2. Ruler and Compass

Ancient Greek mathematics centered around proving results in geometry. For Plato only two shapes were considered perfect, the line and the circle. So, in the classical constructions of Greek mathematics, one was only allowed to use a ruler and a compass.

Unfortunately it was realised, quite early on, that certain constructions were seemingly impossible. Three such problems are especially famous:

(1) *Squaring the circle.* Starting a fixed square, construct a circle of the same area.
(2) *Duplication of the cube.* Starting with a fixed cube, construct a cube of twice the volume.
(3) *Trisection of the angle.* Starting with a fixed angle, construct a third of this angle.

It turns out that it is relatively easy, using the methods we have developed so far, to prove that it is impossible to carry out all three constructions.

We start by beginning the process of formalising the constructions of the ancient Greek geometers. We allow repeated applications of the following two constructions.

(1) Given two points P and Q, which have already been constructed, we can draw the line connecting P to Q.
(2) If we have constructed the point P and the length r, we can draw the circle with centre P and radius r.

By way of illustration, let us consider a typical construction of classical geometry. Suppose that we are given two lines L and M, that meet at a point P. The object is to the bisect the angle between the two lines at P.

The trick is to first observe the following. Given a line N and a point Q of N, we may find a line N' that passes through Q and makes a right angle with N. The line N' is known as the perpendicular bisector. The construction is as follows. Construct a circle of arbitrary radius through Q. This circle will meet N in two points, say R_1 and R_2, which, almost by definition, are equi-distant from Q. Now construct two circles, of the same radius, with centres at R_1 and R_2. Provided the radius of these two circles is greater than the radius of the original circle, these two circles will meet at two points Q_1 and Q_2. It is not hard to see that the line N' is precisely the line connecting Q_1 and Q_2.

Once we know how to construct perpendicular bisectors, it is straightforward to bisect any angle. Indeed, first start by drawing a circle with centre P of any radius. This circle will meet both lines in two points. Pick a pair of points Q_1 and Q_2, either side of the angle to be bisected. Let N_1 and N_2 be the two perpendicular bisectors of L and N at Q_1 and Q_2. N_1 and N_2 meet a unique point R. Draw the line N connecting P to R. N is easily seen to be the perpendicular bisector.

Now we turn to proving that the three constructions above are impossible. The first step is to further formalise the classical constructions. In all constructions, the key point is to keep track of the points we have constructed so far.

So we suppose that A represents the set of all points in the plane, and B denotes the set of all points that be constructed. Of course, it is natural to use Cartesian coordinates to refer to the points of the plane. So we suppose that B contains two points p and q. We arbitrarily assign unit length to the distance between these two points. Further we consider the line connecting p to q to be the x-axis, with p the origin and thing of the direction p to q as being positive. If we pick an orientation of the perpendicular bisector through p, then we have an x and a y-axis. In this case we may identify A with $\mathbb{R}^2$. With this choice of coordinates, $p = (0, 0)$ and $q = (1, 0)$. We let L denote the set of all possible lengths that can be constructed. That is $r \in L$ iff there are two points a and b belonging to B, such that the distance from a to b is r. By definition $L \subset \mathbb{R}$.

Proposition 2.1. *Let x and $y \neq 0$ be two elements of L. Then $x + y$, $x - y$, xy and x/y are elements of L.*

Proof. Using a compass we can mark off the point $(x, 0)$ on the x-axis. The circle with centre $(x, 0)$ meets the x-axis at the two points $(x-y, 0)$ and $(x + y, 0)$. It follows that $x \pm y \in L$.

Now pick any line L of slope positive slope passing through the origin and mark of the point R of distance y lying on this line, in the first quadrant. Consider the line M connecting $(1, 0)$ to R. Construct a line N, parallel to M, passing through $(x, 0)$ (it is easy to construct parallel lines, simply by constructing two perpendicular bisectors). This line will meet L in a point S. It is easy to see, using the theory of similar triangles, that the distance of S to the origin is precisely xy. Thus $xy \in L$.

The construction of x/y proceeds in a similar fashion. □

Corollary 2.2. *L is a field and $B = L^2 \subset \mathbb{R}^2$.*

Proof. The fact that L is a field follows immediately from 2.1. Clearly $L^2 \subset B$. On the other hand, pick a point $(a, b) \in L^2$. Then a and b belong to L and we can mark off, using a compass, the points $(a, 0)$

and $(0, b)$. The intersection of the perpendicular bsicectors of the x and y-axis, passing through $(a, 0)$ and $(0, b)$ is then the point (a, b), so that, by construction, $(a, b) \in B$. □

Thus to determine B it suffices to determine the subfield L of $\mathbb{R}$.

Lemma 2.3. *Let L be the line connecting two points p and q and C a circle containing p of radius r.*

If the coordinates of p and q and the number r lie in a field K, then so do the coefficient of a definining equation for L and C.

Proof. Easy and well-known. □

Here is the main result.

Proposition 2.4. *Suppose that we are given a subfield K of the real numbers, two lines L and M and two circles C and D, all of whose coefficients lie in K. If $P = (u, v)$ is one point of the following sets,*

(1) $L \cap M$, *or*
(2) $L \cap C$, *or*
(3) $C \cap D$,

then u and v belong to the same quadratic extension of K.

Proof. Suppose that L is given as $ax + by = c$ and M is given as $dx+ey = f$, where by assumption a, b, c, d, e and $f \in K$. Then we may solve for x and y, using the standard methods of Gaussian elimination. As the steps of Gaussian elimination only involve addition, subtraction, multiplication and division, it follows that the solution lies in K. Hence (1).

As K is a field, we are free to magnify, translate and rotate our coordinate frame, using a matrix with entries in K. So we may translate the centre of our circle back to the origin and rotate so that the line L is parrallel to the x-axis. In this case the equations for L and C are then

$$\begin{aligned} y &= a \\ x^2 + y^2 &= b. \end{aligned}$$

Thus $v \in K$ and $u^2 \in K$, so that $K(u, v) = K(u)$ has degree two over K. Hence (2).

As before, we may assume that the centre of the first circle is $(0, 0)$ and that the centre of the second is $(a, 0)$. The two equations are then

$$\begin{aligned} x^2 + y^2 &= b \\ (x-a)^2 + y^2 &= c. \end{aligned}$$

Subtracting one from the other we get

$$2ax = b - c + a^2,$$

so that $u \in K$, and $v^2 \in K$. Thus $K(u, v) = K(v)$ has degree two over K. ☐

Lemma 2.5. *Let L/K be a quadratic field extension and suppose that the characteristic of K is not equal to two.*

Then we may find $\alpha \in L$ such that $L = K(\alpha)$ and $\alpha^2 \in K$.

Proof. Pick any $\beta \in L$ not in K. Then $K(\beta) \neq K$ so that by the tower law, $L = K(\beta)$. Thus the minimum polynomial of β must be quadratic. Suppose that it is

$$x^2 + bx + c.$$

Completing the square, we get

$$(x + b/2)^2 + c - b^2/4.$$

Thus

$$\alpha^2 = b^2/4 - c,$$

where $\alpha = \beta + b/2$. Now $\alpha \notin K$, so that $L = K(\alpha)$ whilst $\alpha^2 \in K$. ☐

Corollary 2.6. *$a \in L$ iff there are $\alpha_1, \alpha_2, \dots, \alpha_n$ such that*

$$a \in \mathbb{Q}(\alpha_1, \alpha_2, \dots, \alpha_n),$$

where $\alpha_i^2 \in \mathbb{Q}(\alpha_1, \alpha_2, \dots, \alpha_{i-1})$.

Proof. One direction is an easy induction, using 2.3, 2.4 and 2.5 and the observation that the two basic classicial operations only allow us to construct lines and circles and to intersect some combination of these.

For the other direction, note that it is easy to construct square roots, as in the proof of 2.4. ☐

Corollary 2.7. *If $a \in L$ then a is algebraic and the minimum polynomial of a is a power of two.*

Proof. Let E be the field $\mathbb{Q}(\alpha_1, \alpha_2, \ldots, \alpha_n)$, whose existence is guaranteed by 2.6 and set $F = \mathbb{Q}(a)$. By the Tower law and induction,

$$[E : \mathbb{Q}]$$

is a power of two. Again by the tower applied to the intermediary field F, $[F : \mathbb{Q}]$ is finite and divides a power of two, whence it is a power of two. But the degree of a primitive field extension is nothing more than the degree of the minimal polynomial. □

Theorem 2.8. *It is impossible to square the circle.*

Proof. Start with a square of side 1, so that the area is 1. We would need to construct a circle of radius $1/\sqrt{\pi}$. So if we could square the circle, it then follows that $\sqrt{\pi} \in L$ and hence $\pi \in L$.

But then π would be algebraic, a contradiction. □

Theorem 2.9. *It is impossible to duplicate the cube.*

Proof. Suppose we start with a cube of side one. Then we want to construct a cube of side 2, so that we need to construct a length of $a = \sqrt[3]{2}$. Now a is a root of the polynomial $x^3 - 2$. By Eisenstein, applied with $p = 2$, this is irreducible. Thus $x^3 - 2$ is the minimum polynomial of a. Thus we cannot duplicate the cube by 2.7. □

Theorem 2.10. *It is impossible to trisect an angle.*

Proof. Suppose we start with an angle of $\pi/3$. Then we want to construct an angle of $\theta = \pi/9$. This is equivalent to constructing the point on the unit circle, with coordinates $(a, b) = (\cos\theta, \sin\theta)$.

Now

$$\begin{aligned} \cos(\phi + \psi) &= \cos\phi\cos\psi - \sin\phi\sin\psi \qquad \text{and} \\ \sin(\phi + \psi) &= \cos\phi\sin\psi\sin\phi\cos\psi \end{aligned}$$

so that

$$\begin{aligned} 1/2 &= \cos 3\theta \\ &= \cos 2\theta\cos\theta - \sin 2\theta\sin\theta \\ &= (\cos^2\theta - \sin^2\theta)\cos\theta - 2\sin^2\theta\cos\theta \\ &= \cos^3\theta - 3\sin^2\theta\cos\theta \\ &= 4\cos^3\theta - 3\cos\theta \end{aligned}$$

Thus $2a$ satisfies the equation

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

I claim that this polynomial is irreducible. As it is a cubic, it is reducible iff it has a root. By Gauss' Lemma it suffices to check that it has no roots over $\mathbb{Z}$. The only possibility is ± 1, which it is easy to see are both not roots.

Thus it is impossible to trisect the angle, by 2.7. $\square$