic quadrilateral with one side of length 0.

Problems for Section 8.2

1. Use trigonometry to give another proof of Ptolemy's theorem (Theorem 8.2.1).

2. (Brahmagupta) Let $ABCD$ be a cyclic quadrilateral with perpendicular diagonals. Then the line through the intersection of the diagonals and the midpoint of any side is perpendicular to the opposite side.

3. Use Ptolemy's theorem and the previous problem to give a formula for the lengths of the diagonals of a cyclic quadrilateral in terms of the lengths of the sides.

4. Let $ABCD$ be a cyclic quadrilateral. Prove that the incenters of triangles $\triangle ABC, \triangle BCD, \triangle CDA, \triangle DAB$ form a rectangle.

5. Let $ABCD$ be a cyclic quadrilateral. Prove that the sum of the inradii of $\triangle ABC$ and $\triangle CDA$ equals the sum of the inradii of $\triangle BCD$ and $\triangle DAB$.

8.3 Circumscribed quadrilaterals

The following theorem characterizes circumscribed quadrilaterals; while it can be proved directly using the equal tangents rule, it proves easier to exploit what we already know about incircles and excircles of triangles.

Figure 8.3.1: A circumscribed quadrilateral.

Theorem 8.3.1. *A convex quadrilateral $ABCD$ admits an inscribed circle if and only if $AB + CD = BC + DA$.*

Proof. Let $\overleftrightarrow{AB}$ and $\overleftrightarrow{CD}$ meet at P; without loss of generality, assume A lies between P and B. (We skip the limiting case $AB \parallel CD$.) The quadrilateral $ABCD$ has an inscribed circle if and only if the incircle of $\triangle PBC$ coincides with the excircle of $\triangle PDA$ opposite P. Let Q and R be the points of tangency of $\overleftrightarrow{PB}$ with the incircle of $\triangle PBC$ and the excircle of $\triangle PDA$, respectively; since both circles are tangent to the sides of the angle $\angle CPB$, they coincide if and only if $Q = R$, or equivalently $PQ = PR$. However, by the usual formulae

$$\begin{aligned} PQ &= \frac{1}{2}(PB + PC - BC) = \frac{1}{2}(PD + DC + PA + AB - BC) \\ PR &= \frac{1}{2}(PA + PD + DA) \end{aligned}$$

and these are equal if and only if $AB + CD = BC + DA$. □

Just as with triangles, a convex quadrilateral can have an escribed circle, a circle not inside the quadrilateral but tangent to all four sides (or rather their extensions). We trust the reader can now supply the proof of the analogous

Figure 8.3.2: A quadrilateral having an escribed circle.

characterization of quadraterals admitting an escribed circle.

Fact 8.3.2. *A convex quadrilateral $ABCD$ admits an exscribed circle opposite A or C if and only if $AB + BC = CD + DA$.*

For more problems about circumscribed quadrilaterals, flip back to Section 6.3, where we study them using Brianchon's theorem.

Problems for Section 8.3

1. (IMO 1962/5) On the circle K there are given three distinct points A, B, C. Construct (using only straightedge and compass) a fourth point D on K such that a circle can be inscribed in the quadrilateral thus obtained.

2. (Dick Gibbs) Let $ABCD$ be a quadrilateral inscribed in an ellipse, and let $E = AB \cap CD$ and $F = AD \cap BC$. Show that $ACEF$ can be inscribed in a hyperbola with the same foci as the ellipse. (If you're not familiar with ellipses and hyperbolae, peek ahead to Section 11.3.)

3. (USAMO 1998/6) Let $n \geq 5$ be an integer. Find the largest integer k (as a function of n) such that there exists a convex n-gon $A_1A_2 \cdots A_n$ for which exactly k of the quadrilaterals $A_iA_{i+1}A_{i+2}A_{i+3}$ have an inscribed circle, where $A_{n+j} = A_j$.

8.4 Complete quadrilaterals

A *complete quadrilateral* is the figure formed by four lines, no two parallel and no three concurrent; the *vertices* of a complete quadrilateral are the six pairwise intersections of the lines. This configuration has been widely studied; we present here as problems a number of intriguing properties of the diagram.

In the following problems, let $ABCDEF$ be the complete quadrilateral formed by the lines ABC, AEF, BDF, CDE.

Figure 8.4.1: A complete quadrilateral.

Problems for Section 8.4

1. Show that the orthocenters of the triangles $\triangle ABF$, $\triangle ACE$, $\triangle BCD$, $\triangle DEF$ are collinear. The common line is called the *ortholine* of the complete quadrilateral.

2. Show that the circles with diameters $\overline{AD}, \overline{BE}, \overline{CF}$ are coaxial. Deduce that the midpoints of the segments $\overline{AD}, \overline{BE}, \overline{CF}$ are collinear. (Can you show the latter directly?)

3. Show that the circumcircles of the triangles $\triangle ABF$, $\triangle ACE$, $\triangle BCD$, $\triangle DEF$ pass through a common point, called the *Miquel point* of the complete quadrilateral. (Many solutions are possible.)

4. We are given five lines in the plane, no two parallel and no three concurrent. To every four of the lines, associate the point whose existence was shown in the previous problem. Prove these five points lie on a circle. (This assertion and the previous one belong to an infinite chain of such statements: see W.K. Clifford, *Collected Papers* (1877), 38–54.)

Chapter 9

Geometric inequalities

The subject of geometric inequalities is so vast that it suffices to fill entire books, two notable examples being the volume by Bottema et al. [2] and its sequel [15]. This chapter should thus be regarded more as a sampler of techniques than a comprehensive treatise.

9.1 Distance inequalities

A number of inequalities involve comparing lengths. Useful tools against such problems include:

- Triangle inequality: in triangle $\triangle ABC$, $AB + BC > BC$.
- Hypotenuse inequality: if $\angle ABC$ is a right angle, then $AC > BC$.
- Ptolemy's inequality (Problem 10.3.9): if $ABCD$ is a convex quadrilateral, then $AB \cdot CD + BC \cdot DA \geq AC \cdot BD$, with equality if and only if $ABCD$ is cyclic.
- Erdős-Mordell inequality: see Section 9.4.

Transformations can also be useful, particularly reflection. For example, to find the point P on a fixed line that minimizes the sum of the distances from P to two fixed points A and B, reflect the segment $\overline{PB}$ across the line and observe that the optimal position of P is on the line joining A to the reflection of B.

Figure 9.1.1: Minimizing the distance from a point on a line to two fixed points.

A more dramatic example along the same lines is the following solution (by H.A. Schwarz) to *Fagnano's problem*: of the triangles inscribed in a given acute triangle, which one has the least perimeter? Reflecting the triangle as shown implies that the perimeter of an inscribed triangle is at least the distance from A to its eventual image, with equality when the inscribed triangle makes equal angles with each side. As noted earlier, this occurs for the orthic triangle, which is then the desired minimum.

Figure 9.1.2: Solution of Fagnano's problem.

Problems for Section 9.1

1. For what point P inside a convex quadrilateral $ABCD$ is $PA + PB + PC + PD$ minimized?

2. (Euclid) Prove that the longest chord whose vertices lie on or inside a given triangle is the longest side. (This is intuitively obvious, but make sure your proof is complete.)

3. (Kürschák, 1954) Suppose a convex quadrilateral $ABCD$ satisfies $AB + BD \leq AC + CD$. Prove that $AB < AC$.

4. (USAMO 1999/2) Let $ABCD$ be a cyclic quadrilateral. Prove that

$$|AB - CD| + |AD - BC| \geq 2|AC - BD|.$$

5. (Titu Andreescu and Răzvan Gelca) Points A and B are separated by two rivers. One bridge is to be built across each river so as to minimize the length of the shortest path from A to B. Where should they be placed? (Each river is an infinite rectangular strip, and each bridge must be a straight segment perpendicular to the sides of the river. You may assume that A and B are separated from the intersection of the rivers by a strip wider than the two rivers combined.)

6. Prove that a quadrilateral inscribed in a parallelogram has perimeter no less than twice the length of the shorter diagonal of the parallelogram. (You may want to first consider the case where the parallelogram is a rectangle.)

7. (IMO 1993/4) For three points P, Q, R in the plane, we define $m(PQR)$ as the minimum length of the three altitudes of $\triangle PQR$. (If the points are collinear, we set $m(PQR) = 0$.) Prove that for points A, B, C, X in the plane,
$$m(ABC) \leq m(ABX) + m(AXC) + m(XBC).$$

8. (Sylvester's theorem) A finite set of points in the plane has the property that the line through any two of the points passes through a third. Prove that all of the points are collinear. (As noted in Problem 11.7.8, this result is false in the complex projective plane.)

9. (IMO 1973/4) A soldier needs to check on the presence of mines in a region having the shape of an equilateral triangle. The radius of action of his detector is equal to half the altitude of the triangle. The soldier leaves from one vertex of the triangle. What path should he follow in order to travel the least possible distance and still accomplish his mission?

10. Suppose the largest angle of triangle $\triangle ABC$ is not greater than $120°$. Let D be the third vertex of an equilateral triangle constructed externally on side $\overline{BC}$. For P inside the triangle, show that $PA+PB+PC \geq AD$, and determine when equality holds.

11. Suppose the largest angle of triangle $\triangle ABC$ is not greater than $120°$. Deduce from the previous problem that for P inside the triangle, $PA+PB+PC$ is minimized when $\angle APB = \angle BPC = \angle CPA = 120°$. The point satisfying this condition is known variously as the *Fermat point* or the *Torricelli point*.

12. (IMO 1995/5) Let $ABCDEF$ be a convex hexagon with $AB = BC = CD$ and $DE = EF = FA$, such that $\angle BCD = \angle EFA = \pi/3$. Suppose G and H are points in the interior of the hexagon such that $\angle AGB = \angle DHE = 2\pi/3$. Prove that $AG + GB + GH + DH + HE \geq CF$.

9.2 Algebraic techniques

Another class of methods of attack for geometric inequalities involve invoking algebraic inequalities. The most commonly used is the AM-GM inequality: for $x_1, \dots, x_n > 0$,

$$\frac{x_1 + \cdots + x_n}{n} \geq (x_1 \dots x_n)^{1/n}.$$

Often all one needs is the case $n = 2$, which follows from the fact that

$$(\sqrt{x_1} - \sqrt{x_2})^2 \geq 0.$$

A more sophisticated result is the Cauchy-Schwarz inequality:

$$(x_1^2 + \cdots + x_n^2)(y_1^2 + \cdots + y_n^2) \geq (x_1 y_1 + \cdots + x_n y_n)^2,$$

which one proves by noting that the difference between the left side and the right is

$$\sum_{i<j} (x_i y_j - x_j y_i)^2.$$

A trick that often makes an algebraic approach more feasible, when a problem concerns the side lengths a, b, c of a triangle, is to make the substitution

$$x = s - a, \quad y = s - b, \quad z = s - c,$$

where $s = (a + b + c)/2$. A little algebra gives

$$a = y + z, \quad b = z + x, \quad c = x + y.$$

The point is that the necessary and sufficient conditions $a + b > c, b + c > a, c + a > b$ for a, b, c to constitute the side lengths of a triangle translate into the more convenient conditions $x > 0, y > 0, z > 0$.

Don't forget about the possibility of "algebraizing" an inequality using complex numbers; see Section 2.3.

Problems for Section 9.2

1. (IMO 1988/5) The triangle $\triangle ABC$ has a right angle at A, and D is the foot of the altitude from A. The straight line joining the incenters of the triangles $\triangle ABD, \triangle ACD$ intersects the sides $\overline{AB}, \overline{AC}$ at the points K, L, respectively. S and T denote the areas of the triangles $\triangle ABC$ and $\triangle AKL$, respectively. Show that $S \geq 2T$.

2. Given a point P inside a triangle ABC, let x, y, z be the distances from P to the sides $\overline{BC}, \overline{CA}, \overline{AB}$. Find the point P which minimizes
$$\frac{a}{x} + \frac{b}{y} + \frac{c}{z}.$$

3. If K is the area of a triangle with sides a, b, c, show that
$$ab + bc + ca \geq 4\sqrt{3}K.$$

4. (IMO 1964/2) Suppose a, b, c are the sides of a triangle. Prove that
$$a^2(b + c - a) + b^2(c + a - b) + c^2(a + b - c) \leq 3abc.$$

5. (IMO 1983/6) Let a, b, c be the lengths of the sides of a triangle. Prove that
$$b^2c(b - c) + c^2a(c - a) + a^2b(a - b) \geq 0.$$
(Beware: you may not assume that $a \geq b \geq c$ without loss of generality!)

6. (Balkan, 1996) Let O and G be the circumcenter and centroid of a triangle of circumradius R and inradius r. Show that $OG^2 \leq R^2 - 2Rr$. (This proves Euler's inequality $R \geq 2r$. If you don't know how to compute OG^2, see Problem 7.8.4.)

7. (Murray Klamkin) Let $n > 2$ be a positive integers, and suppose that $a_1, \dots, a_n$ are positive real numbers satisfying the inequality
$$(a_1^2 + \cdots + a_n^2)^2 > (n - 1)(a_1^4 + \cdots + a_n^4).$$
Show that for $1 \leq i < j < k \leq n$, the numbers a_i, a_j, a_k are the lengths of the sides of a triangle.

8. Let $\triangle ABC$ be a triangle with inradius r and circumradius R. Prove that
$$\frac{2r}{R} \leq \sqrt{\cos\frac{A - B}{2} \cos\frac{B - C}{2} \cos\frac{C - A}{2}}.$$

9. (IMO 1995 proposal) Let P be a point inside the convex quadrilateral $ABCD$. Let E, F, G, H be points on sides $\overline{AB}, \overline{BC}, \overline{CD}, \overline{DA}$, respectively, such that $\overleftrightarrow{PE}$ is parallel to $\overleftrightarrow{BC}$, $\overleftrightarrow{PF}$ is parallel to $\overleftrightarrow{AB}$, $\overleftrightarrow{PG}$ is parallel to $\overleftrightarrow{DA}$, and $\overleftrightarrow{PH}$ is parallel to $\overleftrightarrow{CD}$. Let K, K_1, K_2 be the areas of $ABCD, AEPH, PFCG$, respectively. Prove that
$$\sqrt{K} \geq \sqrt{K_1} + \sqrt{K_2}.$$

9.3 Trigonometric inequalities and convexity

A third standard avenue of attack involves reducing a geometric inequality to an inequality involving trigonometric functions. Such inequalities can often be treated using Jensen's inequality for convex functions.

A *convex* function is a function $f(x)$ satisfying the rule
$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$
for all x, y and all $t \in [0, 1]$. Geometrically, this says that the area above the graph of f is a convex set, i.e. that chords of the graph always lie above the graph. Equivalently, tangents to the graph lie below.

Those of you who know calculus can check whether f is convex by checking whether the second derivative of f (if it exists) is always positive. (In some calculus texts, a convex function is called "concave upward", or occasionally is said to "hold water".) Also, if f is continuous, it suffices to check the definition of convexity for $t = 1/2$.

The key fact about convex functions is *Jensen's inequality*, whose proof (by induction on n) is not difficult.

Fact 9.3.1. *Let $f(x)$ be a convex function, and let $t_1, \ldots, t_n$ be nonnegative real numbers adding up to* 1. *Then for all $x_1, \ldots, x_n$,*
$$f(t_1x_1 + \cdots + t_nx_n) \leq t_1f(x_1) + \cdots + t_nf(x_n).$$

For example, the convexity of the function $(-\log x)$ implies the AM-GM inequality.

As a simple example, note that in triangle $\triangle ABC$, we have $\angle A + \angle B + \angle C = \pi$, and the function $f(x) = \sin x$ is concave, so
$$\sin A + \sin B + \sin C \geq 3 \sin \pi/3 = 3\sqrt{3}/2.$$

In other words, the minimum perimeter of a triangle inscribed in a fixed circle is achieved by the equilateral triangle.

Also note that convexity can be used in apparently purely geometric circumstances, thanks to the following fact. (Remember, it suffices to verify this for $t = 1/2$, which is easy.)

Fact 9.3.2. *The distance from a fixed point P is a convex function on the plane. That is, for any points P, Q, R, the distance from P to the point (in vector notation) $tQ + (1-t)R$ is a convex function of t.*

Problems for Section 9.3

1. ([2], 2.7) Show that in triangle $\triangle ABC$, $\sin A \sin B \sin C \leq \frac{3}{8}\sqrt{3}$.
2. Prove that the Brocard angle of a triangle cannot exceed $\pi/6$. (Hint: use Problem 7.6.1, but beware that cot is only convex in the range $(0, \pi/2]$.)
3. ([2], 2.15) Let α, β, γ be the angles of a triangle. Prove that
$$\sin\frac{\beta}{2}\sin\frac{\gamma}{2} + \sin\frac{\gamma}{2}\sin\frac{\alpha}{2} + \sin\frac{\alpha}{2}\sin\frac{\beta}{2} \leq \frac{3}{4}.$$
4. Prove that of the n-gons inscribed in a circle, the regular n-gon has maximum area.
5. ([2], 2.59) Prove that in triangle $\triangle ABC$,
$$1 + \cos A \cos B \cos C \geq \sqrt{3}(\sin A \sin B \sin C).$$
6. Show that for any convex polygon S, the distance from S to a point P (the length of the shortest segment joining P to a point on S) is a convex function of P.
7. (Junior Balkaniad, 1997) In triangle $\triangle ABC$, let D, E, F be the points where the incircle touches the sides. Let r, R, s be the inradius, circumradius, and semiperimeter, respectively, of the triangle. Prove that
$$\frac{2rs}{R} \leq DE + EF + FD \leq s$$
and determine when equality occurs.

8. (MOP 1998) Let $\triangle ABC$ be a acute triangle with circumcenter O, orthocenter H and circumradius R. Show that for any point P on the segment $\overline{OH}$,
$$PA + PB + PC \leq 3R.$$

9.4 The Erdős-Mordell inequality

The following inequality is somewhat more sophisticated than the ones we have seen so far, but is nonetheless useful. It was conjectured by the Hungarian mathematician Pál (Paul) Erdős (1913-1996) in 1935 and first proved by Louis Mordell in the same year.

Theorem 9.4.1. *For any point P inside the triangle $\triangle ABC$, the sum of the distances from P to A, B, C is at least twice the sum of the distances from P to $\overleftrightarrow{BC}, \overleftrightarrow{CA}, \overleftrightarrow{AB}$. Furthermore, equality occurs only when ABC is equilateral and P is its center.*

Proof. The unusually stringent equality condition should suggest that perhaps the proof proceeds in two stages, with different equality conditions. This is indeed the case.

Let X, Y, Z be the feet of the respective perpendiculars from P to $\overline{BC}, \overline{CA}, \overline{AB}$. We will first prove that
$$PA \geq \frac{AB}{BC}PY + \frac{AC}{BC}PZ. \tag{9.4.1.1}$$
The only difference between most proofs of this theorem is in the proof of the above inequality. For example, rewrite (9.4.1.1) as
$$PA \sin A \geq PY \sin C + PZ \sin B,$$
recognize that $PA \sin A = YZ$ by the Extended Law of Sines, and observe that the right side is the length of the projection of $\overline{YZ}$ onto the line $\overleftrightarrow{BC}$. Equality holds if and only if $\overleftrightarrow{YZ}$ is parallel to $\overleftrightarrow{BC}$.

Putting (9.4.1.1) and its analogues together, we get
$$PA+PB+PC \geq PX\left(\frac{CA}{AB} + \frac{AB}{CA}\right)+PY\left(\frac{AB}{BC} + \frac{BC}{AB}\right)+PZ\left(\frac{BC}{CA} + \frac{CA}{BC}\right),$$
with equality if and only if $\triangle XYZ$ is homothetic to $\triangle ABC$; this occurs if and only if P is the circumcenter of $\triangle ABC$ (Problem 1). Now for the

second step: we note that each of the terms in parentheses is at least 2 by the AM-GM inequality. This gives

$$PA + PB + PC \geq 2(PX + PY + PZ),$$

with equality if and only if $AB = BC = CA$. □

Problems for Section 9.4

1. With notation as in the above proof, show that the triangles $\triangle XYZ$ and $\triangle ABC$ are homothetic if and only if P is the circumcenter of $\triangle ABC$.

2. Give another proof of (9.4.1.1) by comparing P with its reflection across the angle bisector of A. (Beware: the reflection may lie outside of the triangle!)

3. Solve problem 7.6.4 using the Erdős-Mordell inequality.

4. (IMO 1996/5) Let $ABCDEF$ be a convex hexagon such that $\overline{AB}$ is parallel to $\overline{DE}$, $\overline{BC}$ is parallel to $\overline{EF}$, and $\overline{CD}$ is parallel to $\overline{FA}$. Let R_A, R_C, R_E denote the circumradii of triangles $\triangle FAB, \triangle BCD, \triangle DEF$, respectively, and let P denote the perimeter of the hexagon. Prove that

$$R_A + R_C + R_E \geq \frac{P}{2}.$$

5. (Nikolai Nikolov) The incircle k of triangle $\triangle ABC$ touches the sides at points A_1, B_1, C_1. For any point K on k, let d be the sum of the distances from K to the sides of the triangle $\triangle A_1B_1C_1$. Prove that $KA + KB + KC > 2d$.

9.5 Additional problems

Now it's your turn. Which technique(s) will help in the following instances?

Problems for Section 9.5

1. Prove that of all quadrilaterals with a prescribed perimeter P, the square has the greatest area. Can you also prove the analogous result for polygons with any number of sides?

2. What is the smallest positive real number r such that a square of side length 1 can be covered by three disks of radius r?

3. Let r be the inradius of triangle $\triangle ABC$. Let r_A be the radius of a circle tangent to the incircle as well as to sides $\overline{AB}$ and $\overline{CA}$. Define r_B and r_C similarly. Prove that

$$r_A + r_B + r_C \geq r.$$

4. Prove that a triangle with angles α, β, γ, circumradius R, and area A satisfies

$$\tan\frac{\alpha}{2} + \tan\frac{\beta}{2} + \tan\frac{\gamma}{2} \leq \frac{9R^2}{4A}.$$

5. Let a, b, c be the sides of a triangle with inradius r and circumradius R. Show that

$$\left|1 - \frac{2a}{b+c}\right| \leq \sqrt{1 - \frac{2r}{R}}.$$

6. Two concentric circles have radii R and R_1 respectively, where $R_1 > R$. $ABCD$ is inscribed in the smaller circle and $A_1B_1C_1D_1$ in the larger one, with A_1 on the extension of CD, B_1 on that of DA, C_1 on that of AB, and D_1 on that of BC. Prove that the ratio of the areas of $A_1B_1C_1D_1$ and $ABCD$ is at least R_1^2/R^2.

7. With the same notation, prove that the ratio of the perimeters of $A_1B_1C_1D_1$ and $ABCD$ is at least R/r.

Part III

Some roads to modern geometry

Chapter 10

Inversive and hyperbolic geometry

One of the features of "modern" geometry is the inclusion of transformations which are more drastic than those considered in Chapter 3. In this chapter, we consider some transformations which preserve angles but not distances or areas or even collinearity. One singularly useful class of examples is the inversions; these give simple proofs both of classic theorems and of competition problems. Moreover, they can be used to give a simple derivation of the basic properties of hyperbolic geometry.

The introduction of inversion requires new concepts of what a "plane" is and what "transformations" are. In particular, though inversion does not preserve lines, it preserves angles in a sense we will make precise. This puts inversion in the rich class of *conformal transformations*, which play a key role in applications (e.g., in physics).

10.1 Inversion

The notion of an inversion is a natural extension of the concept of reflection across a line, once one accepts the idea that lines are really just "circles of infinite radius". Indeed, one can uniformly characterize lines and circles using directed angles: given three points A, B, C, the set of points D for which $\measuredangle ADB = \measuredangle ACB$, together with A, B, C, forms either a line or a circle. So it is not too much of a stretch to imagine a "reflection across a circle"; indeed, this thought seems to have occurred to Apollonius of Perga, who is thought (by virtue of descriptions given by later authors) to have introduced inversion in his lost treatise *Plane Loci.* However, only in modern times did the technique come into common currency; the first surviving appearance of inversion seems to be in the work of the Swiss geometer Jakob Steiner (1796–1863), some of whose profitable use of the technique we will see shortly.

Let O be a point in the plane, and let r be a positive real number. The *inversion* with center O and radius r is the map of the plane minus the point O to itself, carrying the point $P \neq O$ to the point P' on the ray $\overrightarrow{OP}$ such that $OP \cdot OP' = r^2$. Since specifying a point and a positive real number is the same as specifying a circle (the point and the positive real corresponding to the center and radius, respectively, of the circle), we can also speak of inversion through a circle using the same definition.

Figure 10.1.1: Inversion through a circle.

What happens to the point O under inversion? Points near O get sent very far away, in all different directions, so there is no good place to put O itself. To rectify this, we define the *inversive plane* as the usual plane with one additional point ∞, thought of as being a "point at infinity". We view an inversion centered at O as a transformation on the entire inversive plane by declaring that O and ∞ are inverses of each other.

As an aside, we note a natural interpretation of the inversive plane. Under stereographic projection (used in some maps), the surface of a sphere, minus the North Pole, is mapped to a plane tangent to the sphere at the South Pole as follows: a point on the sphere maps to the point on the plane collinear with the given point and the North Pole. Then the point at infinity corresponds to the North Pole, and the inversive plane corresponds to the whole sphere. In fact, inversion through the South Pole with the appropriate radius corresponds to reflecting the sphere through the plane of the equator!

Figure 10.1.2: A stereographic projection.

Returning to Euclidean geometry, we now establish some important properties of inversion. We first make an easy but important observation.

Fact 10.1.1. *If O is the center of an inversion taking P to P' and Q to Q', then the triangles $\triangle OPQ$ and $\triangle OQ'P'$ are oppositely similar.*

In particular, we have that $\measuredangle OP'Q' = -\measuredangle OQP$, a fact underlying our next proof.

By an *inversive circle* in the inversive plane, we will mean either a circle in the Euclidean plane, or a line in the Euclidean plane together with the extra point ∞.

Theorem 10.1.2. *The image of an inversive circle under an inversion is an inversive circle.*

Proof. Let A, B, C, D be four points on an inversive circle, and let A', B', C', D' be the respective images of A, B, C, D under an inversion with center O. We now chase directed angles, using the similar triangles of Fact 10.1.1:

$$\begin{aligned}
\measuredangle A'B'C' &= \measuredangle A'B'O + \measuredangle OB'C' \\
&= \measuredangle BAO + \measuredangle OCB \\
&= \measuredangle BAD + \measuredangle DAO + \measuredangle OCD + \measuredangle DCB \\
&= \measuredangle DAO + \measuredangle OCD \\
&= \measuredangle A'D'O + \measuredangle OD'C' \\
&= \measuredangle A'D'C'.
\end{aligned}$$

We see that A', B', C', D' lie on an inversive circle as well. □

Figure 10.1.3: Proof of Theorem 10.1.2.

Notice the way the angles are broken up and recombined in the above proof. In some cases, inversion can turn a constraint involving two or more angles in different places into a constraint about a single angle, which then is easier to work with. Some examples can be found in the problems.

Inversion also turns out to "reverse the angles between lines". Since lines are sent to circles in general, we will have to define the angle between two circles to make sense of this statement.

Given two inversive circles ω_1 and ω_2, the (directed) angle between them at one of their intersections P is defined as the (directed) angle from the tangent to ω_1 at P to the tangent of ω_2 at P. We say that two inversive circles are *orthogonal* if the angle between them is $\pi/2$. Note that absent a choice between the two points of intersection, the angle between two circles is only well-defined up to sign as an angle modulo π; however, orthogonality does not depend on this choice. Note also that a line and a circle are orthogonal if and only if the line passes through the center of the circle.

Fact 10.1.3. *The directed angle between circles (at a chosen intersection) is reversed under inversion.*

Distances don't fare as well under inversion, but one can say something using Fact 10.1.1.

Fact 10.1.4 (Inversive distance formula)**.** *If O is the center of an inversion of radius r sending P to P' and Q to Q', then*

$$P'Q' = PQ \cdot \frac{r^2}{OP \cdot OQ}.$$

Problems for Section 10.1

1. Deduce Theorem 10.1.2 from Problem 4.2.5 (or use the above proof to figure out how to do that problem).
2. Give another proof of Theorem 10.1.2 using the converse of the power-of-a-point theorem (Fact 6.1.2) and Fact 10.1.4.
3. The angle between two lines through the origin is clearly preserved under inversion. Why doesn't this contradict the fact that inversion reverses angles?
4. (IMO 1996/2) Let P be a point inside triangle $\triangle ABC$ such that

   $$\angle APB - \angle ACB = \angle APC - \angle ABC.$$

   Let D, E be the incenters of triangles $\triangle APB, \triangle APC$, respectively. Prove that $\overline{AP}, \overline{BD}, \overline{CE}$ meet in a point. (Many other solutions are possible; over 25 were submitted by contestants at the IMO!)
5. (IMO 1998 proposal) Let $ABCDEF$ be a convex hexagon such that $\angle B + \angle D + \angle F = 360°$ and

   $$\frac{AB}{BC} \cdot \frac{CD}{DE} \cdot \frac{EF}{FA} = 1.$$

   Prove that

   $$\frac{BC}{CA} \cdot \frac{AE}{EF} \cdot \frac{FD}{DB} = 1.$$

6. Prove that the following are equivalent:

    1. The points A and B are inverses through the circle ω.
    2. The line $\overleftrightarrow{AB}$ and the circle with diameter $\overline{AB}$ are both orthogonal to ω.
    3. ω is a circle of Apollonius with respect to A and B.

    In particular, conclude that a circle distinct from ω is fixed (as a whole, not pointwise) by inversion through ω if and only if it is orthogonal to ω.

7. Give yet another proof of Theorem 10.1.2 using complex numbers and the circle of Apollonius (Theorem 6.5.1).

8. Show that a set of circles is coaxial if and only if there is a circle orthogonal to all of them. Deduce that coaxial circles remain that way under inversion. Also, try drawing a family of coaxial circles and some circles orthogonal to them; the picture is very pretty.

9. Prove that any two nonintersecting circles can be inverted into concentric circles. (This will be used in Theorem 10.2.2 below.)

10.2 Inversive magic

As noted earlier, we know about inversion largely through the work of Jakob Steiner. Steiner used the technique to give dazzlingly simple proofs of a number of difficult-looking statements. Here we present but a few examples.

We start with a classical result attributed to Pappus of Alexandria. It is one of a number of results concerning a figure bounded by three semicircles with diameters $\overline{AB}, \overline{BC}, \overline{AC}$, where A, B, C are three points lying on a line in that order. Such a figure was first consirede by Archimedes, who called it an *arbelos*[1].

Theorem 10.2.1 (Pappus)**.** *Let ω be a semicircle with diameter $\overline{AB}$. Let ω_1 and ω_2 be two semicircles externally tangent to each other at C, and internally tangent to ω at A and B, respectively. Let $C_1, C_2, \dots$ be a sequence of circles,*

[1]The word "arbelos" in Greek refers to a shoemaker's knife, which presumably looked something like the figure Archimedes was considering.

each tangent to ω and ω_1, such that C_i is tangent to C_{i+1} and C_1 is tangent to ω_2 (as in Figure 10.2.1). Let r_n be the radius of C_n, and let d_n be the distance from the center of C_n to $\overleftrightarrow{AB}$. Then for all n,

$$d_n = 2nr_n.$$

Figure 10.2.1: An arbelos, and a theorem of Pappus.

Proof. Perform an inversion with center A, and choose the radius of inversion so that C_n remains fixed. Then ω and ω_1 map to lines perpendicular to $\overleftrightarrow{AB}$ and tangent to C_n, and $C_{n-1}, \dots, C_1$ to a column of circles between the lines, with ω_2' at the bottom of the column. The relation $d_n = 2nr_n$ is now obvious. □

Figure 10.2.2: Proof of Theorem 10.2.1.

The following theorem is known as *Steiner's porism*.

Theorem 10.2.2. *Suppose two nonintersecting circles have the property that one can fit a "ring" of n circles between them, each tangent to the next. Then one can do this starting with any circle tangent to both given circles.*

Proof. By Problem 10.1.9, a suitable inversion takes the given circles to concentric circles, while preserving tangency of circles. The result is now obvious. □

Figure 10.2.3: Steiner's porism (Theorem 10.2.2) and its proof.

Problems for Section 10.2

1. Suppose that, in the hypotheses of Pappus's theorem, we assume that C_0 is tangent to ω, ω_1 and the line $\overleftrightarrow{AB}$ (instead of the semicircle ω_2). Show that in this case $d_n = (2n+1)r_n$.

2. (Romania, 1997) Let ω be a circle, and let $\overleftrightarrow{AB}$ be a line not intersecting ω. Given a point P_0 on ω, define the sequence $P_0, P_1, \dots$ as follows: P_{n+1} is the second intersection with ω of the line $\overleftrightarrow{BQ_n}$, where Q_n is the second intersection of the line $\overleftrightarrow{AP_n}$ with ω. Prove that for a positive integer k, if $P_0 = P_k$ for some choice of P_0, then $P_0 = P_k$ for any choice of P_0.

10.3 Inversion in practice

So far we have seen that inversion can be used to give spectacular proofs of a few results. However, it is much more useful than that; it can often be applied to solve much more mundane problems. The paradigm for doing this is almost always the following: invert the given picture and its conclusion, thus transforming the original problem into a new problem on a new diagram, then solve the new problem. In some cases, one must also superimpose the original and inverted diagrams (as in the proof of Theorem 10.2.1) and/or compare information in the two diagrams (e.g. using Fact 10.1.4).

A general principle behind this method is that it is easier to deal with lines than circles. Hence if one wishes to perform an inversion on a geometric diagram, one should center the inversion at a point which is "busy" in the sense of having many relevant circles and lines passing through it.

Problems for Section 10.3

1. Make up an inversion problem by reversing the paradigm: start with a result that you know, invert about some point, and see what you get. The tricky part is choosing things well enough so that the resulting problem doesn't have an obvious busy point; such a problem would be too easy!

2. Let C_1, C_2, C_3, C_4 be circles such that C_i and C_{i+1} are externally tangent for $i = 1, 2, 3, 4$ (where $C_5 = C_1$). Prove that the four points of tangency are concyclic.

3. (Romania, 1997) Let $\triangle ABC$ be a triangle, let D be a point on side $\overline{BC}$, and let ω be the circumcircle of $\triangle ABC$. Show that the circles tangent to $\omega, \overline{AD}, \overline{BD}$ and to $\omega, \overline{AD}, \overline{DC}$ are tangent to each other if and only if $\angle BAD = \angle CAD$.

Figure 10.3.1: Problem 10.3.3.

4. (Russia, 1995) Draw a semicircle with diameter $\overline{AB}$ and center O, then draw a line which intersects the semicircle at C and D and which intersects line $\overleftrightarrow{AB}$ at M, such that $MB < MA$ and $MD < MC$. Let K be the second point of intersection of the circumcircles of triangles $\triangle AOC$ and $\triangle DOB$. Prove that $\angle MKO$ is a right angle.

5. (USAMO 1993/2) Let $ABCD$ be a convex quadrilateral with perpendicular diagonals meeting at O. Prove that the reflections of O across $\overleftrightarrow{AB}$, $\overleftrightarrow{BC}$, $\overleftrightarrow{CD}$, $\overleftrightarrow{DA}$ are concyclic. (For an added challenge, find a non-inversive proof as well.)

6. (Apollonius's problem) Given three nonintersecting circles, how many circles are tangent to all three? And how can they be constructed with straightedge and compass?

7. (IMO 1994 proposal) The incircle of $\triangle ABC$ touches $\overline{BC}, \overline{CA}, \overline{AB}$ at D, E, F, respectively. Let X be a point inside $\triangle ABC$ such that the incircle of $\triangle XBC$ touches $\overline{BC}$ at D also, and touches $\overline{CX}$ and $\overline{XB}$ at Y and Z, respectively. Prove that $EFZY$ is a cyclic quadrilateral.

8. (Israel, 1995) Let $\overline{PQ}$ be the diameter of semicircle H. The circle O is internally tangent to H and is tangent to $\overline{PQ}$ at C. Let A be a point on H, and let B be a point on $\overline{PQ}$ such that $\overleftrightarrow{AB}$ is perpendicular to $\overleftrightarrow{PQ}$ and is also tangent to O. Prove that $\overleftrightarrow{AC}$ bisects $\angle PAB$.

9. Give an inversive proof of Ptolemy's inequality (Theorem 2.3.1).

10. (IMO 1993/2) Let A, B, C, D be four points in the plane, with C, D on the same side of line $\overleftrightarrow{AB}$, such that $AC \cdot BD = AD \cdot BC$ and $\angle ADB = \pi/2 + \angle ACB$. Find the ratio $(AB \cdot CD)/(AC \cdot BD)$ and prove that the circumcircles of triangles $\triangle ACD$ and $\triangle BCD$ are orthogonal.

11. (Iran, 1995) Let M, N, P be the points of intersection of the incircle of $\triangle ABC$ with sides BC, CA, AB, respectively. Prove that the orthocenter of $\triangle MNP$, the incenter of $\triangle ABC$, and the circumcenter of $\triangle ABC$ are collinear.

12. (MOP 1997) Let $\triangle ABC$ be a triangle and let O be its circumcenter. The lines $\overleftrightarrow{AB}$ and $\overleftrightarrow{AC}$ meet the circumcircle of triangle $\triangle BOC$ again at B_1 and C_1, respectively. Let D be the intersection of lines $\overleftrightarrow{BC}$ and $\overleftrightarrow{B_1C_1}$. Show that the circle tangent to $\overleftrightarrow{AD}$ at A and having its center on $\overleftrightarrow{B_1C_1}$ is orthogonal to the circle with diameter $\overline{OD}$.

13. (Russia, 1993) Let $ABCD$ be a convex cyclic quadrilateral, and let O be the intersection of diagonals $\overline{AC}$ and $\overline{BD}$. Let ω_1 and ω_2 be the circumcircles of triangles $\triangle ABO$ and $\triangle CDO$, respectively, and let ω_1 and ω_2 meet at O and K. The line through O parallel to $\overleftrightarrow{AB}$ meets ω_1 again at L, and the line through O parallel to $\overleftrightarrow{CD}$ meets ω_2 again at M. Let P and Q be points on segments $\overline{OL}$ and $\overline{OM}$, respectively, such that $OP/PL = MQ/QO$. Prove that O, K, P, Q lie on a circle.

10.4 Hyperbolic geometry: an historical aside

One cannot give a survey of "modern" geometry without including hyperbolic, or non-Euclidean, geometry. Originally viewed as a pathological construction, it was later realized in several ways *within* the confines of Euclidean geometry, and thus is no less valid! Subsequently, hyperbolic geometry has become omnipresent within mathematics, and even within physics via Einstein's theory of relativity.

To understand the relevance of hyperbolic geometry, we must momentarily overturn our revisionist construction of the Euclidean plane and go back to the axiomatic definition. It relies on five postulates, which we loosely translate into modern language.

1. Any two points are the endpoints of a line segment.

2. Any line segment can be extended to a straight line.

3. There exists a circle with any given radius and center.

4. Any two right angles are congruent to each other.

5. If two lines intersect a third and the interior angles on one side are both less than π, then the two lines intersect somewhere on that side of the third line.

One cannot know whether Euclid realized it was necessary to include the fifth postulate, the so-called *parallel postulate.* For many centuries, it was felt that Euclid had simply fallen short in simplifying the axioms, and that it would be possible to deduce the parallel postulate from the other four. It was finally realized by Gauss that this is impossible, as there is actually a perfectly sensible (if highly counterintuitive) geometry in which the parallel postulate failed while the other postulates continue to hold. Gauss, no fan of controversy[2], never published his findings, leaving them to be rediscovered independently by János Bolyai (1802–1860) and Nikolai Ivanovich Lobachevsky[3] (1792–1856).

As noted above, one proves the independence of the parallel postulate by constructing a "model geometry" in which the parallel postulate fails while the other postulates continue to hold. This is done by building a geometric situation and carefully relabeling the objects; we will do this in the next section.

10.5 Poincaré's models of hyperbolic geometry

As described in the previous section, one typically builds spaces of "hyperbolic geometry" by realizing them using constructions within Euclidean geometry. We now describe two related methods for doing this, introduced by Henri Poincaré (1854–1912) .

In the *disc model* , we take the underlying set of points to be the interior of an open disc in the Euclidean plane. The lines of the disc model are the lines and circles orthogonal to the boundary of the disc, or rather the pieces of these lying within the disc. See Figure 10.5.1.

In the *halfplane model* , we take the underlying set of points to be those $(x, y) \in \mathbb{R}^2$ with $y > 0$. The lines of the disc model are the lines and

[2]Although the consistency of non-Euclidean geometry is a nonissue for mathematicians in our time, it continues to cause controversy among nonmathematicians, who have trouble shaking the belief that there could be any alternative to the Euclidean setting. An infamous example is columnist Marilyn vos Savant, who published a notorious book attacking the work of Wiles on Fermat's Last Theorem on precisely these long-discredited grounds.

[3]There has been some dispute over whether these rediscoveries were truly independent of each other and of Gauss. This dispute is satirized in a famous song by mathematician/satirist Tom Lehrer.

Figure 10.5.1: Points and lines in the disc model.

circles orthogonal to the x-axis, or rather the pieces of these lying within the halfplane.

Figure 10.5.2: Points and lines in the halfplane model.

At this point it is clear that one can transform the disc model and the halfplane model into each other by inversion: inverting through a point on the boundary of the disc turns the disc model into the halfplane model, and one gets back by inverting through a point below the x-axis. We may thus safely refer to either one as the *hyperbolic plane* , as long as we only refer to concepts which carry identical meanings in both model. This includes lines of the model; by virtue of the angle-preserving property of inversion, this also includes angles between lines.

Fact 10.5.1. *Any two points in the hyperbolic plane lie on a unique line. Any two lines in the hyperbolic plane intersect in at most one point.*

By contrast with the Euclidean plane, there are multiple lines through a given point which fail to intersect a given line not through that point. That is, there are multiple parallels to a given line through a given point not on the line.

Figure 10.5.3: Multiple parallels in the hyperbolic plane.

It is reasonable to ask why one needs two (or more, but two will suffice for now) different models of the hyperbolic plane. One answer is that different symmetries appear more readily in each model, so having multiple models makes it easier to visualize the full set of symmetries of the hyperbolic plane. We illustrate this with an example.

By a *hyperbolic transformation,* we will mean a bijection from the hyperbolic plane to itself carrying lines to lines and preserving angles. For example, in the disc model, one may rotate around the center of the disc; in the halfplane model, one may make a horizontal translation or a homothety with positive ratio and center on the x-axis.

Theorem 10.5.2. *(a) There exists a hyperbolic transformation carrying any given point to any other given point.*

(b) There exists a hyperbolic transformation carrying any given line to any other given line.

(c) Given a point P in the hyperbolic plane, there exists a hyperbolic transformation carrying any given line through P to any other given line through P.

Proof. Fix an isomorphism between the disc model and the halfplane model. Let O be the center of the disc model, and let O' be its image in the halfplane model.

(a) This is clear in the halfplane model, using dilations and translations.

(b) On one hand, any line in the halfplane model can be translated to a line through O'. On the other hand, any two lines through O in the disc model are related by a rotation. This yields the claim.

(c) By (b), it suffices to check this for the point O in the disc model, which is clear.

□

10.6 Hyperbolic distance

We next wish to define a notion of distance between two points in the hyperbolic plane. Before doing this, we first check that there are not "too many" hyperbolic transformations.

Theorem 10.6.1. *Let A, B, C be three distinct points in the hyperbolic plane, which lie on a line in that order. Then there is no hyperbolic transformation fixing A and taking B to C.*

Proof. By Theorem 10.5.2, we may reduce to the case where A is the center of the disc model and the line ℓ through A, B, C is diametric. Let ℓ_1, ℓ_2 be the lines through B, C, respectively, perpendicular to ℓ. Let m be a line through A such that in the Euclidean plane, the extensions of ℓ_2 and m meet on the boundary of the disc. Note that m does not meet ℓ_2 in the hyperbolic plane, but it does meet ℓ_1.

Figure 10.6.1: Proof of Theorem 10.6.1.

Suppose now that there is a hyperbolic transformation fixing A and taking B to C. Then ℓ maps to ℓ, and by angle preservation, ℓ_1 maps to ℓ_2. Again by angle preservation, m maps either to m or to its reflection across ℓ. In either case, the two intersecting lines m, ℓ_1 are carried to nonintersecting lines, contradiction. □

Let A, B be two distinct points in the hyperbolic plane; we now define the *distance* $d_h(A, B)$ as follows. (If $A = B$, we just set $d_h(A, B) = 0$.) Apply Theorem 10.5.2 to map A, B to points A', B' which lie on a vertical line in the halfplane model. Without loss of generality, we may assume A' lies below B'. Let d_A, d_B be the Euclidean distances between A, B and the x-axis, and define

$$d_h(A, B) = \log d_B - \log d_A;$$

this is unambiguous by Theorem 10.6.1, since we cannot use a hyperbolic transformation to move B' up or down the line while fixing A'.

Fact 10.6.2. *The hyperbolic distance satisfies the following properties.*

(a) If points A, B, C lie on a line in that order, then $d_h(A, C) = d_h(A, B) + d_h(B, C)$.

(b) For any points A, B, C, D, we have $d_h(A, B) = d_h(C, D)$ if and only if there is a hyperbolic transformation sending A to B and sending C to D.

Notice something funny going on: the definition of a hyperbolic transformation only required preservation of lines and angles, and yet these also preserve distances. This is in contrast with the Euclidean plane, where there is a clear distinction between similarities and rigid motions. Somehow the hyperbolic plane has an inherent "sense of scale" that the Euclidean plane does not; this can be explained by formalizing the statement that the "curvature" of the Euclidean plane is zero but that of the hyperbolic plane is nonzero.

Problems for Section 10.6

1. Give a formula to compute distances in the disc model.

2. Prove that any line in the hyperbolic plane contains pairs of points whose distance is arbitrarily large; i.e., the length of a line is infinite.

3. Prove that any map of the hyperbolic plane to itself that carries lines to lines preserves (undirected) angles, and hence is a hyperbolic transformation. That is, there is no analogue in the hyperbolic plane of affine transformations which are not rigid motions.

10.7 Hyperbolic triangles

A *(line) segment* in the hyperbolic plane will be the segment or arc between two points on a hyperbolic line; we refer to the distance between the two endpoints also as the *length* of the segment. With this definition, we may now speak about polygons in the hyperbolic plane, and in particular of triangles.

Theorem 10.7.1. *The sum of the angles in a hyperbolic triangle is always strictly less than* π*.*

Proof. Again, fix an isomorphism between the disc model and the halfplane model. Let O be the center of the disc model, and let O' be its image in the halfplane model. Given a triangle in the halfplane model, we can apply dilations and translations corresponding to hyperbolic transformations to create a congruent triangle with O' in its interior. Then the result is clear: each angle of the hyperbolic triangle is less than the corresponding angle of the ordinary triangle with the same vertices. □

Figure 10.7.1: Proof of Theorem 10.7.1.

The difference between π and the sum of the angles in a hyperbolic triangle is called the *angular defect* of the triangle. It is additive in the sense that if A, B, C are three points in the hyperbolic plane, and D is a point on the hyperbolic segment BC, then the angular defect of the hyperbolic triangle ABC is the sum of the angular defects of ABD and ADC. It can thus be used as a measure of area for hyperbolic triangles.

Figure 10.7.2: Additivity of angular defect.

Problems for Section 10.7

1. Prove that any two hyperbolic triangles which have the same angles are congruent. Yes, you read that correctly! This is another case where the hyperbolic plane exhibits an intrinsic "sense of scale".

Chapter 11

Projective geometry

Projective geometry is the study of geometric properties which are invariant under "changes of perspective"; this eliminates properties like angles and distances but retains properties like collinearity and concurrence. The formalism of projective geometry makes a discussion of such properties possible, and exposes some remarkable facts, such as the duality of points and lines.

The history of projective geometry is a remarkable instance of art and science feeding off one another.[1] Based on the optics studies of the Arabic mathematician Alhazen (Ibu Ali al-Hasan ibn al-Haytham) (965–1040), several early Renaissance artists[2] attempted to develop a style of visual depiction that presented the eye with a truer semblance of three-dimensional space than did earlier, flatter styles. The discovery of the principle of linear perspective (the idea that all parallel lines appear to converge at a single point) is credited to Filippo Brunelleschi (1377–1446). This led to a flurry of activity, culminating in the work of Girard Desargues (1591–1661), which introduced projective geometry as we now it.

In the modern era, the real power of projective geometry lies within the realm of algebraic geometry, i.e., the study of geometric objects defined by polynomial equations. This study, implicit in the coordinate geometry with which this book begins, took off in earnest late in the 19th century, and remains one of the most vital branches of present-day mathematics research. We end the chapter with a glimpse in this direction.

---

[1]The MacTutor archive, mentioned in the introduction, includes a nice description of this history.

[2]To be fair, the distinction between artists and scientists was somewhat blurred at this period, whence the modern phrase "Renaissance man" for a versatile individual.

11.1 The projective plane

We begin with a lengthy description of the formalism of the projective plane. The impatient reader may wish to read only the next paragraph at first, then skip to the later sections and come back to this section as needed.

The *projective plane* consists of the standard Euclidean plane, together with a set of points called *points at infinity*, one for each collection of parallel lines. We say that a line passes through the point at infinity corresponding to its direction (and no others), and that all of the points at infinity lie on a *line at infinity*. Note that three parallel lines now indeed have a common point at infinity, which retroactively justifies our calling such lines "concurrent".

An alternate description of the projective plane turns out to be quite useful, and corresponds more closely to the artists' conception. View the Euclidean plane as some plane in three-dimensional space, and fix a point O not on the plane (corresponding to the eye). Then each point on the plane corresponds to a line through O passing through that point, but not all lines through O correspond to points on the Euclidean plane. In fact, they correspond to the points at infinity. In other words, we can identify the projective plane with the set of lines in space passing through a fixed point.

This description also yields a natural coordinate system for the projective plane, using what are known as *homogeneous coordinates*. Each point in the projective plane can be specified with a triple of numbers $[x : y : z]$, where x, y, z are not all zero. Be careful, though: for any nonzero real number λ, $[x : y : z]$ and $[\lambda x : \lambda y : \lambda z]$ are the same point! (Hence the name "homogeneous coordinates".) The colons are meant to remind you that it is the ratios between the coordinates that are well-defined, not the individual coordinates themselves.

How are homogeneous coordinates related to the usual Cartesian coordinates on the Euclidean plane? If we embed the Euclidean plane in space as the plane $z = 1$, then the point with Cartesian coordinates (x, y) has homogeneous coordinates $[x : y : 1]$, and the points at infinity are the points of the form $[x : y : 0]$ for some x, y not both zero.

11.2 Projective transformations

The original definition of a projective transformation corresponded to the process of projecting an image in the "real world" onto an artist's canvas.

Again, fix a point O in three-dimensional space, and now select two planes not passing through O. The mapping that takes each point P on the first plane to the intersection of the line $\overleftrightarrow{OP}$ with the second plane was defined as a projective transformation. (Do you see why this map makes sense over the whole projective plane?)

One can also give an algebraic description of projective transformations that accommodates degenerate cases slightly more easily. In terms of homogeneous coordinates, a projective transformation takes the form

$$[x : y : z] \mapsto [ax + by + cz : dx + ey + fz : gx + hy + iz],$$

where $a, \dots, i$ are numbers such that the 3×3 matrix

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

is invertible, i.e., its determinant

$$aei + bfg + cdh - ceg - afh - bdi$$

is nonzero. From this description it is clear that affine transformations are projective as well, since they occur when $g = h = 0$. Since we have two additional parameters (it looks like three, but by homogeneity one parameter is superfluous), the following analogue of Fact 3.5.2 is no surprise.

Fact 11.2.1. *Any four points, no three collinear, can be mapped to any other four such points by a unique projective transformation.*

The most common use of a projective transformation in problem-solving is to map a particular line to the point at infinity. (As with inversion, it pays to look for a "busy" line for this purpose.) If the statement to be proved is well-behaved under projective transformations, this can yield drastic simplifications. The "well-behaved" concepts mainly consist of incidence properties between points and lines (concurrence, collinearity, and the like); as with affine transformations, angles and distances are not preserved, nor are areas or ratios of lengths along segments (unlike affine transformations).

We demonstrate the power of projection by reproving Desargues's theorem (Theorem 5.2.2).

Theorem 11.2.2. *Given triangles $\triangle ABC$ and $\triangle DEF$, the points $\overleftrightarrow{AB} \cap \overleftrightarrow{DE}$, $\overleftrightarrow{BC} \cap \overleftrightarrow{EF}$, $\overleftrightarrow{CA} \cap \overleftrightarrow{FD}$ are collinear if and only if the lines $\overleftrightarrow{AD}, \overleftrightarrow{BE}, \overleftrightarrow{CF}$ are concurrent.*

Proof. Apply a projective transformation to place the points $\overleftrightarrow{AB} \cap \overleftrightarrow{DE}$ and $\overleftrightarrow{BC} \cap \overleftrightarrow{EF}$ on the line at infinity. If triangles $\triangle ABC$ and $\triangle DEF$ are perspective from a line, they now have parallel sides and so are homothetic; thus the lines $\overleftrightarrow{AD}, \overleftrightarrow{BE}, \overleftrightarrow{CF}$ concur at the center of homothety (or at a point at infinity, in case $\triangle ABC$ and $\triangle DEF$ are now congruent). Conversely, if the lines $\overleftrightarrow{AD}, \overleftrightarrow{BE}, \overleftrightarrow{CF}$ concur at P, consider the homothety centered at A carrying A to D. It preserves the line $\overleftrightarrow{BE}$ and carries the line $\overleftrightarrow{AB}$ to the parallel line $\overleftrightarrow{DE}$ through B, so it maps D to E. Similarly, the homothety maps C to F, and so $\overleftrightarrow{CA}$ and $\overleftrightarrow{FD}$ are also parallel, implying that the points $\overleftrightarrow{AB} \cap \overleftrightarrow{DE}$, $\overleftrightarrow{BC} \cap \overleftrightarrow{EF}$, $\overleftrightarrow{CA} \cap \overleftrightarrow{FD}$ are collinear along the line at infinity. □

Beware that angles, circles, and other "metric" objects are not preserved under projection; we will learn more about getting around this difficulty later in the chapter.

Problems for Section 11.2

1. Use a projective transformation to give an alternate proof of Pappus's theorem.

2. Prove that the center of a circle drawn in the plane cannot be constructed with straightedge alone.

3. (Original) Let $ABCDE$ be the vertices of a convex pentagon, and let $F = \overline{BC} \cap \overline{DE}$, $G = \overline{CD} \cap \overline{EA}$, $H = \overline{DE} \cap \overline{AB}$, $I = \overline{EA} \cap \overline{BC}$, $J = \overline{AB} \cap \overline{CD}$. Show that $\overleftrightarrow{BD} \cap \overleftrightarrow{CE}$ lies on the line $\overleftrightarrow{AF}$ if and only if $\overleftrightarrow{GH} \cap \overleftrightarrow{IJ}$ does.

Figure 11.2.1: Problem 11.2.3.

11.3 A conic section

We now introduce the notion of a conic section, which comes to us from the work of the ancient geometer Apollonius (whose name has arisen already in connection with Theorem 6.5.1).

A *conic section* is classically defined as a cross-section of a right circular cone by a plane not passing through a vertex, where the cone extends infinitely far in *both* directions. The section is a called an *ellipse*, a *parabola*, or a *hyperbola*, depending on whether the angle between the plane and the axis of the cone is greater than, equal to, or less than $\pi/4$.

Figure 11.3.1: Conic sections.

Theorem 11.3.1. *An ellipse is the locus of points whose sum of distances to two fixed points is constant. Similarly, a hyperbola is the locus of points whose (absolute) difference of distances to two fixed points is constant.*

Proof. This was already known to Apollonius, but the following clever proof was found by Germinal Dandelin (1794-1847). We will describe only the case of the ellipse, as the hyperbola case is similar.

Inscribe spheres in the cone on either side of the plane of the ellipse, one on the side of the vertex of the cone, tangent to the plane at A, the other tangent to the plane at B. For any point on the cone between the two

Figure 11.3.2: Dandelin's proof of Theorem 11.3.1.

spheres, the sum of the lengths of the tangents to the two spheres is clearly a constant. On the other hand, for any point on the cone also lying in the plane, the segments to A and B are also tangent to the respective spheres, so the sum of their lengths equals this constant. The result follows. $\square$

The two points alluded to in the above theorem are called *foci* (plural of *focus*). The name comes from the fact that if one has an ellipse made of a reflective material and one places a light source at one focus, all of the light rays will be "focused" at the opposite focus (see Problem 2).

For parabolas, one has the following alternate version of Theorem 11.3.1.

Fact 11.3.2. *A parabola is the locus of points whose distance to a fixed point is equal to the distance to a fixed line.*

The fixed line and point are called the *focus* and *directrix*, respectively, of the parabola.

In modern times, it was noted that conic sections have a nice description in terms of Cartesian coordinates. If $z^2 = x^2 + y^2$ is the equation of the cone, it is evident that any cross-section is defined by setting some quadratic polynomial in x and y to 0. Hence a conic section can alternatively be defined as the zero locus of a quadratic polynomial; one must impose mild extra conditions to avoid degenerate cases, such as a pair of lines (which geometrically arise from planes through the vertex of the cone). Unless we say otherwise, our conic sections will be required to be nondegenerate.

Here are some standard equations for the conic sections:

Type	Standard equation
Ellipse	$x^2/a^2 + y^2/b^2 = 1$
Parabola	$y = ax^2 + bx + c$
Hyperbola	$x^2/a^2 - y^2/b^2 = 1$

Also, the equation $xy = 1$ defines a *rectangular* hyperbola, one with perpendicular asymptotes. (The *asymptotes* of a hyperbola are its tangent lines at its intersections with the line at infinity.)

Problems for Section 11.3

1. Prove that a line tangent to an ellipse makes equal (undirected) angles with the segments from the two foci to the point of tangency. What are the analogous properties of a tangent to a parabola or hyperbola?

Figure 11.3.3: A line tangent to an ellipse (Problem 11.3.1).

2. Prove that two hyperbola branches which share a focus can meet in at most two points (whereas two hyperbolas can meet in four points).

3. (Anning-Erdős) An infinite set of points in the plane has the property that the distance between any two of the points is an integer. Prove that the points are all collinear.

4. Let P and Q be two points on an ellipse. Prove that there exist ellipses similar to the given one, externally tangent to each other, and internally tangent to the given ellipse at P and Q, respectively, if and only if P and Q are antipodes.

5. Use the previous problem to prove that the maximum distance between two points on an ellipse is the length of the major axis *without* doing any calculations.

6. (Original) Prove that the convex quadrilateral $ABCD$ contains a point P such that the incircles of triangles $\triangle PAB$ and $\triangle PBC$ are tangent, as are those of $\triangle PBC$ and $\triangle PCD$, of $\triangle PCD$ and $\triangle PDA$, and of $\triangle PDA$ and $\triangle PAB$, if and only if $ABCD$ has an inscribed circle.

7. Find all points on the conic $x^2 + y^2 = 1$ with *rational* coordinates x, y as follows: pick a point (x, y) with rational coordinates, and project the conic from (x, y) onto a fixed line (e.g. the line at infinity). More generally, given a single rational point on a conic whose defining equation has rational coefficients, this procedure allows you to describe all such points.

11.4 Conics in the projective plane

In this section, we discuss conic sections from the point of view of projective geometry. To start, we rephrase the geometric definition of a conic section.

Fact 11.4.1. *A curve is a conic section if and only if it is the image of a circle under a suitable projective transformation.*

In particular, the theorems of Pascal and Brianchon continue to hold if the circle in the statement of either theorem is replaced with an arbitrary conic. From these one can deduce converse theorems, that a hexagon is inscribed in (resp. circumscribed about) a conic if and only if it satisfies the conclusion of Pascal (resp. Brianchon); thinking of Pappus's theorem, one realizes that the conics in the previous statement must be permitted to be degenerate.

We also note that the classification of conics can be restated in terms of projective geometry.

Fact 11.4.2. *A conic is an ellipse (or a circle) if and only if it does not meet the line at infinity. A conic is a parabola if and only if it is tangent to the line at infinity. A conic is a hyperbola if and only if it intersects the line at infinity in two distinct points.*

Problems for Section 11.4

1. Prove that a hexagon whose opposite sides meet in collinear points is inscribed in a conic (which may degenerate to a pair of lines).

2. Let $\triangle ABC$ and $\triangle BCD$ be equilateral triangles. An arbitrary line through D meets $\overleftrightarrow{AB}$ at M and $\overleftrightarrow{AC}$ at N. Determine the acute angle between the lines $\overleftrightarrow{BN}$ and $\overleftrightarrow{CM}$.

3. (Poncelet-Brianchon theorem) Let A, B, C be three points on a rectangular hyperbola (a hyperbola with perpendicular asymptotes). Prove that the orthocenter of the triangle $\triangle ABC$ also lies on the hyperbola. There are other special points of $\triangle ABC$ which must lie on this hyperbola; can you find any?

4. (*Monthly*, Oct. 1994) Let $A_1, A_2, A_3, A_4, A_5, A_6$ be a hexagon circumscribed about a conic, and form the intersections $P_i = \overleftrightarrow{A_iA_{i+2}} \cap \overleftrightarrow{A_{i+1}A_{i+3}}$ ($i = 1, \dots, 6$, all indices modulo 6). Show that the P_i are the vertices of a hexagon inscribed in a conic.

5. (*Arbelos*) Let A, B, C be three noncollinear points. Draw ellipses E_1, E_2, E_3 with foci B and C, C and A, A and B, respectively. Show that:

    1. Each pair of ellipses meet in exactly two points, where a point of tangency counts twice. (In general, two ellipses can meet in as many as four points.)
    2. The three lines determined by these pairs of points are concurrent.

11.5 The polar map and duality

Fix a circle ω with center O. The *polar map* (or *polar transformation*) with respect to ω interchanges points and lines in the following manner:

1. If P is a finite point other than O, the *pole* of P is the line p through P' perpendicular to $\overleftrightarrow{PP'}$, where P' is the inverse of P through ω.

2. If p is a finite line not passing through O, the *polar* of p is the inverse through ω of the foot of the perpendicular from O to p.

Figure 11.5.1: The polar map with respect to a circle.

3. If P is a point at infinity, the pole of P is the line through O perpendicular to any line through P, and vice versa.

4. If P is O, the pole of P is the line at infinity, and vice versa.

The polar map is also known as *reciprocation*. We keep the notational convention that points are labeled with capital letters and their poles with the corresponding lowercase letters.

Fact 11.5.1. *The polar map satisfies the following properties:*

1. *Every point is the polar of its pole, and every line is the pole of its polar.*

2. *The polar of the line through the points A and B is the intersection of the poles a and b.*

3. *Three points are collinear if and only if their poles are concurrent.*

An obvious consequence of the existence of the polar map is the duality principle.

Fact 11.5.2 (Duality principle)**.** *A theorem of projective geometry remains true if the roles of points and lines are interchanged.*

For example, the dual of one direction of Desargues's theorem is the other direction.

We can now give Brianchon's original proof of his theorem, using Pascal's theorem and the polar map. There's nothing to it, really: given a hexagon circumscribed about a circle ω, apply the polar map with respect to ω. The result is a hexagon inscribed in ω, and the collinearity of the intersections of opposite sides translates back to the original diagram as the concurrence of the lines through opposite vertices.

Problems for Section 11.5

1. Make up a problem by starting with a result that you know and applying the polar map. Beware that circles not concentric with ω do not behave well under the polar map; see below.

2. State the dual of Pappus's theorem. Can you prove this directly? (A projection may help.)

3. State and prove a dual version of problem 8.3.3. Since circles do not dualize to circles, you will have to come up with a new proof!

4. (China, 1996) Let H be the orthocenter of acute triangle $\triangle ABC$. The tangents from A to the circle with diameter $\overline{BC}$ touch the circle at P and Q. Prove that P, Q, H are collinear.

5. Let $\triangle ABC$ be a triangle with incenter I. Fix a line ℓ tangent to the incircle of $\triangle ABC$ (not containing any of the sides). Let A', B', C' be points on ℓ such that

$$\angle AIA' = \angle BIB' = \angle CIC' = \pi/2.$$

   Show that $\overleftrightarrow{AA'}, \overleftrightarrow{BB'}, \overleftrightarrow{CC'}$ are concurrent.

6. (Răzvan Gelca) Let A, B, C, D be four points on a circle. Show that the pole of $\overleftrightarrow{AC} \cap \overleftrightarrow{BD}$ with respect to this circle passes through $\overleftrightarrow{AB} \cap \overleftrightarrow{CD}$ and $\overleftrightarrow{AD} \cap \overleftrightarrow{BC}$. Use this fact to give another solution to Problem 6.2.6 (IMO 1985/6).

7. We know what happens to points and lines under the polar map, but what about a curve? If we view the curve as a *locus*, i.e. a set of points, its dual is a set of lines which form an *envelope*, i.e. they are all tangent to some curve. Show that the dual of a conic, under this definition, is

   Figure 11.5.2: The envelope of a family of lines.

   again a conic. However, the dual of a circle need not be a circle.

8. Let ω be a (nondegenerate) conic. Show that there exists a unique map on the projective plane, taking points to lines and vice versa, satisfying the properties in Fact 11.5.1, and taking each point on ω to the tangent to ω through that point. This map is known as the *polar map with respect to ω* (and coincides with the first definition if ω is a circle).

9. (IMO 1998/5) Let I be the incenter of triangle $\triangle ABC$. Let the incircle of $\triangle ABC$ touch the sides $\overline{BC}, \overline{CA}, \overline{AB}$ at K, L, M, respectively. The line through B parallel to $\overleftrightarrow{MK}$ meets the lines $\overleftrightarrow{LM}$ and $\overleftrightarrow{LK}$ at R and S, respectively. Prove that angle $\angle RIS$ is acute.

11.6 Cross-ratio

From the discussion so far, it may appear that there is no useful notion of distance in projective geometry, for projective transformations do not preserve Euclidean distances, or even ratios of distances along a line (which affine transformations do preserve). There is something to be salvaged here, though; the "ratios of ratios of distances" are preserved.

Given four collinear points A, B, C, D, the *cross-ratio* of these points is defined as the following signed ratio of lengths:

$$\frac{AC \cdot BD}{AD \cdot BC}.$$

In case one of these points is at infinity, the definition can be extended by declaring that the ratio of two infinite distances is 1. We have left the definition where all of the points lie at infinity as an exercise.

In light of duality, we ought to be able to make this definition for four concurrent lines, and in fact we can: the cross-ratio of four lines a, b, c, d is defined as the cross-ratio of the intersections A, B, C, D of a, b, c, d with some line ℓ not passing through the point of concurrency. The cross-ratio is well-defined by the following observation, which follows from several applications of the Law of Sines.

Fact 11.6.1. *Let a, b, c, d be four concurrent lines and ℓ a line meeting a, b, c, d at A, B, C, D, respectively. Then*

$$\frac{AC \cdot BD}{AD \cdot BC} = \frac{\sin \measuredangle(a,c) \sin \measuredangle(b,d)}{\sin \measuredangle(b,c) \sin \measuredangle(a,d)}.$$

Fact 11.6.2. *The cross-ratio is invariant under projective transformations and the polar map.*

In case the cross-ratio is -1, we say C and D are *harmonic conjugates* with respect to A and B (or vice versa). If you did Problem 5.2.3, you witnessed the most interesting property of harmonic conjugates: if P is any point not on the line and Q is any point on $\overleftrightarrow{PC}$ other than P or C, then $\overleftrightarrow{AP} \cap \overleftrightarrow{BQ}$, $\overleftrightarrow{AQ} \cap \overleftrightarrow{BP}$ and D are collinear. (Not surprisingly, this property is projection-invariant.)

One nice application of cross-ratios is the following characterization of conics.

Fact 11.6.3. *Given four points A, B, C, D, the locus of points E such that the cross-ratio of the lines $\overleftrightarrow{AE}, \overleftrightarrow{BE}, \overleftrightarrow{CE}, \overleftrightarrow{DE}$ is constant is a conic.*

Problems for Section 11.6

1. How should the cross-ratio be defined along the line at infinity?
2. Let A, B, C, D be four points on a circle. Show that for E on the circle, the cross-ratio of the lines $\overleftrightarrow{EA}, \overleftrightarrow{EB}, \overleftrightarrow{EC}, \overleftrightarrow{ED}$ remains constant. Then use this to deduce Fact 11.6.3.
3. ("Butterfly problem") Let M be the midpoint of chord $\overline{XY}$ of a circle, and let $\overline{AB}$ and $\overline{CD}$ be chords passing through M. Let $E = \overleftrightarrow{AD} \cap \overleftrightarrow{XY}$ and $F = \overleftrightarrow{BC} \cap \overleftrightarrow{XY}$. Prove that $EM = MF$.
4. The points A, B, C, D, in this order, lie on a straight line. A circle k passes through B and C, and the lines $\overleftrightarrow{AM}, \overleftrightarrow{AN}, \overleftrightarrow{DK}, \overleftrightarrow{DL}$ are tangent to k at M, N, K, L. The lines $\overleftrightarrow{MN}, \overleftrightarrow{KL}$ intersect $\overleftrightarrow{BC}$ at P, Q.
   (a) Prove that P and Q do not depend on k.
   (b) If $AD = a$, $BC = b$, and the segment $\overline{BC}$ moves along $\overline{AD}$, find the minimum length of segment $\overline{PQ}$.

11.7 The complex projective plane

The homogeneous coordinates we have worked with so far also make sense for complex numbers, though visualizing the result is substantially harder. The set of points they define (i.e. the set of proportionality classes of ordered triples of complex numbers, not all zero) is called the *complex projective plane*. We define lines and conics in this new plane simply as the zero loci of linear and quadratic polynomials, respectively.

One handy feature of the complex projective plane is the following characterization of circles.

Fact 11.7.1. *A nondegenerate conic is a circle if and only if it passes through the points* $[1 : i : 0]$ *and* $[1 : -i : 0]$.

These two points are called the *circular points at infinity*, or simply the *circular points* for short.

The fact that complex circles always meet the line at infinity in two points, while real circles to not, is a symptom of the key fact that the complex numbers are *algebraically closed*, i.e. every polynomial with complex coefficients has a complex root. (This is the Fundamental Theorem of Algebra, first proved by Gauss.) This means, for example, that we have the following:

Fact 11.7.2. *In the complex projective plane, two conics meet in exactly four points (counting points of tangency twice).*

In fact, a more general result is true, which we will not prove; it is attributed to Etienne Bézout[3] (1730–1783).

Theorem 11.7.3 (Bézout). *The zero loci of two polynomials, of degrees* m *and* n*, contains exactly* mn *points if the loci meet transversally everywhere (i.e. at each intersection, each locus has a well-defined tangent line, and the tangent lines are distinct).*

If the loci do not meet transversally, e.g. if they are tangent somewhere, one must correctly assign multiplicities to the intersections to make the count work.

An interesting consequence of Bezout's theorem, which we will prove independently, is due to Michel Chasles[4] (1793–1880). The zero locus of a polynomial of degree 3 is known as a *cubic curve*.

Theorem 11.7.4 (Chasles). *Let* C_1 *and* C_2 *be two cubic curves meeting in exactly nine distinct points. Then any cubic curve passing through eight of the points passes through the ninth point.*

[3] There seems to be some disagreement over whether this name is spelled "Bezout" or "Bézout"; we use the MacTutor spelling.

[4] "Chasles" is pronounced "shell".

Proof. The set of homogeneous degree 3 polynomials in x, y, z is a 10-dimensional vector space (check by writing a basis of monomials); let Q_1 and Q_2 be polynomials with zero loci C_1 and C_2, respectively, and let $P_1, \dots, P_9$ be the nine intersections of C_1 and C_2. Note that no four of these points lie on a line and no seven lie on a conic, or else each of C_1 and C_2 would have this line or conic as a component, and their intersection would be infinite rather than nine points.

Let d_i be the dimension of the space of degree 3 polynomials vanishing at $D_1, \dots, D_i$ (and put $d_0 = 10$); then for $i \leq 8$, d_i equals either $d_{i-1} - 1$ or d_i, the latter only if every cubic curve passing through $P_1, \dots, P_{i-1}$ also passes through P_i. However, this turns out not to be the case; see the problems. Thus $d_8 = 2$, and we already have two linearly independent polynomials in this space, namely Q_1 and Q_2. (If they were dependent, they would define the same curve, and again the intersection would be infinite.) Thus if C is a cubic curve defined by a polynomial Q that passes through $P_1, \dots, P_8$, then $Q = aQ_1 + bQ_2$ for some $a, b \in \mathbb{C}$, and so Q also vanishes at P_9, as desired. □

These results are just the tip of a rather sizable iceberg. The modern subject of *algebraic geometry* is concerned with the study of zero loci of sets of polynomials in spaces of any dimension. It interacts with almost every other branch of mathematics, including complex analysis, topology, number theory, combinatorics, and mathematical physics. Unfortunately, the subject as practiced today has become technically involved[5]; the novice should start with a book written in the "classical" style, such as Harris [10] or Shafarevich [16], before proceeding to a "modern" text such as Eisenbud and Harris [6] or Hartshorne [11]. (If it is not already clear from the rhapsodic tone of this section, algebraic geometry, particularly in connection with number theory, ranks among the author's main research interests.)

Problems for Section 11.7

1. Give another proof that there is a unique conic passing through any

[5]Algebraic geometry flourished in Italy in the early 20th century, but it was practiced with a flagrant lack of rigor that led to numerous errors. To fix these, it proved necessary to recast the foundations of the topic; this was accomplished in the 1960s under the guidance of Alexander Grothendieck (1928–??). What the new foundations gain in power and flexibility, they lack in accessibility; the most accessible route to them seems to be via [6].

five points, using the circular points.

2. Make up a problem by taking a projective statement you know and projecting two of the points in the diagram to the circular points. (One of my favorites is the radical axis theorem—which becomes a projective statement if you replace the circles by conics through two fixed points!)

3. Deduce Pascal's theorem from Chasles' theorem applied to a certain degenerate cubic.

4. Prove that given eight or fewer points in the plane, no four on a line and no seven on a conic, one of which is labeled P, there exists a cubic curve passing through all of the points but P.

5. A cubic curve which is nondegenerate, and additionally has no *singular point* (a point where the partial derivatives of the defining homogeneous polynomial all vanish, like the point $[0 : 0 : 1]$ on the curve $y^2z = x^3 + x^2z$) is called an *elliptic curve*.[6] Let E be an elliptic curve, and pick a point O on E. Define "addition" of points on E as follows: given points P and Q, let R be the third intersection of the line $\overleftrightarrow{PQ}$ with E, and let $P + Q$ be the third intersection of the line $\overleftrightarrow{OR}$ with E. Prove that $(P + Q) + R = P + (Q + R)$ for any three points P, Q, R, i.e. that "addition is associative". (If you know what a group is, show that E forms a group under addition, by showing that there exist inverses and an identity element.) For more on elliptic curves, and their role in number theory, see [18].

6. Give another solution to problem 10.2.2 using a well-chosen projective transformation in the complex projective plane.

7. One can define addition on a curve on a singular cubic in the same fashion, as long as none of the points involved is a singular point of the cubic. Use this fact to give another solution to Problem 10.2.2.

8. Let E be an elliptic curve. Show that there are exactly nine points at which the tangent line at E has a triple, not just a double, intersection with the curve (and so meets the curve nowhere else). These points

[6]The geometry of elliptic curves pervades much of modern number theory, e.g., the proof of Fermat's Last Theorem given in 1995 by Andrew Wiles (1953–). See [18] for a gentle introduction, or [17] for a more comprehensive treatment.

are called *flexes*. Also show that the line through any two flexes meets E again at another flex. (Hence the flexes constitute a counterexample to Problem 9.1.8 in the complex projective plane!)

9. (Poncelet's porism) Let ω_1 and ω_2 be two conic sections. Given a point P_0 on ω_1, let P_1 be either of the points on ω_1 such that the line $\overleftrightarrow{P_0P_1}$ is tangent to ω_2. Then for $n \geq 2$, define P_n as the point on ω_1 other than P_{n-2} such that $\overleftrightarrow{P_{n-1}P_n}$ is tangent to ω_2. Suppose there exists n such that $P_0 = P_n$ for a particular choice of P_0. Show that $P_0 = P_n$ for any choice of P_0.

Part IV

Odds and ends

Hints

Here are the author's suggestions on how to proceed on some of the problems. If you find another solution to a problem, so much the better—but it may not be a bad idea to try to find the suggested solution anyway!

We have refrained from including detailed solutions to all of the problems; for the justification of this decision, and for a web location at which solutions can be found, see the Introduction.

1.2.2 Imitate the proof of Theorem 6.5.1.

3.1.5 Consider the triangle $\triangle AB_1C_1$ together with the second intersection of the circumcircles of $\triangle AB_1C_2$ and $\triangle AB_3C_1$. Show that this figure is congruent to the two analogous figures formed from the other triangles. Do this by rotating $\triangle AB_1C_1$ onto $\triangle C_2AB_2$ onto $\triangle B_3C_3A$ and tracing what happens to the figure. (Or apply Theorem 4.3.1.)

3.2.5 Consider the homothety around D taking B to C. If you knew the problem were true, what would that say about the image of E? Once you figure that out, work backwards. (It may help to peek ahead to Chapter 7.)

3.3.2 After applying Theorem 4.3.1, this should bear a strong resemblance to Problem 3.1.5.

3.3.3 How does P_2 depend on P_1?

4.1.6 The octahedron has 4 times the volume of the tetrahedron. What happens when you glue them together at a face?

4.3.1 Prove one assertion, then work backward to prove the other.

4.3.3 Construct two of the intersections of the trisections, complete the equilateral triangle, then show that its third vertex is the third intersection. This is difficult; if you're still stuck, see [5].

5.2.2 Draw 10 points: the 6 vertices of the triangles, the three intersections of corresponding sides, and the intersection of the lines joining two pairs of corresponding vertices. If you relabel these 10 points appropriately, this diagram will turn into a case of the forward direction of Desargues!

6.2.1 Draw the circumcircle of ABC, and apply the radical axis theorem to that circle, ω_1, and ω_2.

6.2.6 There are several solutions to this problem, but no one of them is easy to find. In any case, before anything else, find an extra cyclic quadrilateral.

6.3.4 Work backwards, defining G as the point for which the conclusion holds. Also consider the circumcircle of CDE.

6.4.2 Find a cyclic hexagon.

6.4.3 Use Theorem 4.2.3.

6.6.2 Even using directed angles, the result fails for nonconvex hexagons. Figuring out why may help you determine how to use convexity here.

6.5.1 Given segments AB and CD, what conditions must the center P of a spiral similarity carrying AB to CD satisfy?

6.5.2 By Ceva and Menelaus, one can show $BA_1/A_1C = BA_2/A_2C$. This means the circle with diameter A_1A_2 is a circle of Apollonius with respect to B and C.

7.2.5 The center of the circle lies at C.

7.2.6 The fixed point lies on the circumcircle of ABC.

7.2.7 Show that the point F is the excenter of ACD opposite A.

7.2.8 Use homothety.

7.2.9 The incircle of triangle $O_1O_2O_3$ touches O_2O_3 at C. Reformulate the problem in terms of $O_1O_2O_3$ and get rid of the circles. From there, one way to proceed is to calculate where along ℓ the intersection with AO_1 is.

7.3.5 For (a), write the half-turn as the composition of two other homotheties and locate the fixed point.

7.5.4 Use circles of Apollonius.

7.6.1 What is the locus of points where one of these equalities holds?

7.7.1 The frame shift here is to consider the triangle formed by the excenters.

7.8.4 The distance d satisfies $9d^2 = a^2 + b^2 + c^2$.

7.9.1 Apply the Law of Cosines to the triangles ABD and ACD.

8.2.4 Use Fact 7.2.3.

8.4.2 Show that the orthopole is the radical axis of any two of the circles.

8.3.3 Show that no two consecutive quadrilaterals can both have incircles.

9.1.4 Use the similar triangles formed by the sides and diagonals.

9.2.8 Write everything in terms of $\cot A/2$ and the like. Then turn the result into a statement about homogeneous polynomials using the identity

$$\cot\frac{A}{2} + \cot\frac{B}{2} + \cot\frac{C}{2} = \cot\frac{A}{2}\cot\frac{B}{2}\cot\frac{C}{2},$$

and solve the result.

9.2.9 Use an affine transformation to make $ABCD$ cyclic, and perform a quadrilateral analogue of the $s - a$ substitution.

9.4.4 A certain special case of this result is equivalent to Erdős-Mordell. Modify the proof slightly to accommodate the generalization.

10.1.9 Which circles are orthogonal to two concentric circles?

10.3.6 Reduce to the case where two of the circles are tangent, then invert.

10.3.11 The paradigm does not hold here. Invert through the incircle, then superimpose the original and inverted diagrams.

10.3.12 Note that $AB \cdot AB_1 = AC \cdot AC_1$. Also look at the intersection of OA and B_1C_1.

10.3.13 The busy point is O. After you invert there, the conclusion is that K', P', Q' are collinear, and the hypothesis on P and Q should look like a criterion for collinearity.

11.2.2 Find a projective transformation taking the circle to itself but not preserving its center.

11.4.1 Fix five of the points and compare the locus of sixth points making this condition hold with the conic through the five points.

11.4.3 Apply Pascal's theorem to the hyperbola, using the intersections of the asymptotes with the line at infinity as two of the six points.

11.5.6 Draw the circle with diameter OB, and show that its common chord with the circle centered at O is concurrent with KN and AC.

11.7.4 In fact, there exists a degenerate cubic with this property.

11.7.6 Find a projective transformation taking the circle to a circle and the line to infinity.

11.7.9 As in Steiner's porism, reduce to the case of two concentric circles.

Suggested further reading

The definition of “reading” here is expansive: it includes electronic resources such as software packages (for dynamic geometry) and Web resources (for competitions).

Algebraic geometry

As noted in Section 11.7, one should start with a text written in “traditional” language, such as those by Harris [10], Shafarevich [16], or Cox, Little and O’Shea [3].

Competitions

The Art of Problem Solving web site,

`http://www.artofproblemsolving.com/`

is the premier web resource for students interested in problem solving of the sort appearing in competitions like the USAMO and IMO.

Dynamic geometry

The phrase “dynamic geometry” refers to computer software that can render a geometric configuration in a fashion that allows the user to vary the determining data and witness the change in the resulting configuration in real time. (For example, if it appears that three lines are concurrent, one can test this hypothesis by “jiggling the input data” to see whether the concurrence appears to be coincidental or causal.) There are several outstanding programs for doing this: commercial offerings include *Cabri*, *Cinderella*, and *The Geometer’s Sketchpad*, while slimmer noncommercial alternatives include *Kgeo* and *Kseg*.

Geometric inequalities

The compilation [2] is the definitive source, while its sequel [15] details more recent results.

Hyperbolic geometry

The book [19] is a charming introduction to the topic, spinning a tale of Lewis Carroll, his friend and muse Alice Liddell, and a mysterious stranger as they explore unfamiliar geometric territory.

Miscellaneous

The book [1] is a nice survey of “modern” geometry in various forms: it includes sections on hyperbolic geometry, spherical geometry, projective geometry, and constructibility.

Bibliography

[1] A. Baragar, *A Survey of Classical and Modern Geometries (with Computer Activities)*, Prentice-Hall, New Jersey, 2001.

[2] O. Bottema, R. Ž. Djordjević, R. R. Janić, D. S. Mitrinović, and P.M. Vasić, *Geometric Inequalities*, Wolters-Noordhoff, Groningen, 1968.

[3] D. Cox, J. Little, and D. O'Shea, *Ideals, Varieties and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*, second edition, Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1997.

[4] H.S.M. Coxeter, *Non-Euclidean Geometry*, sixth edition, Mathematical Association of America (MAA Spectrum), Washington, 1998.

[5] H.S.M. Coxeter and S.L. Greitzer, *Geometry Revisited*, Mathematical Association of America (NML 19), Washington, 1967.

[6] D. Eisenbud and J. Harris, *The Geometry of Schemes*, Springer-Verlag (Graduate Texts in Mathematics 197), New York, 2000.

[7] H. Eves, *A Course in Geometry* (2 volumes).

[8] S. Greitzer, *Arbelos*, volumes 1–5, American Mathematics Competitions, 1982–1988.

[9] V. Guillemin and A. Pollock, *Differential Topology*, Prentice-Hall, Englewood Cliffs, 1974.

[10] J. Harris, *Algebraic Geometry: A First Course*, Springer-Verlag (Graduate Texts in Mathematics 133), New York, 1992.

[11] R. Hartshorne, *Algebraic Geometry*, Springer-Verlag (Graduate Texts in Mathematics 52), New York, 1977.

[12] K.S. Kedlaya, B. Poonen, and R. Vakil, *The William Lowell Putnam Competition 1985–2000: Problems, Solutions and Commentary*, Problem Books series, MAA (Washington), 2002.

[13] C. Kimberling, *Triangle Centers and Central Triangles*, Utilitas Mathematica Publishing (Congressus Numerantium 129), 1998.

[14] E. A. Maxwell, *Fallacies in Mathematics*, Cambridge University Press, 1959.

[15] D.S. Mitrinović et al, *Recent Advances in Geometric Inequalities*, Kluwer Academic Publishers, Boston, 1989.

[16] I.R. Shafarevich, *Basic Algebraic Geometry* (translated by Miles Reid), Springer-Verlag, New York, 1994.

[17] J.H. Silverman, *The Arithmetic of Elliptic Curves*, Graduate Texts in Mathematics **106**, Springer-Verlag (New York), 1986.

[18] J.H. Silverman and J. Tate, *Rational Points on Elliptic Curves*, Springer-Verlag (Undergraduate Texts in Mathematics), New York, 1992.

[19] M. Sved, *Journey into Geometries*, MAA Spectrum, Mathematical Association of America (Washington), 1991.

[20] J. van Yzeren, A simple proof of Pascal's hexagon theorem, *American Mathematical Monthly* **100** (1993), 930–931.

[21] I.M. Yaglom, *Geometric Transformations*, Random House (New Mathematical Library 8), New York, 1973.

About the license

This book is being distributed in an "open source" fashion, much in the manner that open source software like the Linux operation system is distributed. This confers certain rights on you, the reader, but also carries certain restrictions that limit your ability to restrict the rights of others. This chapter includes a little information about how the open source model operates both in general and in this particular instance; more specifics can be found in the text of the GNU Free Documentation License (GFDL), , which appears in the following chapter.

Open source for text?

This book has been released under the GNU Free Documentation License (GFDL) in order to promote the free and open dissemination of the ideas contained herein, and as an experiment in collaborative authorship. The GFDL expressly permits unlimited distribution of this document either intact or in modified forms, through either commercial or noncommercial means. The main restrictions it imposes are that:

- the human-editable source file(s) for any modified form of this document must be made freely available;
- all modified forms must themselves carry the GFDL.

Note that the second item means that no material may be incorporated into a modified form of this book which carries a copyright restriction less permissive than the GFDL (i.e., most material carrying an outstanding non-GFDL copyright), unless the use of that material constitutes a *fair use* in the sense of US copyright law. On the other hand, other GFDL material, material on which the copyright has expired, and material in the public domain are fair game.

Source code distribution

Like a computer program, a piece of electronic text is usually used in a "processed" form, such as what you are (probably) reading now, but is created in a "source" form which is easier for the author to manipulate. In the case of this book, the source code is input for Knuth's TeX typesetting system, or more precisely for the variant known as AMSLaTeX. One requirement of the GFDL is that any processed document must either include, or direct the reader to an electronic location of, all source files. In this case, these source files will be distributed at the following URL:

`http://math.mit.edu/~kedlaya/geometryunbound`.

This site will also include some material not included in the present document, such as diagrams and solutions.

History

This section is mandated by the GFDL to carry the modification/derivation history of this document. If you prepare a modification of this document, make sure to update this section!

The present document is the original version to carry the GFDL, released 18 Jan 2006.

GNU Free Documentation License

Preamble

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondarily, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

1. Applicability and Definitions

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The **"Document"**, below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as **"you"**. You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A **"Modified Version"** of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A **"Secondary Section"** is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The **"Invariant Sections"** are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The **"Cover Texts"** are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A **"Transparent"** copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called **"Opaque"**.

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The **"Title Page"** means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

A section **"Entitled XYZ"** means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as **"Acknowledgements"**, **"Dedications"**, **"Endorsements"**, or **"History"**.) To **"Preserve the Title"** of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

2. Verbatim Copying

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3. Copying in Quantity

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

4. Modifications

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.

B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.

C. State on the Title page the name of the publisher of the Modified Version, as the publisher.

D. Preserve all the copyright notices of the Document.

E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.

F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.

G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.

H. Include an unaltered copy of this License.

I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.

J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.

K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.

L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.

M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.

N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.

O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the

Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties–for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

5. Combining Documents

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements", and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements".

6. Collections of Documents

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

7. Aggregation with Independent Works

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an "aggregate" if the copyright resulting from the compilation is not used to limit the legal rights of the compilation's users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

8. Translation

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements", "Dedications", or "History", the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

9. Termination

You may not copy, modify, sublicense, or distribute the Document except as expressly provided for under this License. Any other attempt to copy, modify, sublicense or distribute the Document is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

10. Future Revisions of This License

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See `http://www.gnu.org/copyleft/`.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the

Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation.

Addendum: How to use this License for your documents

To use this License in a document you have written, include a copy of the License in the document and put the following copyright and license notices just after the title page:

> Copyright ©YEAR YOUR NAME. Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.2 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the section entitled "GNU Free Documentation License".

If you have Invariant Sections, Front-Cover Texts and Back-Cover Texts, replace the "with... Texts." line with this:

> with the Invariant Sections being LIST THEIR TITLES, with the Front-Cover Texts being LIST, and with the Back-Cover Texts being LIST.

If you have Invariant Sections without Cover Texts, or some other combination of the three, merge those two alternatives to suit the situation.

If your document contains nontrivial examples of program code, we recommend releasing these examples in parallel under your choice of free software license, such as the GNU General Public License, to permit their use in free software.