A Readable Introduction to Real Mathematics

(PRELIMINARY DRAFT)

Daniel Rosenthal, David Rosenthal and Peter Rosenthal

Contents

Chapter 11

Fundamentals of Euclidean Plane Geometry

In this chapter we describe the fundamentals of Euclidean geometry of the plane in a way that relies on some intuitively-apparent properties of geometric figures. In particular, we begin by making a (reasonable) assumption about the Euclidean plane: between any two points there exists a unique line that extends infinitely in two directions. Later on in the chapter we will need to make an additional, very important, assumption known as the Parallel Postulate. More rigorous axiomatic approaches to Euclidean geometry are possible.

11.1 Triangles

One basic concept is that of a *triangle*, by which we mean a geometric figure consisting of three points (called its *vertices*) which do not all lie on one line, and the line segments joining those vertices (which are called the *sides* of the triangle). Thus a typical triangle is pictured in Figure 1, where its vertices are labeled with capital letters. We often refer to the sides of the triangle as AB (or BA), BC and AC. The triangle in Figure 11.1 might be denoted $\triangle ABC$.

Definition 11.1.1. Two triangles are *congruent*, denoted $\cong$, if their vertices can be paired so that the corresponding angles and sides are equal to each other. That is, $\triangle ABC \cong \triangle DEF$ if $\angle A = \angle D$, $\angle B = \angle E$, $\angle C = \angle F$, and $AB = DE$, $BC = EF$, and $AC = DF$.

If two triangles are congruent, then one can be placed on top of the other so that they completely coincide. More generally, two geometric figures are said

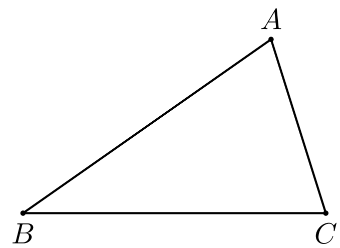


Figure 11.1

to be *congruent* to each other if they can be so placed. It is important to note that congruence of triangles can be established without verifying that all of the pairs of corresponding angles and all the pairs of corresponding sides are equal to each other; equality of some of those pairs implies equality of all of them.

For example, suppose that we fix a side of a triangle and an angle with vertex one of the endpoints of that side, and the length of the next side. That is, for example, suppose in Figure11.2 we fix the angle B and lengths AB and BC. It

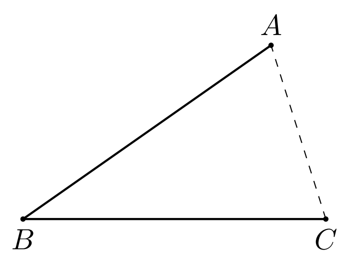


Figure 11.2

seems intuitively clear that any two triangles with the specified sides AB and BC and the angle B between them are congruent to each other; the only way to complete the given data to form a triangle is by joining A to C by a line segment. Thus it appears that any two triangles that have two pairs of equal sides and have equals angles formed by those sides are congruent to each other. We state this as a fundamental axiom.

Definition 11.1.2. (*The Congruence Axiom*, or *Side-Angle-Side*) If two triangles have two pairs of corresponding sides equal and also have equal angles between those two sides, then the triangles are congruent to each other.

We think of this axiom as stating that triangles are congruent if they have "side-angle-side" in common.

Definition 11.1.3. A triangle is said to be *isosceles* if two of its sides have the same length. The angles opposite the equal sides of an isosceles triangle are called the *base angles* of the triangle.

Theorem 11.1.4. *The base angles of an isosceles triangle are equal.*

Proof. Let the given triangle be $\triangle ABC$ with $AB = AC$. Turn the triangle over and denote the corresponding triangle as $\triangle A'C'B'$, as shown in Figure 11.3 . Then $\triangle ABC \cong \triangle A'C'B'$ since $\angle A = \angle A'$ and $AB = A'C' = AC = A'B'$, and thus they have side-angle-side in common. In this congruence $\angle B$ corresponds to $\angle C'$ and $\angle C$ to $\angle B'$, so $\angle B = \angle C'$ and $\angle C = \angle B'$. On the other hand, $\angle C$

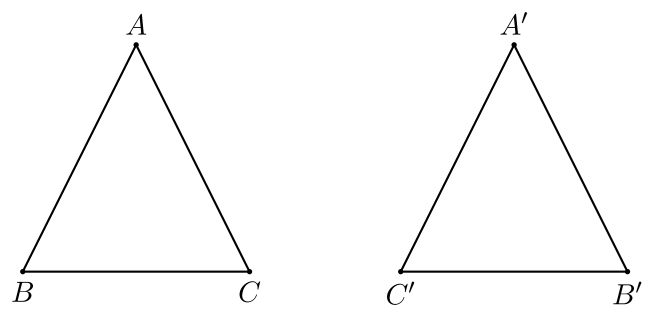


Figure 11.3

was obtained by turning $\angle C$ over, and so $\angle C' = \angle C$. It follows that $\angle B = \angle C$, as was to be proven. $\square$

Definition 11.1.5. A triangle is *equilateral* if all three of its sides have the same length.

Corollary 11.1.6. *All three angles of an equilateral triangle are equal to each other.*

Proof. Any two angles of an equilateral triangle are the base angles of an isosceles triangle, and are therefore equal to each other by the previous theorem. It follows that all three angles are equal. $\square$

It is sometimes convenient to establish congruence of triangles by correspondences other than side-angle-side.

Theorem 11.1.7. *If two triangles have "angle-side-angle" in common, then they are congruent.*

Proof. Suppose that triangles ABC and DEF are given with $\angle A = \angle D$, $AB = DE$ and $\angle E = \angle B$. If $AC = DF$, then the triangles are congruent by side-angle-side. If this is not the case, then one of them is longer; suppose, without loss of generality, that AC is shorter than DF. We will show that is impossible.

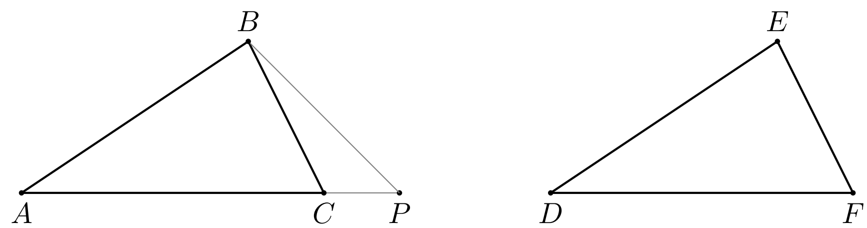


Figure 11.4

Mark the length DF along AC beginning with the point A and ending at a point P, as shown in Figure11.4. Then draw the line connecting B to P. It would follow that $\triangle ABP$ has side-angle-side in common with $\triangle DEF$. This would imply that $\angle ABP = \angle E$. But we are assuming that $\angle ABC = \angle E$. This would give $\angle ABC = \angle ABP$, from which we conclude that $\angle PBC = 0$, so PB lies on BC and hence $AP = AC$. □

If two triangles have equal sides then they automatically also have equal angles.

Theorem 11.1.8. *(Side-Side-Side) If two triangles have corresponding sides equal to each other, then they are congruent.*

Proof. Let triangles ABC and DEF be given with $AB = DE$, $BC = EF$ and $AC = DF$. At least one of the sides is greater than or equal to each of the other two; suppose, for example (the other cases would be proven in exactly the same way) that AB is greater than or equal to each of AC and CB. Then place the triangle DEF under $\triangle ABC$ so that DE coincides with AB as in Figure 11.5. Connect the points C and F by a straight line. Since $AC = DF$, triangle AFC is isosceles and the base angles ACF and AFC are equal to each other (by Theorem 11.1.4). Similarly, the triangle BCF is isosceles, so $\angle BCF = \angle BFC$. Adding the angles shows that $\angle ACB = \angle DFE$. It follows that triangles ABC and DEF agree in side angle side, and are therefore congruent to each other. □

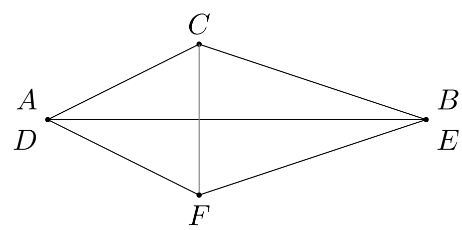


Figure 11.5

Definition 11.1.9. A *straight angle* is an angle that is a straight line. That is, the angle ABC is a straight angle if the points A, B and C all lie on a straight line and B is in between A and C. A *right angle* is an angle that is half the size of a straight angle.

Definition 11.1.10. *Vertical angles* are pairs of angles that occur opposite each other when two lines intersect. In Figure 11.6, the angles BEA and CED are a pair of vertical angles, and the angles BED and CEA are a pair of vertical angles.

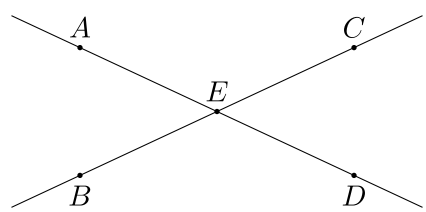


Figure 11.6

Theorem 11.1.11. *Vertical angles are equal.*

Proof. In Figure 11.6, we show that $\angle BEA = \angle CED$, as follows. Angle BEA and angle AEC add up to a straight angle. Angle AEC and angle CED also add up to a straight angle. Hence angle BEA equals angle CED. $\square$

One customary way of denoting the size of angles is in terms of degrees.

Definition 11.1.12. The measure of an angle in degrees is defined so that a straight angle is 180 degrees and other angles are the number of degrees determined by the proportion that they are of straight angles. In particular, a right angle is 90 degrees.

We will prove that the sum of the angles of a triangle is a straight angle. In the approach that we follow, the following partial result is essential.

Theorem 11.1.13. *The sum of any two angles of a triangle is less than 180 degrees.*

Proof. Consider an arbitrary triangle ABC as depicted in Figure 11.7, and extend the side AB beyond A as shown. We will show that the sum of angles CAB and ACB is less than a straight angle.

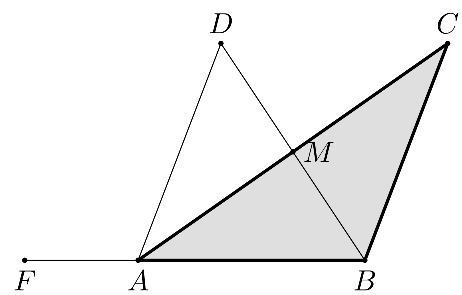


Figure 11.7

Let M be the midpoint of the side AC. Draw the line from B through M and extend it to the other side of M to a point D such that $DM = MB$. Draw the line from D to A. Then $\angle DMA = \angle CMB$ (Theorem 11.1.11). By construction, $AM = MC$ and $DM = MB$. Thus $\triangle CMB \cong \triangle AMD$ by side-angle-side (11.1.2). It follows that $\angle DAM$ is equal to $\angle BCM$. Thus the sum that we are interested in, $\angle BCM + \angle MAB$, is equal to the sum of $\angle DAM + \angle MAB$. But this latter sum is less than a straight angle since it, together with $\angle DAF$, sums to a straight angle. □

11.2 The Parallel Postulate

By a *line* we mean a straight line extending infinitely in both directions; by a *line segment*, we mean a finite part of a line between two given points. Two lines are *parallel* if they lie in the same plane and do not intersect.

For hundreds of years, mathematicians tried to prove the following as a theorem that followed from the other basic assumptions about Euclidean geometry. Finally, in the 1880s, this was shown to be impossible when other geometries were constructed that satisfied the other basic assumptions but not the following (they are now called "non-Euclidean geometries"). Since we cannot prove it, we assume it as an axiom.

Definition 11.2.1. (*The Parallel Postulate*) Given a line and a point that is not on the line, there is one and only one line through the given point that is parallel to the given line.

We will develop a necessary and sufficient condition that two lines be parallel. Given two lines, a third line that intersects both of the first two is said be a *transversal* of the two lines. Given a transversal of two lines, a pair of angles created by the intersections of the transversal with the lines are said to be *corresponding angles* if they lie on the same sides of the given lines. In Figure 11.8, T is a transversal of the lines L_1 and L_2. The angles a and c are a pair of corresponding angles, as are the angles b and d. The four angles between the parallel lines are called *interior angles.* If two interior angles lie on opposite sides of the transversal, they are called *alternate interior angles.* In Figure 11.8, the angles b and e are a pair of alternate interior angles, as are the angles c and f.

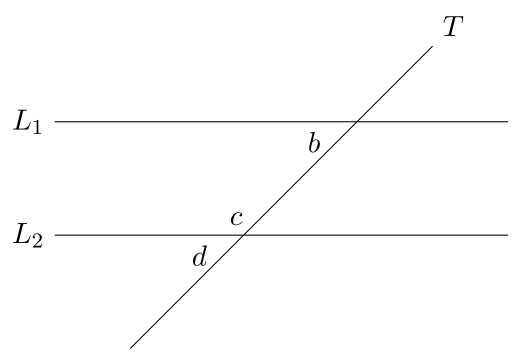


Figure 11.8

Theorem 11.2.2. *If the angles in a pair of corresponding angles created by a transversal of two lines are equal to each other, then the given lines are parallel.*

Proof. If the theorem were not true, then there would be a situation as depicted in Figure 11.9, where $\angle a = \angle c$ but lines L_1 and L_2 intersect in some point S.

Now $\angle a + \angle b$ is clearly a straight angle. Then, since $\angle a = \angle c$, it would follow that that the sum of angles b and c is a straight angle, contradicting Theorem 11.1.13. Hence the lines L_1 and L_2 cannot intersect. $\square$

The converse of this theorem is also true.

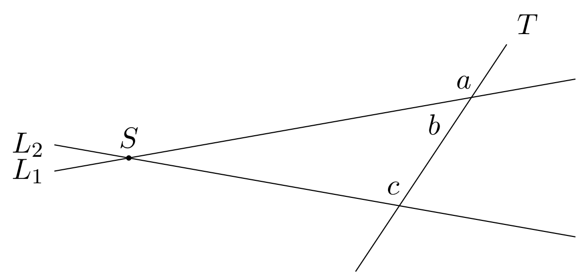


Figure 11.9

Theorem 11.2.3. *If two lines are parallel, then any pair of corresponding angles are equal to each other.*

Proof. Suppose that two lines are parallel and that two corresponding angles differ from each other. Then there would be a situation such as that depicted in Figure 11.8 with two parallel lines L_1 and L_2 and angle b different from angle d. Suppose that angle b is bigger than angle d (the proof where this inequality is reversed would be virtually identical). Then we could draw a line L through Q, the intersection point of L_1 and T, as depicted below in Figure 11.10, such that angle b' is equal to angle d.

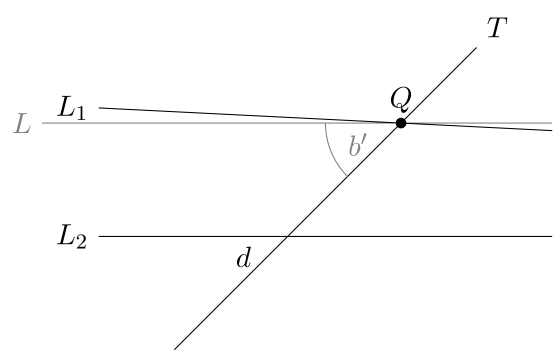


Figure 11.10

But then, by the previous theorem (11.2.2), L would be parallel to L_2. Thus L and L_1 would be distinct lines through the point Q both of which are parallel to L_2, contradicting the uniqueness aspect of the Parallel Postulate (11.2.1). $\square$

Corollary 11.2.4. *If two lines are parallel, then any pair of alternate interior angles are equal to each other.*

Proof. Consider the alternate interior angles b and e in Figure 11.8. From Theorem 11.2.3, we know that angles b and d are equal, and by Theorem 11.1.11, angle d is equal to angle e. Therefore angles b and e are equal. □

We can now establish the fundamental theorem on the angles of a triangle.

Theorem 11.2.5. *The sum of the angles of a triangle is a straight angle.*

Proof. Let a triangle ABC be given. Use the Parallel Postulate (11.2.1) to pass a line through A that is parallel to BC and mark points D and E on that line (see Figure 11.11). By Theorem 11.2.3, $\angle DAB = \angle ABC$ and $\angle EAC = \angle ACB$.

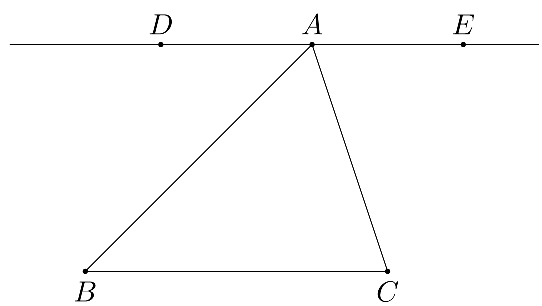


Figure 11.11

Clearly, the sum of the angles DAB, BAC, and EAC is a straight angle. Hence the sum of the angles ABC, BAC , and ACB is also a straight angle. □

The following is an obvious corollary.

Corollary 11.2.6. *If two angles of one triangle are respectively equal to two angles of another triangle, then the third angles of the triangles are also equal.*

Corollary 11.2.7. *If two triangles agree in angle-angle-side, then they are congruent.*

Proof. By the previous corollary, the triangles have their third angles equal as well. Thus the triangles also agree in angle-side-angle, and, by Theorem 11.1.7, they are congruent. □

11.3 Area of Triangles

We require knowledge of the areas of some common geometric figures. We begin with the following definition which forms the basis for the definition of areas of all geometric figures.

Definition 11.3.1. The *area of a rectangle* is defined to be the product of its length and its width.

The areas of other geometric figures can be obtained either by directly comparing them to rectangles or by approximating them by rectangles.

Definition 11.3.2. Lines, or line segments, are said to be *perpendicular* (or *orthogonal*) if they intersect in a right angle.

Definition 11.3.3. A *right triangle* is a triangle, one of whose angles is a right angle. The side opposite the right angle in a right triangle is called the *hypotenuse* of the triangle, and the other two sides are called the *legs*.

Theorem 11.3.4. *The area of a right triangle is one half the product of the legs of the triangle.*

Proof. Let the right triangle $\triangle ABC$ be as pictured in Figure 11.12.

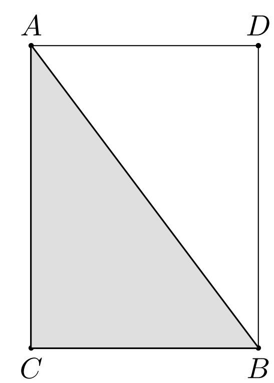


Figure 11.12

By creating perpendiculars to AC at A and to BC at B, complete the triangle to a rectangle as shown. Since the sum of the angles of a triangle is 180 degrees (11.2.5), the sum of angles BAC and ABC is 90 degrees. Since AD is perpendicular to AC, the sum of the angles BAC and BAD is also 90 degrees.

Hence, $\angle ABC = \angle BAD$. Since angle CBD is a right angle, $\angle CAB = \angle ABD$ as they both sum with angle ABC to 90 degrees. It follows that $\triangle ABC \cong \triangle ABD$ since they agree in angle-side-angle (11.1.7). Thus those triangles have equal areas. Since their areas sum to the area of the rectangle whose area is the product of the legs of the triangle ABC, it follows that the area of the triangle ABC is one half of that product. $\square$

Any one of the sides of a triangle may be regarded as a base of the triangle.

Definition 11.3.5. If a side of a triangle is designated as its *base*, then the *height* of the triangle (relative to that base) is the length of the perpendicular from the vertex of the triangle not on the base to the base. It may be necessary to extend the base of the triangle in order to determine its height, as in the second triangle pictured in Figure 11.13. (In both of the triangles depicted in Figure 11.13, h is the height of the triangle to the base AC.)

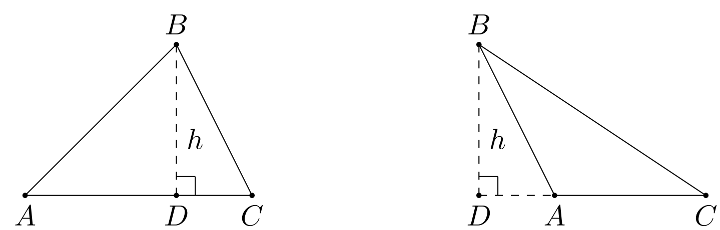


Figure 11.13

Theorem 11.3.6. *The area of any triangle is one half the product of a base of the triangle and the height of the triangle to that base.*

Proof. Suppose that the triangle ABC is as pictured in the first triangle in Figure 11.13, where h is the height to the base AC. Then, by the previous theorem (11.3.4), the area of the right triangle ABD is one half the product of h and AD, and the area of the right triangle DBC is one half the product of h and DC. The area of triangle ABC is the sum of those areas and is therefore $\frac{1}{2}h(AD) + \frac{1}{2}h(DC) = \frac{1}{2}h(AD + DC) = \frac{1}{2}h(AC)$. This finishes the proof in this case.

If the vertex from which the height is dropped has an angle in the triangle that is less than a right angle, the computation is slightly different. This situation is depicted in the second triangle in Figure 11.13. The side AC had to be extended to D as the point at the bottom of the height. In this case the area of $\triangle ABC$ is the difference between the right triangles BDC and BDA. Hence the area is $\frac{1}{2}h(DC) - \frac{1}{2}h(DA) = \frac{1}{2}h(AC)$. $\square$

One of the most famous theorems in mathematics is the Pythagorean Theorem. The easiest way to prove it is by using areas.

Theorem 11.3.7. *(The Pythagorean Theorem) For any right triangle, the square of the length of the hypotenuse is equal to the sums of the squares of the lengths of the legs.*

Proof. Let the right triangle have legs of length a and b and hypotenuse of length c. This proof of the Pythagorean theorem is obtained by placing four copies of the given right triangle inside a square whose sides have length $a + b$, as shown in Figure 11.14. We need to prove that the four sided figure $DEFG$ is a square;

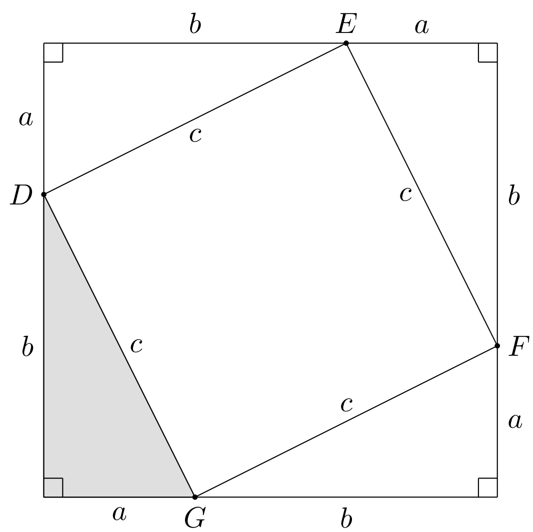


Figure 11.14

i.e., that each of its angles is a right angle. But this follows immediately from the fact that each such angle sums with the two non-right angles of the original triangle to a straight angle. Thus $DEFG$ is a square, whose side has length c. The area of the big square, each of whose sides has length $a+b$, is the sum of the area of the square $DEFG$ and four times the area of the original right triangle. That is, $(a+b)^2 = 4(\frac{1}{2}ab)+c^2$. Thus $a^2+2ab+b^2 = 2ab+c^2$ or $a^2+b^2 = c^2$. $\square$

Definition 11.3.8. Two triangles are *similar* if their vertices can be paired so that the corresponding angles are equal to each other. We use the notation $\triangle ABC \sim \triangle DEF$ to denote similarity.

Of course (by Corollary 11.2.6) it follows that two triangles are similar if they agree in two of their angles. It is an important and non-trivial fact that the corresponding sides of similar triangles are proportional to each other. In other words, if $\triangle ABC \sim \triangle DEF$, then $\frac{AB}{DE} = \frac{AC}{DF} = \frac{BC}{EF}$. The ingenious proof that we present goes back to Euclid.

We begin with a lemma.

Lemma 11.3.9. *If two lines are parallel and two other lines are perpendicular to the parallel lines, then the lengths of the segments of the other lines determined by the parallel lines are equal to each other.*

Proof. In Figure 11.15, we are assuming that L_1 is parallel to L_2 and that AB and DC are perpendicular to both of L_1 and L_2. (By Theorem 11.2.3, if a line is perpendicular to one of two parallel lines it is perpendicular to the other as well.) We must prove that $AB = CD$. Note that $\angle ACB = \angle DBC$ and

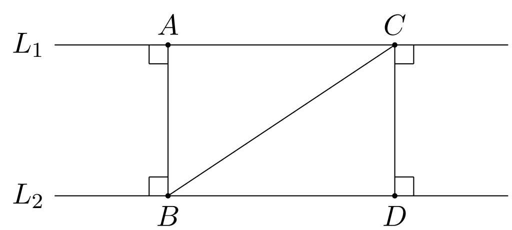


Figure 11.15

$\angle ABC = \angle BCD$, by Corollary 11.2.4. Thus the triangles ABC and BCD are congruent by angle-side-angle, since they also share the side BD. Therefore the corresponding sides AB and CD are equal to each other. $\square$

Our basic approach to the proportionality theorem is based on the following lemma.

Lemma 11.3.10. *If a triangle with area S_1 has the same height with respect to a base b_1 as a triangle with area S_2 has with respect to its base b_2, then $\frac{S_1}{b_1} = \frac{S_2}{b_2}$.*

Proof. Let the common height of the two triangles with respect to the given bases be h. Then $S_1 = \frac{1}{2}hb_1$ and $S_2 = \frac{1}{2}hb_2$. It follows that $\frac{S_1}{b_1} = \frac{1}{2}h = \frac{S_2}{b_2}$. $\square$

Theorem 11.3.11. *If two triangles are similar, then their corresponding sides are proportional. That is, if* $\triangle ABC \sim \triangle DEF$ *then* $\frac{AB}{DE} = \frac{AC}{DF} = \frac{BC}{EF}$.

Proof. It suffices to prove that $\frac{AB}{DE} = \frac{AC}{DF}$; the other equation can be obtained as in the proof below but placing the triangles so that the angle at B coincides with the angle at E.

Place the triangles so that the angle of the first triangle at A coincides with the angle of the second triangle at D. If the length of AB is the same as the length of DE, then the two triangles are congruent and all the proportions are 1. Assume, then, that the length of AB is less than the length of DE. (If the opposite is true, the proof below can be accomplished by interchanging the roles of $\triangle ABC$ and $\triangle DEF$.) The situation is depicted in Figure 11.16.

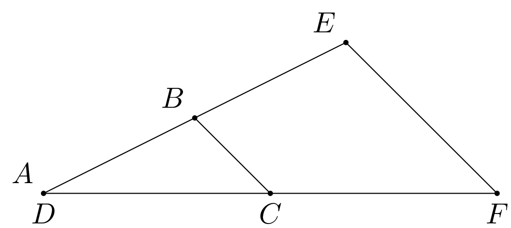


Figure 11.16

We need to construct triangles to which we can apply the preceding lemma. In Figure 11.16, connect B and F by a line and C and E by a line. Note that by Theorem 11.2.3, $\angle ABC = \angle DEF$ implies that the line BC is parallel to the line EF. Regard the triangles EBC and FBC as having a common base BC. Then the corresponding heights of the triangles are the perpendiculars from E to BC and from F to BC respectively. By lemma 11.3.9 those heights are equal to each other. Thus triangles BEC and BFC, having equal bases and heights, have equal areas. Adding those triangles to $\triangle ABC$ establishes that triangles AEC and ABF have equal areas.

We can now use Lemma 11.3.10, as follows. Since $\triangle ABC$ has the same height with respect to its base AB as $\triangle ACE$ has with respect to its base DE, Lemma 11.3.10 implies that $\frac{area(\triangle ABC)}{AB} = \frac{area(\triangle ACE)}{DE}$, or $\frac{ab}{de} = \frac{area(\triangle ABC)}{area(\triangle ACE)}$. Similarly, $\triangle ABC$ has the same height with respect to its base AC as $\triangle ABF$ has with respect to its base DF, so $\frac{area(\triangle ABC)}{AC} = \frac{area(\triangle ABF)}{DF}$, or $\frac{AC}{DE} = \frac{area(\triangle ABC)}{area(\triangle ABF)}$. Since the triangles ABF and ACE have the same area, it follows that $\frac{AB}{DE} = \frac{AC}{DF}$.

□

11.4 Problems

Basic Exercises

1. Which of the following triples cannot be the lengths of the sides of a right triangle?

   (a) 3, 4, 5

   (b) 1, 1, 1

   (c) 2, 3, 4

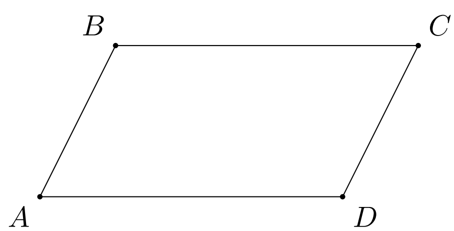
   (d) 1, $\sqrt{3}$, 2

2. Prove that two right triangles are congruent if they have equal hypothenuses and a pair of equal legs.

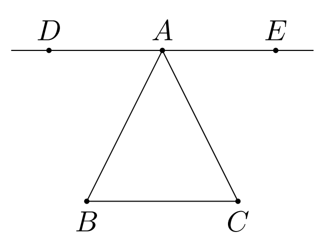
Interesting Problems

3. A *quadrilateral* is a four-sided figure in the plane. Prove that the sum of the angles of a quadrilateral is 360 degrees.

4. For quadrilateral $ABCD$ as shown below, suppose that $\angle ABC = \angle CDA$ and $\angle DAB = \angle BCD$. Prove that $AB = CD$ and $BC = AD$.



5. With reference to the diagram below, prove that $\triangle ABC$ is an isosceles triangle if $\angle DAB = \angle EAC$ and DE is parallel to BC.



6. Prove that if two angles of a triangle are equal, then the sides opposite those angles are equal.

7. A *parallelogram* is a four sided figure in the plane whose opposite sides are parallel to each other. Prove the following:

   (a) The opposite sides of a parallelogram have the same length.

   (b) The area of a parallelogram is the product of the length of any side and the length of a perpendicular to the side from a vertex not on the side.

   (c) If one of the angles of a parallelogram is a right angle, then the parallelogram is a rectangle.

8. A *trapezoid* is a four sided figure in the plane two of whose sides are parallel to each other. The *height* of a trapezoid is the length of a perpendicular from one of the parallel sides to the other. Prove that the area of a trapezoid is its height multiplied by the average of the lengths of the two parallel sides.

9. A *square* is a four sided figure in the plane all of whose sides are equal to each other and all of whose angles are right angles. The *diagonals* of the square are the lines joining opposite vertices. Prove that the diagonals of a square are perpendicular to each other.

10. Show that lines are parallel if there is a transversal such that the alternating interior angles are equal to each other.

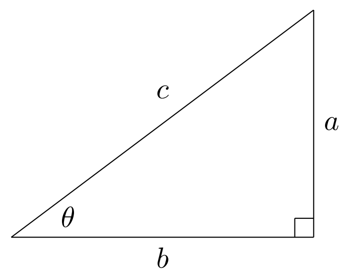
Challenging Problems

11. Give an example of two triangles that agree in angle-side-side but are not congruent to each other.

12. Prove the converse of the Pythagorean Thereom; i.e., show that if the lengths of the sides of a triangle satisfy the equation $a^2 + b^2 = c^2$, then the triangle is a right triangle.

13. The following problem provides some basic results in trigonometry:

    (a) Let θ be any angle between 0 and 90 degrees. Place θ in a right triangle as shown in the diagram below and label the sides as in the diagram. Define $\sin\theta$ to be $\frac{a}{c}$, $\cos\theta$ to be $\frac{b}{c}$ and $\tan\theta$ to be $\frac{a}{b}$. Using Theorem 11.3.11, show that these definitions do not depend on which right triangle θ is placed in.



(b) Label the angles of a triangle A, B and C and label the side opposite A with a, the side opposite B with b, and the side opposite C with c. Prove that

(i) (The law of sines) $\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C}$.

(ii) (The law of cosines) $c^2 = a^2 + b^2 - 2ab \cos C$

Chapter 12

Constructability

The Ancient Greeks were interested in many different kinds of mathematical problems. One of the aspects of geometry that they investigated is the question of what geometrical figures can be constructed using a compass and a straightedge. A compass is an instrument for drawing circles. The compass has two branches that open up like a scissors. One of the branches has a sharp point at the end and the other branch has a pen or pencil at the end. If the compass is opened so that the distance between the two ends is r and the pointed end is placed on a piece of paper and the compass is twirled about that point, the writing end traces out a circle of radius r. The drawing made by any real compass will only approximate a circle of radius r. But we are going to consider constructions theoretically; we will assume that a compass opened up a distance r precisely makes a circle of radius r.

To do geometrical constructions, we will also require (as the Ancient Greeks did) another implement. By a *straightedge* we will mean a device for drawing lines connecting two points and extending such lines as far as desired in either direction. Sometimes people inaccurately speak of constructions with “ruler and compass”. It is important to understand that the constructions investigated by the Ancient Greeks do not allow use of a ruler in the sense of an instrument that has distances marked on it. We can only use such an instrument to connect pairs of points by straight lines.

In this chapter, when we say “construct” or “construction” we always mean “using only a compass and a straightedge”.

We will indicate how to do some basic constructions. But the most interesting part of this chapter will be proving that certain geometrical objects cannot be constructed. In particular, we will prove that an angle of 20 degrees cannot be constructed. This implies that an angle of 60 degrees cannot be trisected (that is, divided into three equal parts) with a straightedge and compass. The Ancient Greeks assumed that there must be some way of trisecting every angle; they thought that they had simply not been clever enough to find a method for doing so. It was only after mathematical advances that took place in the 19th century that it could be proven that there is no way to trisect an angle of 60 degrees with a straightedge and compass. The highlight of this chapter will be a proof of that fact. Although it is hard to imagine how something like that could be proved, we shall see that there is an indirect approach that also establishes many other interesting results.

12.1 Constructions With Straightedge and Compass

Let's start with some very basic constructions.

Definition 12.1.1. By a *line segment*, we mean a straight line of a finite length. A *perpendicular bisector* of a line segment is a line that is perpendicular to the line segment and goes through the middle of the line segment.

Theorem 12.1.2. *Given any line segment, it's perpendicular bisector can be constructed.*

Proof. Given a line segment AB as shown in Figure 12.1, put the point of the compass at A and open the compass to radius the length of AB. Let r equal the length of AB. Then draw the circle with center at A and radius r. Similarly, draw the circle with center at B and radius r. The two circles will intersect at points such as C and D as indicated in Figure 12.1. Take the straightedge and draw the line segment from C to D. We claim that CD is a perpendicular bisector of AB.

To prove this, label the point of intersection of CD and AB as E and then draw the line segments AC, CB, BD, and DA. We must prove that $AE = EB$ and that $\angle CEA$ (and/or any of the other three angles at E) is a right angle. First note that AC, CB, BD, and AD all have the same length, r, since they are all radii of the two circles of radius r. Thus triangle ACD is congruent to triangle DCB, since the third side of each is CD and they therefore agree in side-side-side (11.1.8). It follows that $\angle ACE = \angle BCE$. Thus triangle ACE is congruent to triangle CEB (by side-angle-side, 11.1.2). Therefore $AE = EB$. Moreover, $\angle AEC = \angle BEC$ so, since those two angles sum to a straight angle, each of them is a right angle.

$\square$

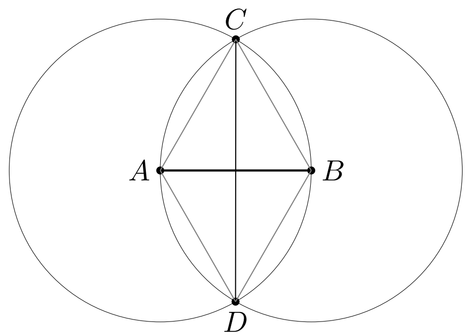


Figure 12.1

Definition 12.1.3. A *bisector of an angle* is a line from the vertex of the angle that divides it into two equal sub-angles.

Theorem 12.1.4. *Given any angle, its bisector can be constructed.*

Proof. Consider an angle ABC as pictured in Figure 12.2 and draw any circle centered at B. Label the points of intersection of the circle with AB and with BC as E and F, respectively. Let r be the distance from E to F. Use the compass to draw a circle of radius r centered at E and a circle of radius r centered at F. These two circles intersect in some point G within the angle ABC, as shown in figure 2. Use the straightedge to draw the line segment connecting B to G. We claim that this line segment bisects the angle ABC.

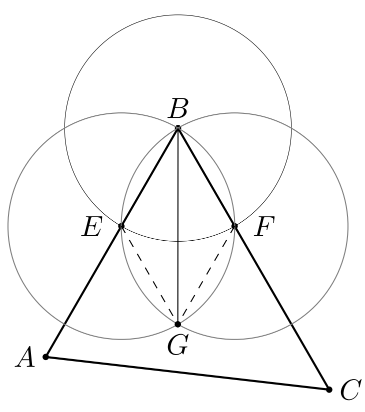


Figure 12.2

To prove this, draw the lines EG and FG. We prove that triangle BEG is congruent to triangle BFG. Note that $BE = BF$, since they are both radii of the original circle centered at B. Note also that $EG = FG$, since they are each radii of circles with radius r. Since $BG = BG$, it follows (11.1.8) that triangle BEG is congruent to triangle BFG. Thus $\angle EBG = \angle GBF$ and BG a bisector of angle ABC. □

Theorem 12.1.5. *Given any line segment, the line segment can be copied anywhere using only a straightedge and compass.*

Proof. Suppose a line segment AB is given as pictured in Figure 12.3 and it is desired to copy it on another line. Choose any point C on the other line, then open the compass to a radius the length of AB. Put the point of the compass at C and draw any portion of the resulting circle that intersects the other line. Label the point of intersection D.

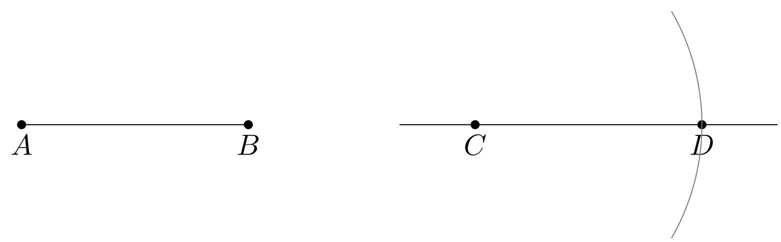


Figure 12.3

Then CD is copy of AB. □

Theorem 12.1.6. *Given any angle, the angle can be copied anywhere using only a straightedge and compass.*

Proof. Let an angle ABC be given as in Figure 12.4. We construct an angle equal to $\angle ABC$ with vertex G on any other line. To do this, draw any arc of any circle (of radius, say, r) centered at B that intersects both BA and BC. Label the points of intersection D and E. Draw the circle of radius r centered at G. Use H to label the point where that circle intersects the line containing G. Then adjust the compass to make circles of radius DE. Put the point of the compass at H and draw a portion of the circle that intersects the circle centered at G; call that point of intersection I. Draw line segments connecting D to E and I to H.

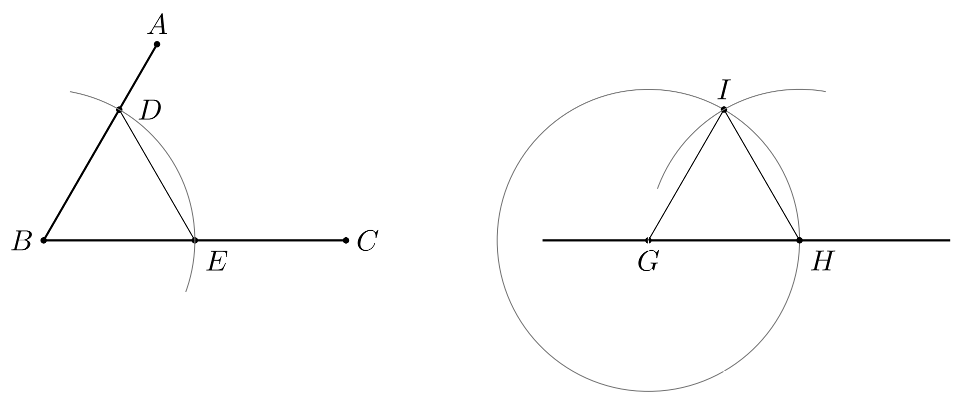


Figure 12.4

Then $DE = IH$, since they are each radii of circles with the same radius. Also draw the line segment GI. The lengths of BD, BE, GI, and GH are all equal to r. It follows (by side-side-side, 11.1.8) that $\triangle BDE \cong \triangle GIH$. Thus $\angle IGH$ is a copy of $\angle ABC$. $\square$

Corollary 12.1.7. *If the angles α and β are constructed, then*

(i) the angle $\alpha + \beta$ can be constructed, and

(ii) for every natural number n, the angle $n\alpha$ can be constructed.

Proof. (i) Let the angles α and β be given, as pictured in Figure 12.5. To construct the angle $\alpha + \beta$, simply copy the angle ABC with one side DE and the other side outside the original angle β as shown in the third diagram in Figure 12.5.

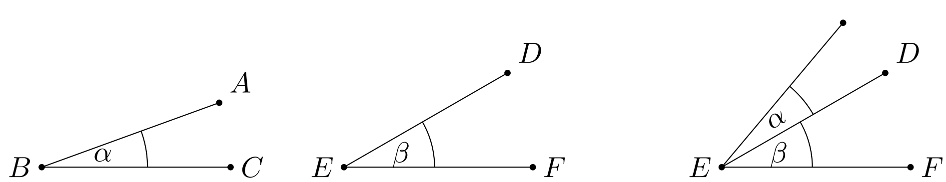


Figure 12.5

(ii) This clearly follows from repeated application of part i), starting with angles α and β that are equal to each other. (This can be proven more formally using mathematical induction.) $\square$

Theorem 12.1.8. *Given any line segment and any natural number n, the line segment can be divided into n equal parts using only a straightedge and compass.*

Proof. Fix a natural number n. Let a line segment AB be given as shown in Figure 12.6. Use the straightedge to draw any line segment AC emanating from A, as shown. Open the compass to any radius s less than one nth of the length of AC. Beginning at A use the compass to mark off n consecutive segments of AC of length s, as illustrated in figure 6. Label the points of intersection of the arcs and AC as $P_1, P_2, P_3, \ldots, P_{n-1}$. Label with D the point of intersection of the line and the last arc drawn. Use a straightedge to connect D to B. We then construct lines parallel to DB through each point of intersection of an arc with AD, after which we will show that the intersections of those lines with AB divide AB into n equal segments.

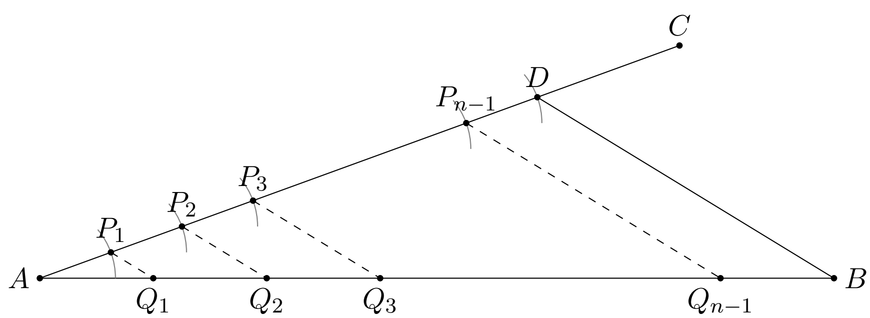


Figure 12.6

To construct the parallel lines, copy the angle ADB (12.1.6) at each point of intersection of an arc with AD so that one side of the new angle lies on AD, and the other side points downward and is extended to intersect the line AB. These are the dotted lines in Figure 12.6. Label the points of intersection of the dotted lines with AB as $Q_1, Q_2, Q_3, \ldots, Q_{n-1}$, as shown. We claim that the points $\{Q_1, Q_2, Q_3, \ldots, Q_{n-1}\}$ divide the segment AB into n equal parts. To see this, note that, for each j, the triangle AP_jQ_j has two angles ($\angle P_jAQ_j$ and $\angle AP_jQ_j$) equal to corresponding angles of $\triangle ADB$. Thus $\triangle AP_jQ_j$ is similar to $\triangle ADB$ (11.2.6). Therefore, the corresponding sides are proportional (11.3.11). The ratio of AP_j to AD is $\frac{j}{n}$. Thus the length of AQ_j divided by the length of AB is also $\frac{j}{n}$. $\square$

Even though a line segment can be divided into any number of equal parts, some angles, such as 60 degrees, cannot be divided into three equal parts using only a straightedge and compass. We now begin preparation for an indirect approach to establishing that theorem.

12.2 Constructible numbers

We consider constructing numbers instead of constructing geometric objects, although we will use geometric constructions to construct the numbers.

We begin by imagining a line on which a point is arbitrarily marked as 0 and another point, to the right of it, is arbitrarily marked as 1. We consider the question of what other numbers can be obtained by starting with the length 1 (that we take as the distance between the points marked 0 and 1) and doing geometrical constructions in the plane to obtain other lengths.

Definition 12.2.1. A real number is *constructible* if the point corresponding to it on the number line can be obtained from the marked points 0 and 1 by performing a finite sequence of constructions using only a straightedge and compass.

Theorem 12.2.2. *Every integer is constructible.*

Proof. The numbers 0 and 1 are given as constructible. The number 2 can easily be constructed: simply take a compass, open it up to radius 1 by placing one side at the point 0 and the other side at the point 1, and then place the pointed side on the point marked 1 and draw the circle of radius 1 with that point as center. The point where that circle meets the number line to the right of 1 is the number 2, so 2 has been constructed. Then clearly 3 can be constructed by placing the compass with radius 1 so as to make a circle centered at 2. Similarly, all the natural numbers can be constructed. To construct the number -1, simply make the circle of radius 1 centered at 0 and mark the intersection to the left of 0 of that circle with the number line. Then -2 can be constructed by marking the point with the circle centered at -1 meets the number line to the left of the point -1. Every negative integer can be constructed similarly. □

What about the rational numbers?

Theorem 12.2.3. *Every rational number is constructible.*

Proof. To construct, for example, the number $\frac{1}{3}$, simply divide the interval between 0 and 1 into three equal parts (see 12.1.8) and mark the right-most point of the first part as $\frac{1}{3}$. Similarly, for any natural number n, dividing the unit interval in n equal parts shows that $\frac{1}{n}$ is constructible. Then, for any natural number m, $\frac{m}{n}$ can be constructed by placing m segments of length $\frac{1}{n}$ next to each other on the number line with the first of those segments beginning at 0.

We have therefore shown that all of the positive rational numbers are constructible. If x is a negative rational number, construct $|x|$ and then make a circle of radius $|x|$ centered at 0; the point to the left of 0 where that circle intersects the number line is x. Thus every rational number is constructible. □

We need to get information about the set of all constructible numbers.

It is essential to the development of this approach that doing arithmetic with constructible numbers produces constructible numbers.

Theorem 12.2.4. *If a is constructible, then $-a$ is constructible.*

Proof. Place a compass on the number line with its point at 0 and the other end opened to a. Then draw the circle. The number $-a$ will be the point of intersection of the circle and the number line opposite to that of a. (If a is positive, then $-a$ is negative, but if a is negative, then $-a$ is positive.) □

Theorem 12.2.5. *The sum of two constructible numbers is constructible.*

Proof. Suppose that a and b are constructible. There are several cases, depending upon the signs of a and b. If both of a and b are non-negative, open the compass to length b, place the point on the number line at a, and make the circle of radius b. The point to the right of a at which the circle intersects the number line is $a + b$.

If a is non-negative and b is negative, open the compass to length $|b|$, place the point of the compass at a and draw the circle with radius $|b|$. The point $a+b$ will be the point of intersection of the circle and the number line that is to the left of a. If b is non-negative and a is negative, simply interchange the roles of a and b in this construction.

The remaining case is where both of a and b are negative. In this case, open the compass to radius $|b|$, place the point of the compass at a and draw the circle. The point $a + b$ is the point of intersection of the circle and the number line that lies to the left of a. □

We also need to construct products and quotients. These constructions are a little more complicated; we begin with the following.

Theorem 12.2.6. *If a and b are positive constructible numbers, then $\frac{a}{b}$ is constructible.*

Proof. We consider the two different cases, where b is greater than 1 and b is less than 1, respectively.

For the case where b is greater than 1, draw the numbers 0, 1 and b on the number line. Use the straightedge to draw a line of length greater than a starting from 0, making any angle greater than 0 and less than 90 degrees with the number line. Since a is constructible, we can open the compass to radius a. Place the point of the compass at 0 and mark a on the line above the number line. Use the straightedge to connect the point a on the new line to the point b on the number line. Copy the angle, β, at b to the point 1 on the number line so that the lower side of the angle is the number line itself. Use the straightedge to extend the other side of the angle beyond the new line. The intersection of the other side of the angle and the new line is a point that we have thereby constructed. Let the distance from the origin to that point be x. We can open the compass to radius x and thereby mark x on the number line. So x is a constructible number. The relationship between x and a and

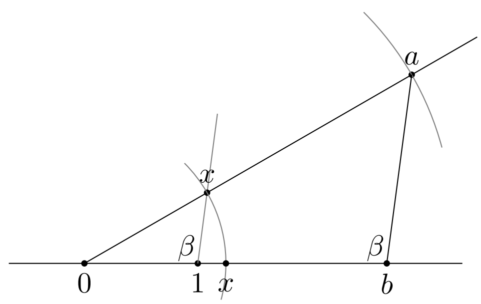


Figure 12.7

b can be determined by observing that the two triangles formed by the above construction are similar to each other, and therefore the corresponding sides are in proportion (11.3.11). It follows that $\frac{x}{a} = \frac{1}{b}$. Thus $x = \frac{a}{b}$, so we have constructed $\frac{a}{b}$.

The case where b is less than 1 is very similar. In this case, 1 is to the right of b on the number line. Use the straightedge to make a side of an angle starting at 0 above the number line. Since a is constructible, we can open the compass to radius a and mark a point on the new line that is distance a from the vertex of the angle. Then use the straightedge to draw a straight line between that point and the point b on the number line. Copy the angle, β, at the point b on the number line to the point 1 on the number line and extend the side of the angle so that it intersects the other line. The compass can then be opened to radius equal to the distance from that point of intersection to the origin. If x denotes that radius, then the fact that the corresponding sides of similar triangles are proportional gives $\frac{a}{x} = \frac{b}{1}$, so that $x = \frac{a}{b}$. Thus $\frac{a}{b}$ is constructible. $\square$

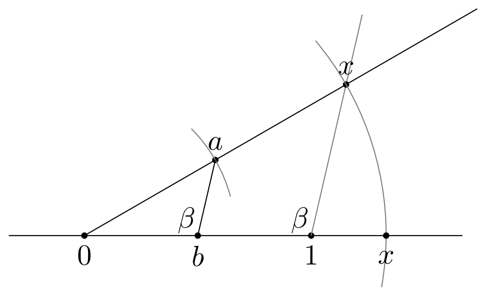


Figure 12.8

It is easy to extend the above to negative numbers.

Corollary 12.2.7. *If a and b are constructible numbers, then ab is constructible and, if $b \neq 0$, $\frac{a}{b}$ is constructible.*

Proof. First suppose that a and b are both positive. Then $\frac{a}{b}$ is constructible by the previous theorem. Let $c = \frac{1}{b}$; then c is constructible by the previous theorem using $a = 1$. Since c is constructible, the previous theorem implies that $\frac{a}{c}$ is constructible. But $\frac{a}{c} = \frac{a}{(\frac{1}{b})} = ab$, so ab is constructible.

If one or both of a and b is negative, the above can be applied to $|a|$ and $|b|$. Then $ab = |a| \cdot |b|$ if a and b are both negative, and $ab = -(|a| \cdot |b|)$ if exactly one of them is negative. Similarly, $\frac{a}{b}$ is equal to one of $\frac{|a|}{|b|}$ or $-(\frac{|a|}{|b|})$. Since we can construct the negative of any constructible number, it follows that ab and $\frac{a}{b}$ are constructible in this case as well. $\square$

A "field" is an abstract mathematical concept. In this book we do not need to consider general fields; we only need to consider fields of real numbers. The following definition forms the basis for the rest of this chapter.

Definition 12.2.8. A *number field* is a set $\mathcal{F}$ of real numbers satisfying the following properties:

i) The numbers 0 and 1 are both in $\mathcal{F}$.

ii) If x and y are in $\mathcal{F}$ then $x + y$ and $x \cdot y$ are in $\mathcal{F}$ (i.e., $\mathcal{F}$ is "closed under addition" and "closed under multiplication").

iii) If x is in $\mathcal{F}$, then $-x$ is in $\mathcal{F}$.

iv) If x is in $\mathcal{F}$ and $x \neq 0$, then $\frac{1}{x}$ is in $\mathcal{F}$.

There are many different number fields. Of course, $\mathbb{R}$ itself is a number field. So is the set $\mathbb{Q}$ of rational numbers. It is clear that $\mathbb{R}$ is the biggest number field; it is almost as obvious that $\mathbb{Q}$ is the smallest number field, in the following sense.

Theorem 12.2.9. *If $\mathcal{F}$ is any number field, then $\mathcal{F}$ contains all rational numbers.*

Proof. To see this, first note that $0 \in \mathcal{F}$, $1 \in \mathcal{F}$ and property ii of a number field implies that 2 is in $\mathcal{F}$, and 3 is in $\mathcal{F}$, and so on. That is, $\mathcal{F}$ contains all the natural numbers. Property iii then implies that $\mathcal{F}$ contains all integers. By property iv, $\mathcal{F}$ contains the reciprocals of every integer other than 0, so by property ii, $\mathcal{F}$ contains all rational numbers. □

Definition 12.2.10. The set of all constructible numbers will be denoted $\mathcal{C}$.

The following is an important fact.

Theorem 12.2.11. *The set $\mathcal{C}$ of constructible numbers is a number field.*

Proof. This follows immediately from Theorems 12.2.4, 12.2.5, and 12.2.6 and Corollary 12.2.7. □

One of the fundamental theorems in this chapter will provide an alternate characterization of the field $\mathcal{C}$ of constructible numbers (12.3.12).

Example 12.2.12. The set $\mathbb{Q}(\sqrt{2})$ defined by

$$\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a \in \mathbb{Q}, b \in \mathbb{Q}\}$$

is a number field.

Proof. It is clear that $\mathbb{Q}(\sqrt{2})$ contains 0 $(= 0 + 0 \cdot \sqrt{2})$ and 1 $(= 1 + 0 \cdot \sqrt{2})$. Moreover,

$$(a_1 + b_1\sqrt{2}) + (a_2 + b_2\sqrt{2}) = (a_1 + a_2) + (b_1 + b_2)\sqrt{2}$$

hence $\mathbb{Q}(\sqrt{2})$ is closed under addition. Also,

$$(a_1 + b_1\sqrt{2})(a_2 + b_2\sqrt{2}) = (a_1a_2 + 2b_1b_2) + (a_1b_2 + a_2b_1)\sqrt{2}$$

so $\mathbb{Q}(\sqrt{2})$ is closed under multiplication. Futhermore, $-(a+b\sqrt{2}) = (-a) + (-b)\sqrt{2}$.

It remains to be shown that $\frac{1}{a+b\sqrt{2}}$ is in $\mathbb{Q}(\sqrt{2})$ whenever a and b are not both 0. But

$$\frac{1}{a+b\sqrt{2}} = \frac{a-b\sqrt{2}}{(a+b\sqrt{2})(a-b\sqrt{2})} = \frac{a-b\sqrt{2}}{a^2-2b^2} = \frac{a}{a^2-2b^2} + \frac{-b}{a^2-2b^2}\sqrt{2}$$

which is the sum of a rational number and a number that is the product of rational number and $\sqrt{2}$, and is therefore in $\mathbb{Q}(\sqrt{2})$. (Of course, the above expression would not make sense if $a^2-2b^2=0$. However, this cannot be the case, since $a^2-2b^2=0$ would imply $(\frac{a}{b})^2=2$, and we know that $\sqrt{2}$ is irrational (8.2.5.)) □

The number field $\mathbb{Q}(\sqrt{2})$ is the field obtained by starting with the field $\mathbb{Q}$ and "adjoining $\sqrt{2}$" to $\mathbb{Q}$; it is called "the extension of $\mathbb{Q}$ by $\sqrt{2}$". This is a special case of a much more general situation.

Theorem 12.2.13. *Let $\mathcal{F}$ be any number field and suppose that r is a positive number in $\mathcal{F}$. If $\sqrt{r}$ is not in $\mathcal{F}$, and*

$$\mathcal{F}(\sqrt{r}) = \{a+b\sqrt{r} : a \in \mathcal{F}, b \in \mathcal{F}\},$$

then $\mathcal{F}(\sqrt{r})$ is a number field.

Proof. The proof is very similar to the proof given above for the special case of $\mathbb{Q}(\sqrt{2})$. It is very easily seen that 0 and 1 are in $\mathcal{F}(\sqrt{r})$ and that $\mathcal{F}(\sqrt{r})$ is closed under addition. To see that it is closed under multiplication, note that

$$(a_1+b_1\sqrt{r})(a_2+b_2\sqrt{r}) = (a_1a_2+rb_1b_2)+(a_1b_2+a_2b_1)\sqrt{r}$$

This is in $\mathcal{F}(\sqrt{r})$ since r is in $\mathcal{F}$ and $\mathcal{F}$ itself is a number field. Also,

$$\frac{1}{a+b\sqrt{r}} = \frac{a-b\sqrt{r}}{(a+b\sqrt{r})(a-b\sqrt{r})} = \frac{a-b\sqrt{r}}{a^2-rb^2} = \frac{a}{a^2-rb^2} + \frac{-b}{a^2-rb^2}\sqrt{r}$$

Note that $a^2-rb^2 \neq 0$ unless a and b are both 0, because $\sqrt{r} \notin \mathcal{F}$ ($a^2-rb^2=0$ and $b \neq 0$ then $(\frac{a}{b})^2 = r$, and it would follow that $\sqrt{r} \in \mathcal{F}$ □

Definition 12.2.14. If $\mathcal{F}$ is a number field and r is a positive number that is in $\mathcal{F}$ such that $\sqrt{r}$ is not in $\mathcal{F}$, then the number field

$$\mathcal{F}(\sqrt{r}) = \{a+b\sqrt{r} : a \in \mathcal{F}, b \in \mathcal{F}\}$$

is the number field obtained by adjoining $\sqrt{r}$ to $\mathcal{F}$ and is called *the extension of $\mathcal{F}$ by $\sqrt{r}$.*

Example 12.2.15. The extension of $\mathbb{Q}(\sqrt{2})$ by $\sqrt{5}$ is

$$\{a+b\sqrt{5} : a \in \mathbb{Q}(\sqrt{2}), b \in \mathbb{Q}(\sqrt{2})\} = \{(c+d\sqrt{2})+(e+f\sqrt{2})\sqrt{5} : c,d,e,f \in \mathbb{Q}\}$$

For present purposes, we are interested in adjoining square roots to number fields because that can be done "constructibly".

Theorem 12.2.16. *If r is a positive constructible number, then $\sqrt{r}$ is constructible.*

Proof. Mark the number $1+r$, labeled A, on the number line. Let $M = \frac{r+1}{2}$; m is constructible. Make a circle with centre M and radius M. The circle then goes through the point A and also the point corresponding to 0, which we label O. Use D to denote the point corresponding to r on the number line. Erect a perpendicular to the number line at D and let C be the point above the number line at which that perpendicular intersects the circle.

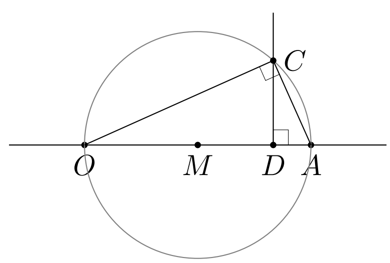


Figure 12.9

The angle OCA is 90 degrees, since it is inscribed in a semicircle. Therefore the sum of the angles OCD and DCA is 90 degrees, from which it follows that the angle COD equals the angle DCA. Thus triangle OCD is similar to triangle DCA, so their corresponding sides are proportional (11.3.11). Let x denote the length of the perpendicular from D to C. Then $\frac{x}{1} = \frac{r}{x}$, so $x^2 = r$. Hence $x = \sqrt{r}$ and $\sqrt{r}$ is constructible. $\square$

It follows immediately from this theorem (12.2.16) and the fact that the constructible numbers form a number field (Theorem 12.2.11) that every number in $\mathbb{Q}(\sqrt{2})$ is constructible. More generally, every element of $\mathbb{Q}(\sqrt{r})$ is constructible for every positive rational number r such that $\sqrt{r}$ is irrational. Even more generally, if $\mathcal{F}$ is a number field consisting of constructible numbers and r is a positive number in $\mathcal{F}$ such that $\sqrt{r}$ is not in $\mathcal{F}$, then $\mathcal{F}(\sqrt{r})$ consists of constructible numbers. Thus if we start with $\mathbb{Q}$ we can keep on adjoining square roots and get constructible numbers.

Definition 12.2.17. A *tower of number fields* is a finite sequence $\mathcal{F}_0, \mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_n$ of number fields such that $\mathcal{F}_0 = \mathbb{Q}$ and for each i from 1 to n there is a positive number r_i in $\mathcal{F}_{i-1}$ such that $\sqrt{r_i}$ is not in $\mathcal{F}_{i-1}$ and $\mathcal{F}_i = \mathcal{F}_{i-1}(\sqrt{r_i})$.

Note that a tower can be described as a sequence $\{\mathcal{F}_i\}$ of number fields such that

$$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \subset \ldots \subset \mathcal{F}_n$$

with $\mathcal{F}_0 = \mathbb{Q}$ and each $\mathcal{F}_i$ is obtained from its predecessor $\mathcal{F}_{i-1}$ by adjoining a square root.

12.3 Surds

We will show that the constructible numbers are exactly those real numbers that are in fields that are in towers. There is another name that is frequently used for such numbers.

Definition 12.3.1. A *surd* is a number that is in some number field that is in a tower. That is, x is a surd if there exists a tower

$$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \subset \ldots \subset \mathcal{F}_n$$

such that x is in $\mathcal{F}_n$.

Theorem 12.3.2. *The set of all surds is a number field. Moreover, if r is a positive surd then $\sqrt{r}$ is a surd.*

Proof. To show that the set of surds is a number field, it must be shown that the arithmetical operations applied to surds produce surds. This follows immediately if it is shown that for any surds x and y there exists a number field $\mathcal{F}$ that occurs in some tower and that contains both x and y. If $\{\sqrt{r_1}, \sqrt{r_2}, \ldots \sqrt{r_m}\}$ are the numbers adjoined in making a tower that contains x and $\{\sqrt{s_1}, \sqrt{s_2} \ldots \sqrt{s_n}\}$ are the numbers adjoined in making a tower containing y, then adjoining all of those numbers produces a number field that contains both x and y. Thus the set of surds is a number field

If r is a positive surd, then r is in some field $\mathcal{F}$ that is in a tower. If $\sqrt{r}$ is in $\mathcal{F}$ then $\sqrt{r}$ is clearly a surd. If $\sqrt{r}$ is not in $\mathcal{F}$ then $\sqrt{r}$ is in $\mathcal{F}(\sqrt{r})$, which is clearly in a tower that has one more field than the tower leading to $\mathcal{F}$. $\square$

Theorem 12.3.3. *Every surd is constructible.*

Proof. This follows immediately from the results that the rational numbers are constructible (12.2.3), that the constructible numbers form a number field (12.2.11) and that the square root of a positive constructible number is constructible (12.2.16). □

The fundamental theorem that we will need is that the constructible numbers are exactly the surds. To establish this, we must show that starting with the numbers 0 and 1 and performing constructions with straightedge and compass never produces any numbers that are not surds. Since constructions take place in the plane, we will have to investigate what points in the plane can be constructed.

Definition 12.3.4. We say that the point (x, y) in the plane is *constructible* if that point can be obtained from the points (0,0) and (1,0) by performing a sequence of constructions with straightedge and compass.

Theorem 12.3.5. *The point (x, y) is constructible if and only if both of the coordinates x and y are constructible numbers.*

Proof. If x and y are constructible numbers, then the point (x, y) can be constructed by constructing the point x on the x axis, erecting a perpendicular to the x axis at the point x and constructing y on that perpendicular.

Conversely, if the point (x, y) has been constructed, then the number x can be constructed by dropping a perpendicular from (x, y) to the x axis and the number y can be constructed by dropping a perpendicular to the y axis. □

Definition 12.3.6. The *surd plane* is the set of all points (x, y) in the x, y plane such that the coordinates x and y are both surds.

By what we have shown above, every point in the surd plane is constructible. We need to show that every constructible point is in the surd plane.

After we have constructed some points, how can we construct others? We can use our straightedge to make lines joining any two points we have constructed and we can use our compass to construct a circle centered at any point that has already been constructed and having radius that has been constructed. New points can be constructed as points of intersection of lines or circles that we have constructed.

Any one line in the plane has many different equations, as does any one circle. We need to know that there are equations with surd coefficients for lines and circles that arise in constructions.

Theorem 12.3.7. *If a line goes through two points in the surd plane, then there is an equation for that line that has surd coefficients.*

Proof. Suppose that (x_1, y_1) and (x_2, y_2) are distinct points in the surd plane. We consider two cases. If $x_1 \neq x_2$ then

$$\frac{y - y_1}{x - x_1} = \frac{y_2 - y_1}{x_2 - x_1}$$

is an equation of the line. Since the surds form a field, the coefficients in this equation are all surds.

If $x_1 = x_2$, then $x = x_1$ is an equation of the line. □

Theorem 12.3.8. *A circle whose center is in the surd plane and whose radius is a surd has an equation where the coefficients are all surds.*

Proof. Let the center be (x_1, y_1) and the radius be r. Then one equation of the circle is $(x - x_1)^2 + (y - y_1)^2 = r^2$. Expanding this equation and using the fact that surds form a number field shows that this equation has surd coefficients. □

Theorem 12.3.9. *A point of intersection of two lines which each go through two points in the surd plane is itself in the surd plane.*

Proof. By 12.3.7, each of the lines has an equation with surd coefficients. Let such equations be $a_1x + b_1y = c_1$ and $a_2x + b_2y = c_2$. If $a_1 = 0$ then $y = \frac{c_1}{b_1}$, so $a_2x + b_2\frac{c_1}{b_1} = c_2$, from which it follows that the intersection of the two lines has coordinates $x = \frac{c_2}{a_2} - \frac{b_2}{a_2}\frac{c_1}{b_1}$ and $y = \frac{c_1}{b_1}$, both of which are surds.

If $a_1 \neq 0$, then $x = -\frac{b_1}{a_1}y + \frac{c_1}{a_1}$. Substituting this in the second equation yields $a_2(-\frac{b_1}{a_1}y + \frac{c_1}{a_1}) + b_2y = c_2$. Since all the coefficients are surds, it is clear that y is also a surd. Hence so is x and the theorem is proven in this case as well. □

We next consider the points of intersection of a line and a circle.

Theorem 12.3.10. *The points of intersection of a line that has an equation with surd coefficients and a circle that has an equation with surd coefficients lie in the surd plane.*

Proof. Consider a line with equation $ax + by = c$ and a circle with equation $(x - f)^2 + (y - g)^2 = r^2$ where all of the coefficients are surds. Consider first the case where $a = 0$. In this case, $y = \frac{c}{b}$. Substituting this in the equation of the circle yields $(x-f)^2+(\frac{c}{b}-g)^2=r^2$. This is a quadratic equation in x. It has 0, 1, or 2 real number solutions depending upon whether the line does not intersect the circle, is tangent to the circle, or intersects the circle in two points. The quadratic formula shows that solutions that exist are obtained from the coefficients by the ordinary arithmetic operations and the extracting of a square root. All of these operations on surds produce surds. Thus any solutions x are surds, proving the theorem in this case.

If $a \neq 0$, $x=-\frac{b}{a}y+\frac{c}{a}$. Substituting this value in the equation of the circle yields $(-\frac{b}{a}y+\frac{c}{a}-f)^2+(y-g)^2=r^2$. As above, any solutions of this equation are also surds. Thus the theorem holds in this case too. $\square$

The remaining case is the intersection of two circles.

Theorem 12.3.11. *All points of intersection of two circles that have equations with surd coefficients lie in the surd plane.*

Proof. The equations of the circles can be written in the form

$$(x-a_1)^2+(y-b_1)^2=r_1^2$$
$$(x-a_2)^2+(y-b_2)^2=r_2^2$$

or

$$x^2-2a_1x+a_1^2+y^2-2b_1y+b_1^2=r_1^2$$
$$x^2-2a_2x+a_2^2+y^2-2b_2y+b_2^2=r_2^2$$

Subtracting the second equation from the first shows that any point (x,y) that lies on both circles also lies on the line with equation

$$(-2a_1+2a_2)x+a_1^2-a_2^2+(-2b_1+2b_2)y+b_1^2-b_2^2=r_1^2-r_2^2$$

Since this equation has surd coefficients, all points of intersection of this line with either circle lie in the surd plane (12.3.10). $\square$

Theorem 12.3.12. *The field of constructible numbers is the same as the field of surds.*

Proof. We already showed that every surd is constructible (12.3.3). On the other hand, Theorems 12.3.9 12.3.10 12.3.11 show that the only constructible points in the plane are points with surd coordinates. Since every constructible number is a coordinate of a constructible point in the plane, it follows that every constructible number is a surd. $\square$

This characterization of the constructible numbers is the key to the proof that certain angles cannot be trisected. One of the relationships between constructible angles and constructible numbers is the following; we restrict the discussion to acute angles (i.e., angles less than a right angle) simply to avoid having to describe several cases.

Theorem 12.3.13. *The acute angle θ is constructible with a straightedge and compass if and only if $\cos\theta$ is a constructible number.*

Proof. Suppose first that the angle θ is constructible. Place the angle so that its vertex lies at the point 0 on the number line and one of its sides is the positive part of the number line and the other side is on top of it, as in Figure 12.10. Use the compass to mark a point on the upper side of the angle that is one unit from the point 0. Drop a perpendicular from that point to the number line. Then that perpendicular meets the number line at $\cos\theta$ so $\cos\theta$ is constructed.

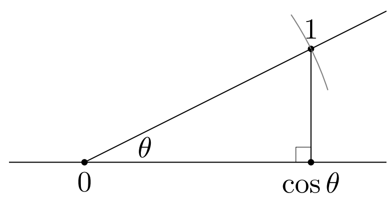


Figure 12.10

Conversely, if $\cos\theta$ is constructed, erect a perpendicular upwards from the point $\cos\theta$ on the number line. Construct the number $a = \sqrt{1-\cos^2\theta}$ and mark the point on the perpendicular that is that distance above the number line, as in Figure 12.11.

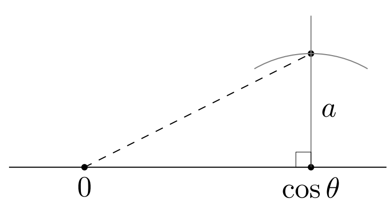


Figure 12.11

Connecting the point 0 to that marked point by a straightedge produces the angle θ. $\square$

With this background we can now determine exactly which angles with an integral number of degrees are constructible. First note the following.

Theorem 12.3.14. *An angle of 60 degrees is constructible.*

Proof. This is an immediate consequence of 12.3.13, for the cosine of 60 degrees equals $\frac{1}{2}$ and $\frac{1}{2}$ is a constructible number. There is also an easy direct proof: simply construct an equilateral triangle using a straightedge and compass, each angle of the equilateral triangle is 60 degrees. □

Corollary 12.3.15. *The following angles are all constructible: 30 degrees, 15 degrees, 45 degrees and 75 degrees.*

Proof. We begin with the fact that an angle of 60 degrees is constructible (12.3.14). An angle of 30 degrees can be constructed by bisecting an angle of 60 degrees, and an angle of 15 degrees can be constructed by bisecting an angle of 30 degrees. An angle of 45 degrees can be constructed by placing an angle of 15 degrees next to one of 30 degrees and an angle of 75 degrees can be constructed by placing an angle of 15 degrees next to an angle of 60 degrees. □

The material about constructible numbers was developed primarily to prove that some angles are not constructible. We need some additional preliminary results.

Theorem 12.3.16. *For any angle θ,* $\cos 3\theta = 4\cos^3\theta - 3\cos\theta$.

Proof. First note that
$$\cos 2\theta = \cos^2\theta - \sin^2\theta.$$
Then
$$\begin{aligned}\cos(3\theta) &= \cos(2\theta+\theta) = \cos 2\theta\cos\theta - \sin 2\theta\sin\theta \\ &= (\cos^2\theta - \sin^2\theta)\cos\theta - 2\sin\theta\cos\theta\sin\theta.\end{aligned}$$
Using $\sin^2(\theta) = 1 - \cos^2(\theta)$ gives,
$$\begin{aligned}\cos(3\theta) &= (2\cos^2\theta - 1)\cos\theta - 2\sin^2\theta\cos\theta \\ &= 2\cos^3\theta - \cos\theta - 2(\cos\theta - \cos^3\theta) \\ &= 4\cos^3\theta - 3\cos\theta\end{aligned}$$

Therefore $\cos 3\theta = 4\cos^3\theta - 3\cos\theta$. □

The case where θ equals 20 degrees is of particular interest.

Corollary 12.3.17. *If $x = 2\cos 20°$, then $x^3 - 3x - 1 = 0$.*

Proof. Using the formula for $\cos 3\theta$ given above and the fact that cosine of 60 degrees equals $\frac{1}{2}$ gives $\frac{1}{2} = 4\cos^3 20° - 3\cos 20°$. This is equivalent to the equation $8\cos^3 20° - 6\cos 20° - 1 = 0$. Let $x = 2\cos 20°$; then $x^3 - 3x - 1 = 0$. $\square$

We will show that the cubic equation $x^3 - 3x - 1 = 0$ does not have a constructible root.

Theorem 12.3.18. *If the roots of the cubic equation $x^3 + bx^2 + cx + d = 0$ are r_1, r_2, and r_3, then $b = -(r_1 + r_2 + r_3)$. (It is possible that two or even three of the roots are the same as each other.)*

Proof. By the factor theorem (9.3.4), and the fact that the coefficient of x^3 is 1, the cubic equation is the same as $(x - r_1)(x - r_2)(x - r_3) = 0$. Multiplying out these three factors shows that the coefficient of x^2 is $-(r_1 + r_2 + r_3)$; hence $b = -(r_1 + r_2 + r_3)$. $\square$

We need the concept of a conjugate for elements of an $\mathcal{F}(\sqrt{r})$, analagous to the conjugate of a complex number.

Definition 12.3.19. If $a + b\sqrt{r}$ is an element of an $\mathcal{F}\sqrt{r}$, then the conjugate of $a + b\sqrt{r}$, denoted by placing a bar on top of the number, is a $a - b\sqrt{r}$.

Theorem 12.3.20. *The conjugate of the sum of two elements of $\mathcal{F}(\sqrt{r})$ is the sum of the conjugates and the conjugate of the product of two elements of $\mathcal{F}(\sqrt{r})$ is the product of the conjugates.*

Proof. For the first assertion simply note that

$$\overline{(a + b\sqrt{r}) + (c + d\sqrt{r})} = \overline{(a + c) + (b + d)\sqrt{r}}$$
$$= a + c - (b + d)(\sqrt{r})$$
$$= (a - b\sqrt{r}) + (c - d\sqrt{r})$$
$$= \overline{(a + b\sqrt{r})} + \overline{(c + d\sqrt{r})}$$

For products, note that

$$\overline{(a + b\sqrt{r})(c + d\sqrt{r})} = \overline{(ac + rbd) + (ad + cb)\sqrt{r}}$$

$$= (ac + rbd) - (ad + cb)\sqrt{r}$$

Then

$$\begin{aligned}(\overline{(a + b\sqrt{r})})(\overline{(c + d\sqrt{r})}) &= (a - b\sqrt{r})(c - d\sqrt{r}) \\ &= (ac + bdr) - (ad + cb)\sqrt{r} \\ &= \overline{(a + b\sqrt{r})(c + d\sqrt{r})}\end{aligned}$$

□

Theorem 12.3.21. *If $a+b\sqrt{r}$ is a root of a polynomial with rational coefficients, then $a - b\sqrt{r}$ is also a root of the polynomial.*

Proof. Suppose that $a_n(a+b\sqrt{r})^n+a_{n-1}(a+b\sqrt{r})^{n-1}+\ldots+a_1(a+b\sqrt{r})+a_0 = 0$. Since each of the coefficients a_k is rational, $\overline{a_k} = a_k$ for all k. Using this fact and the facts that the conjugate of a product is a product of the conjugates and the conjugate of a sum is the sum of the conjugates, it follows that $a_n\overline{(a + b\sqrt{r})}^n + a_{n-1}\overline{(a + b\sqrt{r})}^{n-1} + \ldots + a_1\overline{(a + b\sqrt{r})} + a_0 = 0$. Thus $\overline{a + b\sqrt{r}}$ is also a root of the polynomial. □

Theorem 12.3.22. *If a cubic equation with rational coefficients has a constructible root, then the equation has a rational root.*

Proof. By dividing through by the leading coefficient, we can assume that the coefficient of x^3 is 1. Then, by 12.3.18, the sum of the three roots of the cubic equation is rational.

We first show that if the equation has a root in any $\mathcal{F}(\sqrt{r})$ then it has a root in $\mathcal{F}$. To see this, suppose the equation has a root in $\mathcal{F}(\sqrt{r})$ of the form $a+b\sqrt{r}$ with $b \neq 0$. Then, by 12.3.21, the conjugate $a - b\sqrt{r}$ is also a root. If r_3 is the third root and s is the sum of all three roots, then $s = r_3+a+b\sqrt{r}+a-b\sqrt{r} = r_3+2a$. Thus $r_3 = s - 2a$. Since s is rational, $\mathcal{F}$ contains all rational numbers and a is in $\mathcal{F}$, it follows that the root r_3 is in $\mathcal{F}$ itself.

The preliminary result obtained in the previous paragraph allows us to prove the theorem as follows. If the polynomial has a constructible root, then, since every constructible number is a surd (12.3.12), the root is in a number field that occurs at the end of a tower. Consider the number field at the end of the shortest tower that contains any root of the given cubic equation. We claim that that number field is $\mathbb{Q}$. To see this, simply note that if that root was in an $\mathcal{F}(\sqrt{r})$, the previous paragraph would imply that the root was in $\mathcal{F}$, which would be at the end of a shorter tower than $\mathcal{F}(\sqrt{r})$ is. Hence that root must be in $\mathbb{Q}$. Thus the equation has rational root. □

We can now prove that an angle of 20 degrees cannot be constructed.

Theorem 12.3.23. *An angle of 20 degrees cannot be constructed with straightedge and compass.*

Proof. If an angle of 20 degrees could be constructed with straightedge and compass, then $\cos 20°$ would be a constructible number (12.3.13). Then $2\cos 20°$ would be a constructible number, and the polynomial $x^3-3x-1=0$ would have a constructible root (12.3.17). It follows from the previous theorem (12.3.22) that this polynomial would have to have a rational root. Thus to establish that an angle of 20 degrees is not constructible, all that remains to be shown is that the polynomial $x^3-3x-1=0$ does not have a rational root.

Suppose that m and n are integers with $n \neq 0$ and that $\frac{m}{n}$ is in lowest terms and is a root of the equation $x^3-3x-1=0$. Then $\frac{m^3}{n^3}-3\frac{m}{n}-1=0$ implies that $m^3-3mn^2-n^3=0$. Since $n^3=m(m^2-3n^2)$, every prime number dividing m also divides n^3 and hence also divides n (4.1.3). Since m and n are relatively prime, there are no primes that divide m so m is either 1 or -1. Similarly since $m^3=n(3mn+n^2)$, any prime that divides n also divides m, from which it follows that n is 1 or -1. Hence $\frac{m}{n}$ is 1 or -1. Thus the only possible rational roots of $x^3-3x-1=0$ are $x=1$ or $x=-1$. Substituting those values for x in the equation shows neither of those is a root, so the theorem is proven. □

Corollary 12.3.24. *An angle of 60 degrees cannot be trisected with straightedge and compass.*

Proof. As we have seen, an angle of 60 degrees can be constructed with a straightedge and compass (12.3.14). If an angle of 60 degrees could be trisected with straightedge and compass, then an angle of 20 degrees would be constructible. But an angle of 20 degrees is not constructible 12.3.23. □

12.4 Constructions of Geometric Figures

Another problem that the Ancients Greeks raised but could not solve was what they called "duplication of the cube". This was the question of whether or not a side of a cube of volume 2 could be constructed by straightedge and compass.

Theorem 12.4.1. *The side of a cube of volume 2 cannot be constructed with a straightedge and compass.*

Proof. If x is the length of the side of a cube of volume 2, then, of course, $x^3 = 2$, or $x^3 - 2 = 0$. By Theorem 12.3.22, this equation has a constructible root if and only if it has a rational root. Since the cube root of 2 is irrational, there is no constructible solution, and the cube cannot "be duplicated" using only a straightedge and compass. □

The question of which regular polygons can be constructed is very interesting.

Definition 12.4.2. A *polygon* is a figure in the plane consisting of line segments that bound a finite portion of the plane. A *regular polygon* is a polygon all of whose angles are equal and all of whose sides are equal.

An equilateral triangle is a regular triangle with three sides. Equilateral triangles can easily be constructed with straightedge and compass, as follows. Construct an angle of 60 degrees (12.3.14) use the compass to mark equal lengths on the two sides of the angle and then use the straightedge to connect the ends of those lengths. A square is a regular polygon with four sides. It is also very easy to construct a square. Simply use the straightedge to draw any line segment, and erect perpendiculars at each end of the line segment. Then use the compass to "measure" the length of the line segment and mark points that are that distance above the original line segment on each of the perpendicular bisectors. Using the straightedge to connect those points yields a square. For each natural number n bigger than or equal to 3, there is a regular polygon of n equal sides. This can be seen as follows.

Theorem 12.4.3. *For each natural number n greater than or equal to 3 there is a regular polygon with n sides inscribed in a circle.*

Proof. Given a natural number n bigger than or equal to 3, take a circle and draw successive adjacent angles of size $\frac{360}{n}$ at the centre as shown in Figure 12.12. Then draw the line segments connecting adjacent points determined by the sides of the angles intersecting the circumference of the circle. We must show that those line segments are all equal in length and that the angles formed by each pair of adjacent line segments are equal to each other.

Consider, for example, the triangle OAB and OCD in Figure 12.12. The angles AOB and COD are each equal to $\frac{360}{n}$. The sides OA, OB, OC and OD are all radii of the given circle and are therefore equal to each other. It follows that $\triangle OAB$ is congruent to triangle OCD by side-angle-side (11.1.2). The same proof shows that all of the triangles constructed are congruent to each other. It follows that all of the sides of the polygon, which are the sides opposite the angles of $\frac{360}{n}$ in the triangle, are equal to each other. The angles of the polygon

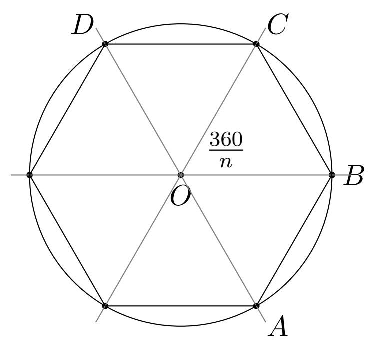


Figure 12.12

are angles such as $\angle ABC$ and $\angle BCD$ in the diagram. Each of them is the sum of two base angles of the drawn triangles and thus the angles of the polygon are equal to each other as well. $\square$

Definition 12.4.4. A *central angle* of a regular polygon of n sides is the angle of $\frac{360}{n}$ degrees that has a vertex at the center of the polygon, as in the above proof.

Theorem 12.4.5. *A regular polygon is constructible if and only if its central angle is a constructible angle.*

Proof. Suppose that a regular polygon can be constructed with straightedge and compass. Then its center (a point equidistant from all of its vertices) can be constructed as the point of intersection of the perpendicular bisectors of two adjacent sides of the polygon (See Problem 10). Now the central angle can be constructed as the angle formed by connecting the center to two adjacent vertices of the polygon. All such angles are equal to each other, since the corresponding triangles are congruent. There are n such angles, the sum of which is 360 degrees, so each central angle is $\frac{360}{n}$ degrees.

Conversely suppose that an angle of $\frac{360}{n}$ degrees is constructible for some natural number $n \geq 3$. Then a regular polygon with n sides can be constructed as follows. Make a circle. Construct an angle of $\frac{360}{n}$ degrees with vertex at the center of the circle. Then construct another such angle adjacent to the first and so on until n such angles have been constructed. Connecting the adjacent points of intersection of the sides of those angles with the circle constructs a regular polygon with n sides. $\square$

Corollary 12.4.6. *A regular polygon with 18 sides cannot be constructed with a straightedge and compass.*

Proof. A regular polygon with 18 sides has a central angle of $\frac{360}{18} = 20$ degrees. We proved (12.3.23) an angle of 20 degrees is not constructible, so the previous theorem implies a regular polygon with 18 sides is not constructible. □

Theorem 12.4.7. *If m is a natural number greater than 2, then a regular polygon with $2m$ sides is constructible if and only if a regular polygon with m sides is constructible.*

Proof. Using 12.4.5, the theorem follows by either bisecting or doubling the central angle of the already constructed polygon. (Alternatively, having constructed a regular polygon of $2m$ sides, use the straightedge to connect alternate vertices, yielding a regular polygon of m sides, as can be established by using congruent triangles. In the other direction, given a regular polygon of m sides, inscribe it in a circle and then double the vertices by adding the points of intersections of the perpendicular bisectors of the sides and the circle.) □

Corollary 12.4.8. *A regular polygon of 9 sides is not constructible.*

Proof. This follows immediately from the fact that a regular polygon with 18 sides is not constructible (12.4.6), and the above (12.4.7). □

It is useful to make the following connection between constructible polygons and constructible numbers.

Theorem 12.4.9. *A regular polygon with n sides is constructible if and only if the length of the side of a regular polygon with n sides that is inscribed in a circle of radius 1 is a constructible number.*

Proof. In the first direction suppose that a regular polygon with n sides is constructible. Then such a polygon can be constructed so that it is inscribed in a circle of radius 1 (for example, by putting its constructible central angle in a circle of radius 1). The length of the side can be constructed by using the compass to "measure" the side of the constructed polygon.

Conversely if s is a constructible number and is the length of the side of a regular polygon with n sides inscribed in a circle of radius 1, the regular polygon can be constructed simply by marking any point on the circle and then using the compass to successively mark points that are at distance s from the last marked one. The marked points will be vertices of a regular polygon with n sides. □

Can a pentagon (a regular polygon with 5 sides) be constructed using only a straightedge and compass? The answer is affirmative, but it is not at all easy to see directly. We will approach this by considering a regular polygon with 10 sides.

Theorem 12.4.10. *A regular polygon with 10 sides is constructible.*

Proof. By 12.4.9 it suffices to show that the length of a side of such a polygon inscribed in a circle of radius 1 is a constructible number. We determine the length of such a side by using a little geometry. The central angle of a regular polygon with 10 sides is 36 degrees. Consider such an angle with vertex O at the center of a circle of radius 1, as shown in Figure 12.13. Label the points of intersection of the sides of that central angle with the circle A and B respectively. Let s denote the length of the line segment from A to B. Let AC be the bisector of $\angle OAB$. Since $\angle OAB$ is 72 degrees (the sum of the degrees of the equal

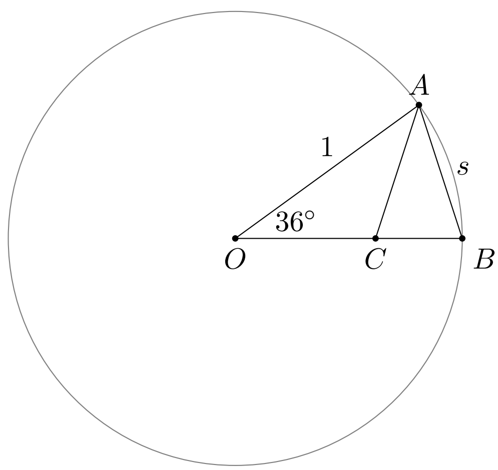


Figure 12.13

angles OAB and ABO must be $180° - 36°$). It follows that angles OAC and CAB are each 36 degrees. Thus triangles AOB and CAB are similar to each other, so corresponding sides are in proportion (11.3.11). Therefore triangle ACB is isosceles, and AC has length s.

Since $\angle AOB = 36° = \angle OAC$, $\triangle OAC$ is also isosceles so OC has length s, from which it follows that BC has length $1 - s$. The side opposite the 36 degree angle in $\triangle OAB$, length s, is to the side opposite the 36 degree of $\triangle CAB$, length $1 - s$, as is the side opposite the 72 degree angles of $\triangle OAB$, length 1 (the radius of the circle), is to the side opposite the 72 degree angle of $\triangle BAC$, which has length s. That is, $\frac{s}{1-s} = \frac{1}{s}$. Thus the length we are interested in, s, satisfies the equation $s^2 - 1 = s$, or $s^2 + s - 1 = 0$. The positive solution of this equation (s is a length) is $\frac{-1+\sqrt{5}}{2}$. Thus s is a constructible number (12.3.12) from which it follows that the regular polygon with 10 sides is constructible. $\square$

Corollary 12.4.11. *A regular pentagon is constructible.*

Proof. This follows immediately from the above theorem and 12.4.7. $\square$

What regular polygons are constructible? Those with 3, 4 and 5 sides are, and thus so are 6 and 8 and 10 (12.4.7). We proved that a regular polygon with 9 sides is not constructible 12.4.8.

What about a polygon with 7 sides? We can approach this question using some facts that we learned about complex numbers. As we have seen, REFERENCE for each natural number greater than 2 the complex solutions to the equation $z^n = 1$ are the vertices of an n sided regular polygon inscribed in a circle of radius 1. We will approach the problem by considering the solutions of $z^7 = 1$,

Theorem 12.4.12. *A regular polygon with 7 sides is not constructible.*

Proof. If a regular polygon with 7 sides was constructible, then one could be constructed inscribed in a circle of radius 1 such that one of the vertices lies on the x axis at the point corresponding to the number 1. We will analyze the next vertex above the x axis. Let that vertex lie at the complex number z_0. If the regular polygon was constructible then z_0 would be a constructible point and therefore the real part of z_0 would be constructible (simply construct a perpendicular from z_0 to the x axis). It would follow that twice the real part is constructible. Let x_0 be twice that real part. We will show that x_0 satisfies a cubic equation that is not satisfied by any constructible number. Now $x_0 = z_0 + \overline{z}_0$. Since $|z_0| = 1$, $\overline{z}_0 = \frac{1}{z_0}$, so $x_0 = z_0 + \frac{1}{z_0}$.

The cubic equation satisfied by x_0 will be obtained from the equation of degree 7 satisfied by z_0. Now $z_0^7 = 1$ and $z_0 \neq 1$. Also $z^7 - 1 = (z-1)(z^6 + z^5 + z^4 + z^3 + z^2 + z + 1)$. Since $z_0 - 1 \neq 0$, $z_0^6 + z_0^5 + z_0^4 + z_0^3 + z_0^2 + z_0 + 1 = 0$. Dividing through by z_0^3 yields

$$z_0^3 + z_0^2 + z_0 + 1 + \frac{1}{z_0} + \frac{1}{z_0^2} + \frac{1}{z_0^3} = 0.$$

Note that $(z_0+\frac{1}{z_0})^3 = z_0^3+3z_0+\frac{3}{z_0}+(\frac{1}{z_0})^3$, and also that $(z_0+\frac{1}{z_0})^2 = z_0^2+2+(\frac{1}{z_0})^2$. It follows that

$$z_0^3 + z_0^2 + z_0 + 1 + \frac{1}{z_0} + \frac{1}{z_0^2} + \frac{1}{z_0^3} = (z_0 + \frac{1}{z_0})^3 + (z_0 + \frac{1}{z_0})^2 - 2(z_0 + \frac{1}{z_0}) - 1.$$

Then since $x_0 = z_0 + \frac{1}{z_0}$, x_0 satisfies the equation

$$x_0^3 + x_0^2 - 2x_0 - 1 = 0.$$

As indicated, to show that a regular polygon with 7 sides is not constructible, it suffices to show that x_0 is not a constructible number. Since x_0 satisfies this cubic equation with rational coefficients, the result will follow if it is shown that this cubic equation has no rational root (12.3.22). Suppose that the rational number $\frac{m}{n}$ satisfied this cubic equation. We can and do assume that m and n have no common integral factor other than 1 and -1. Then $(\frac{m}{n})^3 + (\frac{m}{n})^2 - 2(\frac{m}{n}) - 1 = 0$, or $m^3 + m^2n - 2mn^2 - n^3 = 0$. Now if p was a prime number that divided m, it follows from the above that p would divide n^3 and hence also divide n. Since m and n are relatively prime there is no such prime number p, and we conclude that m is either 1 or -1. Similarly, n is equal to 1 or n is equal to -1. Thus $\frac{m}{n}$ equals 1 or -1. But $1^3 + 1^2 - 2 - 1$ is not 0, nor is $(-1)^3 + (-1)^2 + 2 - 1$. Hence there is no rational solution, and the theorem is proven. □

STATE GAUSS -WENTZEL THEOREM

We can determine exactly which angles of a natural number of degrees are constructible.

Theorem 12.4.13. *If n is a natural number, then an angle of n degrees is constructible if and only if n is a multiple of 3.*

Proof. Recall that we proved that a regular polygon with 10 sides is constructible (12.4.10), and, hence, that an angle of 36 degrees is constructible (12.4.5). Since an angle of 30 degrees is constructible 12.3.15, we can "subtract" a 30 degree angle from a 36 degree one by placing the 30 degree angle with the vertex and one of its sides coincident with the vertex and one of the sides of the 36 degree angle 12.1.6. Then bisecting the constructed angle of 6 degrees yields an angle of 3 degrees.

Once an angle of 3 degrees is constructed, an angle of $3k$ degrees can be constructed by simply placing k angles of 3 degrees appropriately.

To establish the converse, suppose that an angle of n degrees is constructible. We must show that n is congruent to $0 \pmod 3$. If n was congruent to either 1 or 2 mod 3, then we could construct an angle of 1 degree or 2 degrees accordingly by "subtracting" an appropriate number of angles of 3 degrees from the angle of n degrees. If the resulting angle is 2 degrees, bisecting it would yield an angle of 1 degree. Thus if an angle of n degrees was constructible and n was not a multiple of 3, then an angle of 1 degree could be constructed. But an angle of 1 degree is not constructible, for if it was, placing 20 of them together would contradict the fact that an angle of 20 degrees is not constructible. □

We have shown that some angles, such as an angle of 60 degrees, cannot be trisected with a straightedge and compass. But what about the following?

Example 12.4.14 (Trisection of arbitrary acute angles.)**.** Let θ be any acute angle. Mark any two points on your straightedge and let the distance between them be r. Draw the angle θ and construct the circle with radius r whose center is at the vertex of θ. Label the center of the circle O. Extend one of the sides of θ in both directions. Move the marked straightedge so that the point marked to the left is on the extended line, the point marked to the right stays on the circle and the straightedge passes through the intersection of the circle and the side of $\angle\theta$ that was not extended, and label the points of intersection A, B, C, as shown in Figure 12.14. Draw the line BO. Then the line segments AB, BO and OC all have length r.

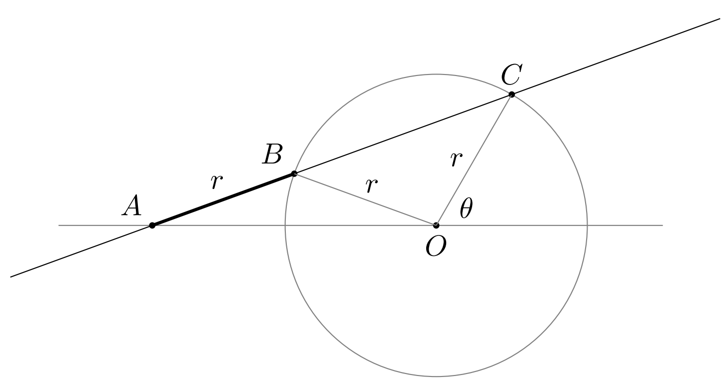


Figure 12.14

Now let the equal base angles of $\triangle ABO$ be x and the equal base angles of $\triangle OBC$ be y, and let $\angle BOC$ be z, as shown in Figure 12.15. Then the sum of $\angle ABO$ and $2x$ is a 180° and the sum of $\angle ABO$ and y is also 180°; hence $y = 2x$.

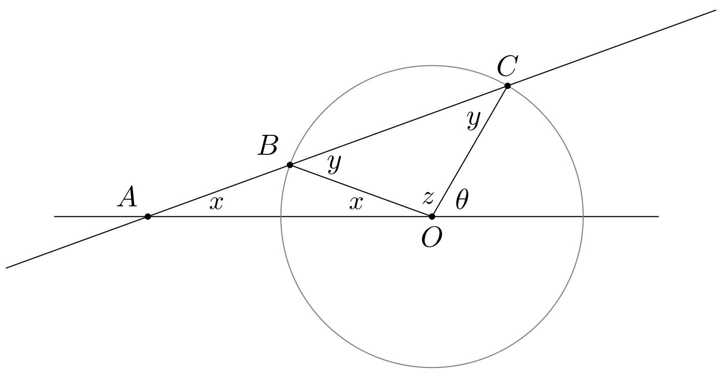


Figure 12.15

It is clear that $x+z+\theta$ is 180°, on the other hand, $z+2y$ is 180°. Since $y=2x$, $4x+z$ is also 180°. It follows that $4x+z=x+\theta+z$, or $3x=\theta$. Thus the angle x is one third of θ, so θ has been trisected. $\square$

You may think that the construction we have just done contradicts our earlier proof that an angle of 60 degrees cannot be trisected. However, the construction in the example above violated the classical rules of constructions that we were adhering to before this example. Namely we marked two points on the straightedge. What we have shown is that it is possible to trisect arbitrary angles with a compass and straightedge on which two (or more) points are marked. Therefore, in particular, any angle can be trisected using a *ruler* and compass, but not with a straightedge and compass.

12.5 Problems

Basic Exercises

1. Determine which of the following numbers are constructible.

   (a) $\frac{1}{\sqrt{3+\sqrt{2}}}$

   (b) $\sqrt[6]{79}$

   (c) 3.146891

   (d) $\sqrt[16]{79}$

   (e) $\sqrt{6+\frac{\sqrt[3]{4}}{2}}$

   (f) $\sqrt{7+\sqrt{5}}$

   (g) $\sqrt{3+4\sqrt{2}+\sqrt{5}}$

   (h) $\sin(\frac{\pi}{16})$

   (i) π^{10}

   (j) $\sqrt{\pi^2+4}$

   (k) $\sqrt[3]{\frac{9}{10}}$

   (l) $\tan(2.5°)$

   (m) $\sin(20°)$

(n) $\sqrt{\frac{3792}{1419}}$
(o) $\sin(75^\circ)$
(p) $\cos(50^\circ)$
(q) $\cos(5^\circ)$
(r) $\cos(10^\circ)$
(s) $\cos\pi$
(t) $11^{\frac{2}{3}}$
(u) $11^{\frac{3}{2}}$
(v) $\tan\frac{\pi}{4}$
(w) $2^{\frac{1}{6}}$
(x) $2^{\frac{3}{2}}$
(y) $\sqrt[3]{\frac{\sqrt{2}}{4}}$
(z) $\sqrt{7\cos(15^\circ)}$

2. Determine which of the following angles are constructible.

(a) 6°
(b) 5°
(c) 10°
(d) 30°
(e) 35°
(f) 15°
(g) 75°
(h) 80°
(i) 92.5°
(j) 37.5°
(k) 7.5°
(l) 120°
(m) 160°

3. Determine which of the following angles can be trisected.

(a) 10°
(b) 30°

4. Is there a constructible angle θ such that $\cos\theta = \frac{\pi}{6}$?

Interesting Applications

5. Determine which of the following polynomials have at least one constructible root.

(a) $x^4 - 3$
(b) $x^8 - 7$
(c) $x^4 + \sqrt{7}x^2 - \sqrt{3} - 1$
(d) $x^3 + 6x^2 + 9x - 10$
(e) $x^3 - 3x^2 - 2x + 6$
(f) $x^3 - 2x - 1$
(g) $x^3 + 4x + 1$
(h) $x^3 + 2x^2 - x - 1$
(i) $x^3 - x^2 + x - 1$

(j) $2x^3 - 4x^2 + 1$

6. Determine which of of the following regular polygons can be constructed with straightedge and compass?
   (a) A regular polygon with 14 sides.
   (b) A regular polygon with 20 sides.
   (c) A regular polygon with 36 sides.
   (d) A regular polygon with 240 sides.

7. Explain how to construct a regular polygon with 24 sides using straightedge and compass.

8. True or False:
   (a) If the angle of θ degrees is constructible and the number x is constructible, then the angle of $x\theta$ degrees is constructible.
   (b) x^y is a surd if x and y are surds.
   (c) If $\frac{x}{z}$ is constructible, then x and z are each constructible.
   (d) There is an angle θ such that $\cos\theta$ is constructible but $\sin\theta$ is not constructible.

9. Show that $\tan\theta$ is a constructible number if and only if θ is a constructible angle.

Challenging Problems

10. Prove that given a regular polygon, its center can be constructed using only a straightedge and compass. (Hint: The center can be determined as the point of intersection of the perpendicular bisectors of two adjacent sides of the polygon. To prove that this point is indeed the center, prove that all the right triangles with one side a perpendicular bisector of a side of the polygon, another side a half of a side of the polygon and the third side the line segment joining the "center" to a vertex of the polygon are congruent to each other.)

11. Prove that the side of a cube with volume a natural number n is constructible if and only if $n^{\frac{1}{3}}$ is a natural number.

12. Prove that an acute angle θ cannot be trisected with straightedge and compass if $\cos\theta$ equals:

(a) $\frac{3}{7}$
(b) $\frac{2}{5}$
(c) $\frac{1}{5}$
(d) $\frac{3}{5}$
(e) $\frac{1}{4}$

13. Can a polynomial of degree 4 with rational coefficients have a constructible root without having a rational root?

14. Prove that the cube cannot be tripled in the sense that starting with an edge of a cube of volume 1, an edge of a cube of volume 3 cannot be constructed with straightedge and compass.

15. Prove that the following equation has no constructible solutions:

$$x^3 - 6x + 2\sqrt{2} = 0.$$

Hint: You can use Theorem 12.3.22 if you make an appropriate substitution.

16. Using mathematical induction, prove that for every integer $n \geq 2$, a regular polygon with $3 \cdot 4^n$ sides can be constructed with straightedge and compass.

17. Let t be a transcendental number. Prove that $\{a + bt \mid a, b \in \mathbb{Q}\}$ is not a number field.

18. Let t be a transcendental number. Prove that t cannot be a root of any equation of the form $x^2 + ax + b = 0$ where a and b are constructible numbers.

19. What is the cardinality of the set of all constructible points in the plane?

20. Is there a line in the plane such that every point on it is constructible?

21. Say that a complex number $a + bi$ is constructible if the point (a, b) is constructible (equivalently, if a and b are both constructible real numbers). Show that the cube roots of $\frac{1}{2} + \frac{\sqrt{3}}{2}i$ are not constructible.

22. Find the cardinality of each of the following sets.

    (a) The set of roots of polynomials with constructible coefficients.
    (b) The set of constructible angles.
    (c) The set of all points (x, y) in the plane such that x is constructible and y is irrational.

(d) The set of all sets of constructible numbers.

23. Let $\mathcal{F}$ be the smallest number field containing π. Show that $\mathcal{F}$ is countable.

24. Which of the following subsets of $\mathbb{R}$ are number fields?

    (a) $\{a\sqrt{2} \ : \ a \in \mathbb{Q}\}$

    (b) $\{a + b\pi \ : \ a, b \in \mathbb{Q}\}$ (Use the fact that π is transcendental.)

25. Is the set of all towers countable? (Recall that a *tower* is a finite sequence of number fields, the first of which is $\mathbb{Q}$, such that the other number fields are obtained from their predecessors by adjoining square roots.)

26. (Very challenging) Use a straightedge and compass to directly construct a regular pentagon.

27. Prove the following.

    (a) If x_0 is a root of a polynomial with coefficients in $\mathcal{F}(\sqrt{r})$, then x_0 is a root of a polynomial with coefficients in $\mathcal{F}$.

    (b) Every constructible number is algebraic.

    (c) The set of constructible numbers is countable.

    (d) There is a circle with center at the origin that is not constructible.