Algebra Notes

Oct. 28: Impossible Constructions

Geoffrey Scott

Today, we'll prove that certain constructions are impossible with a straightedge and compass. We begin by describing a procedure for taking any sequence of straightedge and compass operations, and associating to it a sequence of field extensions of $\mathbb{Q}$ – whenever we draw a new point as one of our operations, we adjoin the x and y coordinates of the point. We will show that these field extensions are always degree 2. However, some geometric constructions, such as trisecting certain angles, necessitate constructing a field extension whose degree is divisible by 3, which we will show cannot occur as a sequence of degree-2 extensions.

Constructible numbers $\iff$ constructible coordinates

Consider the proceduce of starting with the points (0, 0) and (1, 0). Obviously, the x-axis and y-axis are constructible lines and every rational number is a constructible number. The following fact is less obvious.

Claim: The point (x, y) is a constructible point if and only if x and y are constructible numbers.

Proof:

$\Rightarrow$: If (x, y) is a constructible point, draw the lines through it that are parallel to the x and y axes. These lines will intersect the axes at the points $(x, 0)$ and $(0, y)$, which are at distance $|x|$ and $|y|$ to the origin, respectively.

$\Leftarrow$: If x and y are constructible, then drawing circles of radius $|x|$ and $|y|$ around the origin and intersecting this circle with the x and y axes, we construct the points $(x, 0)$ and $(0, y)$. Drawing the lines through these two points parallel to the axes and intersecting them, we construct the point (x, y).

Now imagine you're trying to find all the constructible numbers. You already know that every rational number is constructible, so by the above claim, every point whose coordinates are rational. To find more constructible points, you try to find a way to construct new points with non-rational coordinates. You know that the only way of doing this is by intersecting lines and/or circles together. But each of these lines and/or circles must be "drawn using points having coordinates in $\mathbb{Q}$". That is, each such line must go through two points with rational coordinates, and each such circle must have a point with rational coordinates as its centre, and its radius must be equal to the distance between two points with rational coordinates.

Suppose we succeed in constructing a point (a, b), whose coordinates are not both rational (so a genuinely new point). Because we know that the set of constructible numbers is a field, we know it must be $\mathbb{Q}(a, b)$. Now in order to find more constructible numbers, you repeat the entire procedure of the above paragraph, except you replace every instance of the phrase "rational number" with "number in $\mathbb{Q}(a, b)$". To understand this procedure, we need to understand, for a given subfield F of the real numbers, what the equations of intersection for lines and circles look like when the lines and circles are "drawn using points having coordinates in F".

Equations for Lines and Circles

Let's remind outselves of the equations for lines and circles in $\mathbb{R}^2$. Let F be a subfield of $\mathbb{R}$.

Line: If (x_1, y_1) and (x_2, y_2) are two distinct points with $x_1, y_1, x_2, y_2 \in F$, then the line through them is given by the equation

$$(y_1 - y_2)x + (x_2 - x_1)y + y_2x_1 - x_2y_1 = 0$$

From this expression, one can deduce that the lines that can be drawn "using points in F" are precisely the lines with equations of the form

$$ax + by + c = 0 \qquad \text{where } a, b, c \in F, \text{where } a \neq 0 \text{ or } b \neq 0$$

Circle: If d is the distance between (x_1, y_1) and (x_2, y_2), and (x_3, y_3) is a third point, with $x_1, y_1, x_2, y_2, x_3, y_3 \in F$, then the circle through (x_3, y_3) of radius d is given by the equation

$$(x - x_3)^2 + (y - y_3)^2 = d^2$$

From this expression, one can deduce that the circles that can be drawn "using points in F" are precisely the circles with equations of the form

$$x^2 + y^2 + ax + by + c = 0 \qquad \text{where } a, b, c \in F$$

Now let's study what happens when we take lines or circles of this form and intersect them. *What extension field of F will the coordinates of these new points live in?*

Intersecting Lines and Circles

Proposition: Let F be a subfield of $\mathbb{R}$, and suppose two lines with equations

$$\begin{aligned} a_1x + b_1y + c_1 &= 0, \qquad & a_1, b_1, c_1 \in F \\ a_2x + b_2y + c_2 &= 0, \qquad & a_2, b_2, c_2 \in F \end{aligned}$$

intersect at a point (p, q). Then $p, q \in F$.

Proof: Solve one of the equations for y in terms of x, and then substitute it into the other equation. This gives a linear polynomial with coefficients in F whose root equals p, so $p \in F$. Solving for one of the equations for x in terms of y and following the same logic yields $q \in F$.

Proposition: Let F be a subfield of $\mathbb{R}$, and suppose a line and circle with equations

$$\begin{aligned} a_1x + b_1y + c_1 &= 0, \qquad & a_1, b_1, c_1 \in F \\ x^2 + y^2 + a_2x + b_2y + c_2 &= 0, \qquad & a_2, b_2, c_2 \in F \end{aligned}$$

intersect in a point (p, q). Either $p, q \in F$, or $F(p, q)$ is a degree-2 field extension of F.

Proof: First suppose $b_1 \neq 0$. Then solve the line equation for y in terms of x, and substitute it into the circle equation. This shows that the x-coordinate of intersection, p, is the root of a degree-2 polynomial with coefficients in F, so $[F(p) : F]$ is either 2 or 1. Plugging p into the x-variable of the line equation gives a linear polynomial (with variable y) with coefficients in $F(p)$ that has q as a root. Therefore, $q \in F(p)$, so $F(p, q) = F(p)$. If $b_1 = 0$, then repeat the proof above with the roles of x and y reversed.

Proposition: Let F be a subfield of $\mathbb{R}$, and suppose two circles with equations

$$\begin{aligned} x^2 + y^2 + a_1x + b_1y + c_1 &= 0, \qquad & a_1, b_1, c_1 \in F \\ x^2 + y^2 + a_2x + b_2y + c_2 &= 0, \qquad & a_2, b_2, c_2 \in F \end{aligned}$$

intersect in a point (p, q). Either $p, q \in F$, or $F(p, q)$ is a degree-2 field extension of F.

Proof: By subtracting the second circle equation from the first, we see that (p, q) must be on the line

$$(a_1 - a_2)x + (b_1 - b_2)y + (c_1 - c_2) = 0$$

Then apply the previous proposition.

Associating field extensions to straightedge and compass operations

Suppose we start with a collection of points in $\mathbb{R}^2$ that have coordinates in $\mathbb{Q}$. If we perform a sequence of straightedge and compass operations on these points, we can construct a sequence of field extensions of $\mathbb{Q}$ – every time you draw a point (p, q), adjoin p and q to the current field F. By the above three propositions, any point you draw at the intersection of these lines or circles will correspond to a field extension of degree 1 or 2.

Trisecting the Angle is Impossible

Last class, we studied examples of complicated geometric operations that can be performed as a sequence of the three basic geometric operations, such as constructing parallel lines, perpendicular lines, and equilateral triangles. When people say "It is impossible to trisect an angle using a straightedge and compass," they mean "There exist constructible angles θ for which the angle $\theta/3$ is not constructible." In other words, the geometric operation "trisecting an angle" is *not* an operation that can be performed as a sequence of the three basic geometric operations.

Beware: Make sure you don't confuse the phrase "There exist constructible angles θ for which the angle $\theta/3$ is not constructible" (which is a true statement) with the phrase "For every constructible angle θ, the angle $\theta/3$ is not constructible" (which is a false statement).

Proposition: If an angle θ is constructible, then so are the numbers $\cos(\theta)$ and $\sin(\theta)$.

Proof: Suppose the angle is formed by the lines L_1 and L_2 meeting at the point c. Draw a circle of radius 1 around c, and suppose it meets L_2 at the point q. The line perpendicular to L_1 through q will intersect L_1 at a point $\cos(\theta)$ away from c and $\sin(\theta)$ away from q.

Proposition: The angle $\pi/3$ is constructible, but the angle $\pi/9$ is not constructible.

Proof: The fact that the angle $\pi/3$ is constructible is easy: take two points of distance 1 apart, draw a line through them and draw circles of radius 1 around each of them. Then, draw a line through the intersection point of the circles with either of your original points.

The fact that $\pi/6$ is *not* constructible relies on two facts that I will ask you to believe (I will prove the second fact if there is time). The first fact is a little-known trigonometric identity called the triple angle formula:

$$\cos(3\theta) = 4\cos^3(\theta) - 3\cos(\theta)$$

And the second fact is that the polynomial $4x^3 - 3x - 1/2 \in \mathbb{Q}[x]$ is irreducible.

If we apply the triple angle formula to $\theta = \pi/9$, we see that

$$1/2 = 4(\cos(\pi/9))^3 - 3(\cos(\pi/9))$$

So $\cos(\pi/9)$ is a root of the irreducible polynomial $4x^3 - 3x - 1/2 \in \mathbb{Q}[x]$, so $[\mathbb{Q}(\cos(\pi/9)) : \mathbb{Q}] = 3$.

Now assume towards a contradiction that we can find a sequence of straightedge and compass operations that constructed the angle $\pi/9$. Then we could construct the number $\cos(\pi/9)$, and the point $(\cos(\pi/9), 0) \in \mathbb{R}^2$. This means that the sequence of field extensions corresponding to these operations

$$\mathbb{Q} \subseteq F_1 \subseteq F_2 \subseteq \cdots \subseteq F_n$$

has $\cos(\pi/9) \in F_n$. But this is impossible, each extension has degree 2 and $[\mathbb{Q}(\cos(\pi/9)) : \mathbb{Q}] = 3$.

Constructing a regular 7-gon is impossible

Proposition: It is impossible to construct a regular 7-gon.

Proof: Assume towards a contradiction that is is possible to construct a regular 7-gon. The interior angle of a 7-gon is $5\pi/7$, so it is possible to construct the angle $\pi - 5\pi/7 = 2\pi/7$, and therefore it is possible to construct $\cos(2\pi/7)$. A true fact about the number $\cos(2\pi/7)$ (which we will not prove) is that it is a root of the polynomial

$$8x^3 + 4x^2 - 4x - 1 \in \mathbb{Q}[x].$$

Using the same technique as in the last section, we can check that this polynomial has no roots in $\mathbb{Q}$ so is irreducible in $\mathbb{Q}[x]$. Therefore, $[\mathbb{Q}(\cos(2\pi/7)) : \mathbb{Q}] = 3$, so $\cos(2\pi/7)$ cannot be contained in any field extension obtained through straightedge and compass operations.

Squaring the circle is impossible

Definition: Let E be a field extension of F, and $\alpha \in E$. If there is a polynomial in $F[x]$ that has α as a root, then α is called **algebraic over** F. Otherwise, k is **transcendental over** F. If every element of E is algebraic over F, then the field extension E of F is called an **algebraic** extension; otherwise it is called a **transcendental** extension.

For example, every n^{th} root of a rational number is algebraic over $\mathbb{Q}$ because it will be a root of a polynomial of the form $x^n - a \in \mathbb{Q}[x]$. On the other hand, π, e, e^a (where a is any nonzero algebraic number) are all transcendental over $\mathbb{Q}$, though this is hard to prove. The set of real numbers which are algebraic over $\mathbb{Q}$ is countable, so the set of real numbers which are transcendental over $\mathbb{Q}$ is uncountable. In this sense, there are "more" transcendental numbers than real numbers.

Proposition: If α is transcendental over F, then $F(\alpha)$ has infinite degree over F.

Proof: Assume towards a contradiction that $[F(\alpha) : F]$ is some finite number n. Then the set $\{1, \alpha, \ldots, \alpha^n\}$ has $n + 1$ elements in it, so must have some equation of linear dependence

$$a_0 1 + a_1\alpha + a_2\alpha^2 + \cdots + a_n\alpha^n \qquad \text{where } a_0, \ldots, a_n \in F.$$

Then α is a root of the polynomial $f(x) = a_0 + a_1x + \cdots + a_nx^n \in F[x]$, a contradiction to the fact that α is transcendental.

We will use the fact (without proof) that π is transcendental in the next proposition.

Proposition: It is impossible to construct a square whose area equals the area of a circle of radius 1.

Proof: If it were possible, we could construct a sequence of field extensions

$$\mathbb{Q} \subseteq F_1 \subseteq F_2 \subseteq \cdots \subseteq F_n$$

such that F_n contains the number $\sqrt{\pi}$ (this is the side length of a square with area π). But then F_n contains the field $\mathbb{Q}(\pi)$, and $[\mathbb{Q}(\pi) : \mathbb{Q}]$ is infinite. But $[F_n : \mathbb{Q}]$ is not infinite.