Classical Geometric Construction

Tom Davis
tomrdavis@earthlink.net
http://www.geometer.org/mathcircles
December 4, 2002

1 Example

The best introduction to classical geometric construction is to show an example.

In Section 3, the "official" rules for geometric construction will be presented, but we'll begin with just an intuitive idea of what can be done with a straightedge and compass.

What we would like to do is to begin with a line segment, and using only a straightedge and compass, divide that line into two equal segments. In other words, if we are given a line segment AB, we would like to construct the midpoint M of that segment. Then we have bisected the segment, since $AM = MB$.

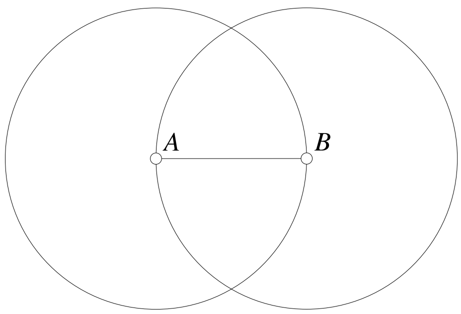

Figure 1: Bisecting a segment—steps 1 and 2

Figure 1 shows the first two steps. A compass is used with center at A to draw a circle passing through B, and in the same way, another circle is drawn centered at B and passing through point A.

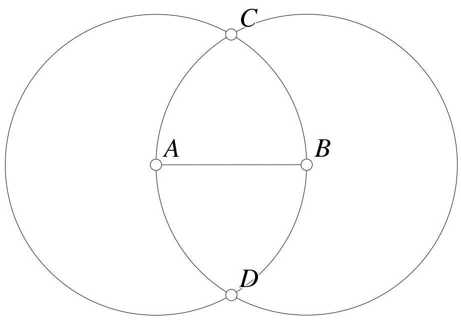

Figure 2: Bisecting a segment—steps 3 and 4

In Figure 2, two new points are determined—the intersections of the two circles that were drawn in step 1. Label these two points C and D.

The final two steps are shown in Figure 3. The straightedge is used to connect points C and D with a new line, and finally, the intersection M of that line with the original segment AB is determined. M is the required midpoint. (The

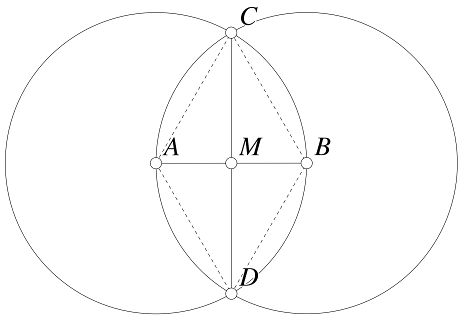

Figure 3: Bisecting a segment—steps 5 and 6

other segments in the figure, such as CA, CB, DA, and DB are not part of the construction, but we will use them below.)

It may seem obvious from the figure that M is the midpoint, but part of any construction is a proof that the required point has been found. In this case it is very easy to do—notice that all the lengths CA, CB, DA, and DB are equal, since they are radii of two circles with radius AB. Therefore the quadrilateral $ACBD$ is a parallelogram, and we know that the diagonals of a parallelogram bisect each other. There are dozens of other easy proofs.

2 Who Cares?

Why should anyone care about what figures can be constructed using only these two specialized tools? Why not choose different tools, or more tools?

One reason for the choice is that the fundamental objects in classical plane Euclidean geometry are points, lines, and circles, and a straightedge and compass are the minimal set of tools that will allow any circle or any line to be drawn. So it is obviously interesting to look at what can be drawn (constructed) using only the two tools that are obviously necessary to draw all the sorts of figures that are interesting.

The other reason for the choice of a straightedge and compass is that this minimal set of tools enables one to make an amazingly large collection of constructions. Also, for thousands of years, many people thought that any "reasonable" construction could be done with just those two tools. There were, however, 3 famous constructions (see Section 6) that nobody seemed to be able to do (using just a straightedge and compass, that is).

Finally, doing classical straightedge and compass constructions provides a wonderful selection of problems that helps to reinforce important geometric concepts from all parts of geometry.

3 Rules

OK, so what are the "official rules" for straightedge and compass constructions? For example, can marks be made on the straightedge? Can you hold the compass against the straightedge to use it as a measuring device for one particular length? If you have a circle drawn already, can you put a straightedge against it so that it just touches (is tangent to) the circle? Et cetera.

3.1 The Official Rules

The rules below are actually a little more flexible than the "official" rules, but they are easy to understand, and the difference between these rules and the official rules is explained in the footnote. Here they are:

A construction problem begins with a set of given points, lines, and circles, and with some desired point, line or circle to construct, based on the given objects. In the example in Section 1, we were given two points and the segment between them, and the object of the construction was to find the point that lies on the segment midway between the two original points.

Notice that we could reduce the statement above to require that only certain points be constructed. Obviously, if you want a line, you can simply require that two different points be found on it, or if you want a circle you can require that the center and a point on the circle be found.

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

1. You may draw a straight line of any length through two existing points. (This means, of course, that the straightedge is as long as you need it to be, so it is better than a real ruler in that sense.)

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

3. You may construct a circle centered at any existing point having a radius equal to the distance between any two existing points. In other words, you can set the size of the compass from any two points A and B, and then you can move the point of the compass to another point C without changing the setting and draw a circle of radius AB about the point C[1]. (Of course this includes drawing a circle given its center and a point on the edge—you use the center and the edge to set the compass size, and then you re-use the center point as the center of the circle.) As with the straightedge, there is no limit to the size of a circle that can be drawn, so the mathematical compass is better than any real one could be.

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

3.2 Other Possible Rules

In fact, constructions using other sets of tools have been examined, but we will not go into that here. For example, you might be interested in what can be constructed using only a straightedge, or just a compass, or perhaps a straightedge and a "rusty compass"—a compass that is frozen into one particular setting, or even constructions using more interesting tools.

Or what if you are allowed to put marks on the straightedge, effectively turning it into a ruler? What if the straightedge is only a fixed finite length? What if one circle has already been drawn, but you no longer have a compass? There are any number of other possibilities.

3.3 Some Final Warnings

Obviously, with a physical straightedge, compass, and pencil, no construction can be carried out to infinite accuracy; what is required here is a construction that would be to infinite accuracy, assuming that all the tools work to infinite accuracy.

In other words, it is not good enough to get an answer that is accurate to one part in a million or to one part in a trillion—an infinitely accurate result is required. As we will see later, if you "only" need accuracy to one part in a trillion, you can construct anything with a straightedge and compass.

Finally, the construction has to be completed in a finite number of steps; you cannot require an infinite set of operations to construct a point that is the limit of these operations. If an infinite number of steps were allowed, again, anything can be constructed. Simply construct points to an accuracy of one in a million, then one in a billion, then one in a trillion, and continue forever. The limit of these points will be the exact solution, but only after an infinite number of steps.

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

4 Basic Constructions

Here are a list of elementary constructions that can be carried out with a straightedge and compass. They are arranged roughly in order of difficulty, and if you can do most of these, you can do most standard geometric constructions.

1. Copy a segment. In other words, mark off a segment that exactly matches the length of a given segment on a different straight line.
2. Copy an angle. Given an angle, make another angle of exactly the same size somewhere else.
3. Bisect a segment. This problem was already solved in Section 1
4. Bisect an angle. Given an angle, find a line through the vertex that divides it in half.
5. Construct a line perpendicular to a given line through a point on the given line.
6. Construct a line perpendicular to a given line and passing through a point not on the given line.
7. Given a line L and a point P not on L, construct a new line that passes through P and is parallel to L.
8. Construct an angle whose size is the sum or difference of two given angles.
9. Given three segments, construct a triangle whose sides have the same lengths as the segments.
10. Construct the perpendicular bisector of a line segment.
11. Given three points, construct the circle that passes through all of them.
12. Given a circle, find its center.
13. Given a triangle T, construct the inscribed and circumscribed circles. The inscribed circle is a circle that fits inside the triangle and touches all three edges; the circumscribed circle is outside the triangle except that it touches all three of the vertices of the triangle.
14. Construct angles of $90°$, $45°$, $30°$, $60°$, and if you want a challenge, $72°$.
15. Construct a regular pentagon. (A regular pentagon is a five sided figure all of whose sides and angles are equal.)
16. Given a point P on a circle C, construct a line through P and tangent to C.
17. Given a circle C and a point P not on C, construct a line through P and tangent to C.
18. Given two circles C_1 and C_2, find lines internally and externally tangent to both.

5 Algebraic Equivalent

The original construction problems began with the Greeks, and for thousands of years, the methods were the same. Problems would be stated, a construction would be found, and then a standard geometric proof was supplied to show that the construction in fact behaved as advertised.

This worked great for thousands of years, except that it did not provide any method to show that certain constructions were impossible (see Section 6). But nobody thought that impossible constructions existed, so there was no real reason to do so.

Now we know that (in a sense) "almost all" constructions are impossible. In this section we will not prove that fact, but we will provide some overwhelming evidence. But to do so, we'll have to look at geometric construction from an algebraic point of view.

We will simplify the problem of construction to be that of finding only points. As we stated in Section 3.1, this is good enough, since any lines or circles that you might want can be identified in terms of a couple of points.

5.1 Arithmetic with Straightedge and Compass

One way to look at arithmetic in terms of geometry is to let the lengths of line segments represent numbers. If we have two different line segments (or in our simplification, two different sets of points), we have two lengths, and it is easy to see how to add or subtract the lengths, given only a straightedge and compass (see the arithmetic exercises in Section 5.2).

Addition and subtraction is easy, but a problem occurs when we talk about multiplication—a problem having to do with units. If you have two segments, for example that are one meter long, if you multiply them, you'll get something that is a square meter. How can you compare the area of a square with the length of a line? In fact, you cannot—if you think the answer is 1, since $1 \times 1 = 1$, why isn't the answer 10000, since the lines are 100 centimeters long, and $100 \times 100 = 10000$? In fact, an equally good case can be made for any such result. Geometrically, you can only compare lengths with lengths, areas with areas, volumes with volumes, et cetera.

A similar problem arises with division, square roots, cube roots, et cetera. There are a couple of ways to get around the problems. The easiest is when something like multiplication or square roots occurs is to give an additional segment whose length is defined to be 1, and to work in terms of that. Another is so make sure that the units are right, so it is possible to construct a segment of length AB/C or $\sqrt{AB}$, if A, B, and C are lengths of given segments.

5.2 Arithmetic Exercises

In the exercises below, assume that you are given pairs of points whose distance apart represents a length.

1. **Addition and subtraction.** Given segments of lengths A and B, construct a segment of length $A+B$ or $A-B$.
2. **Multiplication.** Given segments of lengths A, B, and 1, construct a segment of length AB.
3. **Division.** Given segments of lengths A and 1, construct a segment of length $1/A$. Using the result of the problem above, show that if you also have a segment of length B, you can construct a segment of length B/A.
4. **Square roots.** Given segments of lengths A and B, construct a segment whose length is $\sqrt{AB}$. Clearly, if one of the segments has length 1, you can construct a segment of length $\sqrt{A}$ using the same technique.
5. **Square a rectangle.** Given a rectangle, construct a square with exactly the same area.
6. **Construct a pentagon.** Construct a regular pentagon. For a hint, see the footnote[2].

5.3 Simple Algebraic Concepts

Now imagine that the given points and the required points are drawn in some nice Cartesian coordinate system. Then finding new points is equivalent to finding their coordinates. In other words, a construction problem can be restated in an algebraic form something like this:

- Given a set of pairs of coordinates: $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$, construct some other pairs of coordinates $(X_1, Y_1), (X_2, Y_2), \ldots, (X_m, Y_m)$ using the following algebraic transformations ....

As we'll see below, those "algebraic transformations" are pretty simple to describe, and such a description will allow us to identify the sorts of sets of coordinates that can be constructed using a straightedge and compass.

Here's the idea—we just need to know how to how each of the allowed geometric constructions translates into algebra. Let's take a look:

- **You can draw a line through two points.** Well, it is easy to write down the equation of a line passing through (x_1, y_1) and (x_2, y_2):
$$(y - y_1) = \frac{(y_2 - y_1)}{(x_2 - x_1)}(x - x_1).$$

[2] The cosine of 72° is $(\sqrt{5} - 1)/4$

- **You can draw a circle with any point as center and any other point on the boundary.** If (x_1, y_1) is the center, and (x_2, y_2) is a point on the boundary, we know the radius r is given by:

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

  Then the equation of a circle centered at (x_1, y_1) and having radius r is given by:

$$(x - x_1)^2 + (y - y_1)^2 = r^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2.$$

- **You can find the intersection of any two lines.** Again, this is easy—if the given lines have the equations $A_1x + B_1y + C_1 = 0$ and $A_2x + B_2y + C_2 = 0$, a simultaneous solution of these two equations is simply a matter of algebra, and the result is just an equation in the variables A_1, B_1, C_1, A_2, B_2, and C_2.

- **You can find points that are the intersection of a line and a circle.** In this case, suppose the line has equation $Ax + By + C = 0$, and that the circle has equation $(x - D)^2 + (y - E)^2 = R^2$. Then we can solve the equation of the line for x, giving $x = -(By + C)/A$, and substituting this into the equation for the circle. Solving will involve a quadratic equation, but the solution will at worst require the quadratic formula, and thus a square root of some numbers we already have. Note that there are two, one, or zero roots, depending on whether the line cuts the circle, is tangent to it, or misses it entirely.

- **Finally, we need to be able to find the intersections of two circles.** In this case, suppose that the equations of the two circles are given by: $(x - A_1)^2 + (y - B_1)^2 = R_1^2$ and $(x - A_2)^2 + (y - B_2)^2 = R_2^2$. But if you multiply these out, you get two equations that look like this:

$$\begin{aligned} x^2 + y^2 + ax + by + c &= 0 \\ x^2 + y^2 + dx + ey + f &= 0. \end{aligned}$$

  If you subtract them, you can see that the system above is equivalent, algebraically, to the system:

$$\begin{aligned} x^2 + y^2 + ax + by + c &= 0 \\ (a - d)x + (b - e)y + (c - f) &= 0. \end{aligned}$$

  This can be solved exactly as we did in the line-circle case, giving up to two solutions, and requiring nothing more than a square root in the quadratic formula.

So if you look at the algebraic result of any of the above constructions, you can see that nothing is much worse than a combination of addition, subtraction, multiplication, division, and square roots of the numbers (coordinates) you started with before the construction.

It is also not hard to show that any such arithmetic operation can be emulated with a proper geometric construction, so all of those sorts of numbers (coordinates) can be constructed. So here's the idea of where you can get after a certain number of steps, beginning only with the number 1:

1. You can make all the integers: $1, 2, 3, \ldots$.
2. You can make all the fractions (usually called "rational numbers": i/j, where i and j are integers).
3. You can make anything that looks like $q_1 + q_2\sqrt{q_3}$, where q_1, q_2 and q_3 are any rational number.
4. You can make anything that looks like $x_1 + x_2\sqrt{x_3}$, where x_1, x_2 and x_3 are any numbers you generate in step 3.
5. You can make anything that looks like $y_1 + y_2\sqrt{y_3}$, where y_1, y_2 and y_3 are any numbers you generate in step 4.
6. You can make anything that looks like $z_1 + z_2\sqrt{z_3}$, where z_1, z_2 and z_3 are any numbers you generate in step 5.

7. Et cetera…

Thus, numbers like the following are “easily” obtainable:

$$\sqrt{\frac{314159265358}{2718281828}+\frac{\frac{17}{35\sqrt{\sqrt{\sqrt{1111}}}}-5\sqrt{46}}{1-\sqrt{\sqrt{1/13}+\frac{81}{91}\sqrt{\frac{3}{1198}\sqrt{\frac{11}{17}}+\sqrt{2}}}}}. \tag{1}$$

Another way of looking at it is that any number that can be expressed as a “tower” of square roots and arithmetic operations can be constructed with a straightedge and compass.

But let’s look at an example that’s a little simpler than expression 1, above. Consider:

$$x = 15+\sqrt{17+4\sqrt{\frac{3}{5}}},$$

where we’ve called the constructed number x.

The equation above can be converted as follows:

$$\begin{aligned} x-15 &= \sqrt{17+4\sqrt{\frac{3}{5}}} \\ (x-15)^2 &= 17+4\sqrt{\frac{3}{5}} \\ (x-15)^2-17 &= 4\sqrt{\frac{3}{5}} \\ ((x-15)^2-17)^2 &= 16\frac{3}{5}=\frac{48}{5}. \end{aligned}$$

The final equation above is just (if you multiply it out) an equation containing powers of x that are powers of 2, and exactly the same thing can be done with any such constructable number.

It may be obvious (but it certainly requires proof, and the proof is beyond the scope of this paper) that no matter how many square roots you apply, you can never get to a cube root (or a fifth root, or sixth root, or seventh root, or ninth root…). So if the number you would like to construct involves a cube root (or fifth root, or sixth root, …), you are out of luck—it cannot be constructed using the methods described in Section 3.

6 Three Impossible Constructions

The three classical construction problems of antiquity are known as “squaring the circle”, “trisecting an angle”, and “doubling a cube”. Here is a short description of each of these three problems:

- **Squaring the Circle.** Given a circle, construct a square that has exactly the same area as the circle.
- **Trisecting an Angle.** Given an angle, construct an angle whose measure is exactly 1/3 the measure of the original angle.
- **Doubling a Cube.** Given the length of the side of a (three-dimensional) cube, construct a length so that a cube with an edge of this length will have exactly double the area of the original cube.

It is beyond the scope of this paper to prove that these three problems are impossible to solve, but given the ideas that were presented in Section 5, we can at least indicate how the formal proofs work.

The easiest to show (to be impossible) is the problem of doubling the cube. Clearly if the volume of the required cube is double the volume of the original cube, if the length of the side of the original cube is L, the desired length is $\sqrt[3]{2}L$. In other words, from an algebraic point of view, we can construct the cube root of 2, or equivalently, we can solve the following equation:

$$x^3 = 2,$$

which is irreducible over the rationals, and is not going to be in any quadratic extension of the rationals.

The problem of trisecting an angle is similar, but what is usually done is to show that there is a particular angle that cannot be trisected, and that angle is typically chosen to be 60°. If we can't trisect this particular one, then we know that not all angles can be trisected. (Remember that a solution to the problem requires a general method that will work for any angle. Certain angles, such as 90° angles, can obviously be trisected, since we can construct a 30° angle from scratch. Also, we can construct a 60° angle from scratch, so if this one cannot be trisected, the general problem of trisection is clearly unsolvable.)

If we can trisect a 60° angle, that is equivalent to constructing a 20° angle from scratch, which is equivalent to constructing the cosine of 20°.

Now the cosine of 60° is $1/2$, so

$$\cos 60^\circ = \cos(20^\circ + 20^\circ + 20^\circ) = 1/2.$$

Using the standard formulas for the sine and cosine of sums of angles, this equation above can be converted to:

$$4\cos^3 20^\circ - 3\cos 20^\circ = 1/2,$$

or if $x = \cos 20^\circ$, to

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

which is also an irreducible cubic equation that cannot possibly have a solution in any quadratic extension of the rationals.

Much more difficult is to show that a circle cannot be squared. If the original circle has radius 1, its area will be π, so squaring a circle is equivalent to the construction of a length equal to $\sqrt{\pi}$. It turns out (but is not easy to prove) that π is transcendental—it is not the solution to *any* polynomial equation, so it is surely not the solution to any combination of quadratic polynomials.

7 Problems

Here are a few more construction problems to try your hand at.

1. Given a semicircle centered at a point C with diameter AB, find points I and J on AB, and points H and G on the semicircle such that the quadrilateral $GHIJ$ is a square.

2. Given a quadrant of a circle (two radii that make an angle of 90° and the included arc), construct a new circle that is inscribed in the quadrant (in other words, the new circle is tangent to both rays and to the quarter arc of the quadrant).

3. Given a point A, a line L that does not pass through A, and a point B on L, construct a circle passing through A that is tangent to L at the point B.

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

5. Given two points A and B that both lie on the same side of line L, find a point C on L such that $AC + BC$ is as small as possible. (Hint: This problem is related to the construction problem 4. Also remember that the shortest distance between two points is a line.)

6. Given two non-parallel lines L_1 and L_2 and a radius r, construct a circle of radius r that is tangent to both L_1 and L_2.