Math/CS 120: Intro. to Math — Professor: Padraic Bartlett

Lecture 6: Constructible Numbers

Weeks 7-8 — *UCSB 2014*

> As the geometer who sets himself
> To square the circle and who cannot find,
> For all his thought, the principle he needs,
> Just so was I on seeing this new vision.
>
> Dante Alighieri, Paradise, canto XXXIII

1 Constructible Numbers

Construction, in a sense, was the unofficial theme of the first five weeks of this class. We started our lectures by first defining how to make proofs; from there, we used our language of "proof" to construct $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$.

However, the method in which we constructed these objects was remarkably abstract; we started with nothing,[1] and in the words of János Bolyai, "Out of nothing [we] have created a strange new universe." While this was beautiful in its simplicity — given absolutely nothing but a small handful of logical axioms, we were able to build our way to the foundations of modern mathematics in a few weeks — it also makes one wonder if there are perhaps different or easier ways to approach the same constructions. That is: when we conceive of the number 4, we usually don't think of

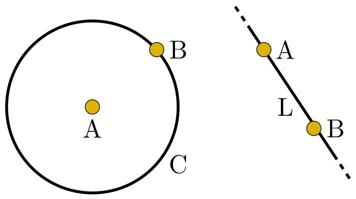
$$\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}\}.$$

These notes are meant to offer a different construction for many of the numbers we care about: the **constructible numbers**! We define them as follows:

Definition. Suppose that you are given an infinite piece of paper (think of this as $\mathbb{R}^2$,) and two tools:

1. A **compass**. Given any two points A, B on our paper, this device will let us draw the unique circle centered at A that passes through the point B.

2. An (infinitely long) **straightedge**. Given any two distinct points A, B on our paper, this device will let us draw the unique straight line that passes through both A and B.



[1] Or, to be very precise, the set containing nothing.

Given these two tools, what can we do? Well: to do anything we need to apparently have some points! Let's give ourselves a pencil as well:

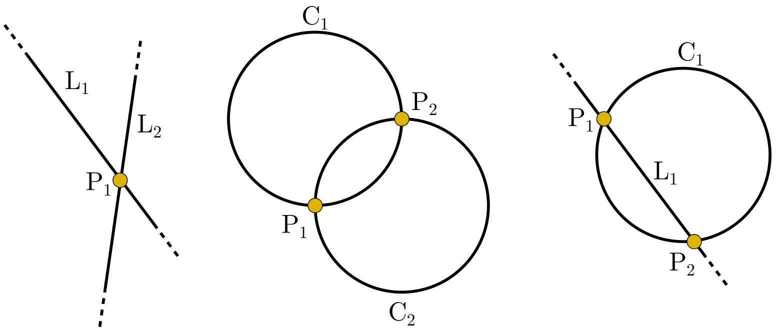
3. A **pencil**[2]. We can use a pencil "freehand" to sometimes draw points when we don't really care where they are. More precisely, we can do the following:
   - We can place a point at random on our paper.
   - Given any line or circle A, we can pick either "side" of A (i.e. left or right/inside our outside), and put a point at random on that side of A.
   - Given any point X on a line L, we can pick either "side" of X, and put a point at random on that side of X that is on the line L.
   - Given any two points X, Y on a circle C, we can pick either "left of X" or "right of X," and put a point at random between X, Y with that relative positioning to X.
   - Given any finite combination of the above claims that has a solution, we can place points at random that satisfy all of the desired claims.

In particular, this lets us make the following claims:

4. Given any two lines l_1, l_2 that intersect at some unique point, we can label that point of intersection.

5. Given any two circles C_1, C_2 that intersect at one or two points, we can label those points of intersection.

6. Given any circle C and line l that intersect at one or two points, we can label those points of intersection.



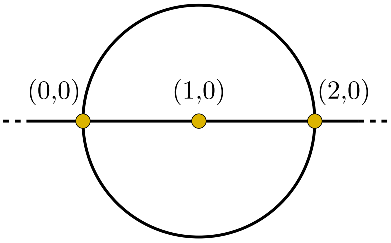
Excellent! We have our ground rules for what we can and cannot create. The only thing that's missing at the moment is some place to **start**. So, let's start! Take our paper, and mark two distinct points on it:

[2] For future reference; most people skip this step, and just assume these operations are "intuitive." As we've seen thus far in mathematics, however, seemingly "intuitive" things are often very strange when looked at up close! This is why we write our our claims here.

Call one of them $(0, 0)$, and the other $(1, 0)$, and think of them as you would think of the two points $(0, 0), (1, 0) \in \mathbb{R}$.

These names motivate the following question: what **other** points can we make? For example, we can construct the point $(2, 0)$ by doing the following:

1. Draw the line L through $(0, 0), (1, 0)$.
2. Draw a circle C centered at $(1, 0)$ through $(0, 0)$.
3. There is a second point P at which L and C intersect, because L goes through the center of C (and is thus not tangent to C.) The distance of P from $(0, 0)$ is exactly twice the radius of C, i.e. 2, and P is also distance 1 from the center of C, which is $(1, 0)$; moreover, because P is on a line between two points with zero y-coördinate, the y-coördinate of P is 0 as well. Consequently, P must be $(2, 0)$!



In fact, via induction, we can make **any** point $(n, 0)$, for any $n \in \mathbb{N}$! To make this rigorous, we make the following claim:

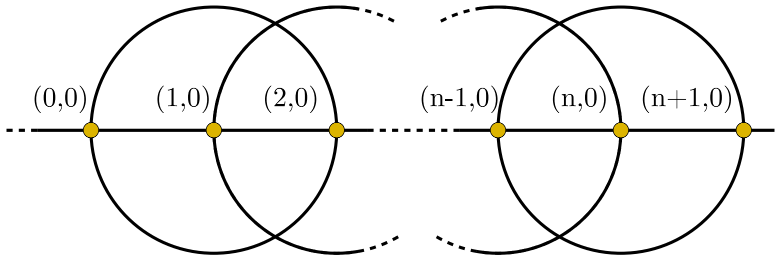
Definition. A real number $r \in \mathbb{R}$ is called **constructible** if there is a finite sequence of compass-and-straightedge constructions that, when performed in order, will always create a point P with at least one coördinate equal to r.

We showed above that 2 is constructible, and claim that n is constructible here:

Theorem. All of the elements of $\mathbb{N}$ are constructible.

Proof. We proceed by induction. Our base cases are already established above; it suffices then to proceed to the inductive step. Assume, for induction, that we have succeeded in constructing all of the natural numbers between 1 and n; specifically, suppose that we constructed these numbers as $(k, 0)$'s for $0 \leq k \leq n$. We seek to prove that we can construct $n + 1$ as well.

To do this, simply construct the line L through $(n - 1, 0)$ and $(n, 0)$, and also draw the circle C centered at $(n, 0)$ through $(n - 1, 0)$. As before, there is a second point P at which L and C intersect, because L goes through the center of C (and is thus not tangent to C.) The distance from P to $(n - 1, 0)$ is 2 and the distance from P to $(n, 0)$ is 1, because C is a circle of radius 1. Finally, P's y-coördinate is 0, because L is a line through multiple points with 0 y-coördinate; therefore, P is precisely the point $(n + 1, 0)$, and we have proved our claim.



□

Identical logic starting with $(0,0),(1,0)$ and working our way backwards lets us extend this result to the integers:

Theorem. All of the elements of $\mathbb{Z}$ are constructible.

Thus far in our constructions, we've limited ourselves to effectively working with the x-axis. This seems like an unnecessary limitation! So, let's attempt to prove some useful lemmas that will let us travel in the y-direction as well. One natural object to hope to construct in geometry is a **perpendicular line**. We can do this as follows:

Theorem. Suppose that L is a line, and A is a point not on L. Then we can construct a line M through A that is **perpendicular**[3] to L; that is, we can create a line M such that the two lines M, L intersect at an angle of $\pi/2$ radians.

Proof. First, before reading this proof, go and try to construct this yourself! This is generally good advice for most math texts, but it's especially relevant here; the interesting thing about these geometric constructions is often not seeing what works, but rather in playing around with a number of failed constructions and seeing why they all do **not** work!

With that said; let's create a perpendicular! To do this, we need to create a line, and thus need a second point to connect A to. There are many constructions one can use here; perhaps the simplest relies on the following intuitive idea:

The line connecting A to its "reflection" through L is perpendicular to L.

This isn't something that we know yet (in fact, we're about to prove it!,) or is even something that we really know what the words mean here (what is a "reflection?") However, this should be *somewhat* intuitive; even without definitions, we have an idea of what it would mean to "flip" an object over a line, and drawing a few test cases should persuade yourself that this claim might be true.

The nice thing about this reflection process is that it's easily done with our tools:

1. Pick at random some point X_1 on our line L.

2. Draw the circle C_1 centered at X_1 through A.

[3] There are many equivalent notions for perpendicular that can be used here; one of the more natural ones to choose is "find the closest point P on L to A. Then the line through P and A is perpendicular to L. You could use this definition instead in the proof below and get a solution as well, if you wanted! Try it out.

3. The line L passes through the center of this circle, and thus has two intersection points with our line L. Pick either one of them, and call that point X_2.

4. Draw the circle C_2 centered at X_2 through A.

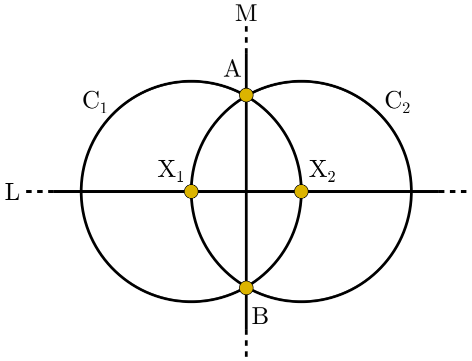
5. Notice that if two circles C_1, C_2 with centers on the same line L are tangent, then this point of tangency must occur on the line L. To see this, simply rotate space so that this line L is the x-axis to simplify coördinates, and assume our circles are centered at some points $(a_1, 0), (a_2, 0)$ with radii r_1, r_2. Then, in this situation, saying that these two circles are tangent is equivalent to claiming that the system of equations

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

has a unique solution. However: if we had any solution (x, y) for this system where y was nonzero, then $(x, -y)$ would also be a solution, because we cannot distinguish y^2 from $(-y)^2$! Therefore, the uniqueness of this solution forces it to be on the x-axis. The act of rotating space doesn't change our observations here; therefore, our original claim still holds.

6. Consequently, because our two circles C_1, C_2 intersect at A, which is not on L, they intersect at some other point B as well!

7. Draw the line M through A and B.



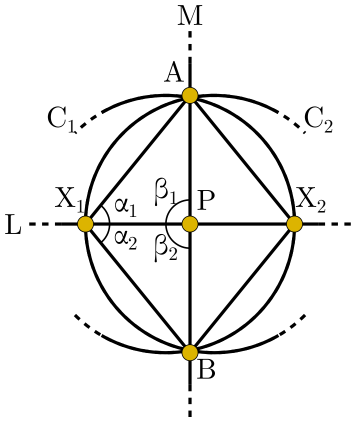
I claim[4] that M is our desired perpendicular line! To see this, draw the line segments connecting X_1, X_2 to A, B, label the intersection of the lines L, M as P, and make the following observations:

1. The triangle $\Delta X_1 A X_2$ is similar to the triangle $\Delta X_1 B X_2$; this is because they have the same side lengths (one with the radius of C_1, one with the radius of C_2, and one with length $\overline{X_1 X_2}$.)

[4] Notice how our geometric proofs have two stages; we first construct an object, and then prove that it has the desired properties. Proofs need both parts to count; don't forget either when solving problems!

2. Therefore, the angles $\alpha_1 = \angle AX_1P$, $\alpha_2 = \angle BX_1P$ are the same. So we can say that the triangles ΔAX_1P, ΔBX_1P are similar, as they share two sides and an angle in common.

3. Consequently, the angle $\beta_1 = \angle APX_1$ is equal to the angle $beta_2 = \angle BPX_1$. But the sum of these two angles is $\pi = 180°$, as M is a straight line passing through A, P and B. Therefore any one of these angles is $\pi/2 = 90°$. In other words, the lines M, L are orthogonal, as claimed.



□

A related result that we might want is the ability to construct **parallel lines**:

Theorem. Suppose that L is a line, and A is a point not on L. Then we can construct a line M through A that is **parallel** to L; that is, we can create a line M such that the two lines M, L never intersect.

Proof. This is much easier now that we can make perpendicular lines!

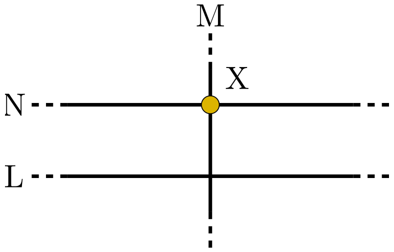
1. Take our line L.

2. Pick any point X not on L.

3. Draw a line M perpendicular to L through X.

4. If this line passes through A, we have really bad luck. No worries, though; simply pick a point X' not on M, and draw the perpendicular to L through X'. Delete our old X, M, and call this pair the new X, M.

5. Now, we have that M is perpendicular to L and that A is not on M. Therefore, we can construct a line perpendicular to M through A! Do so, and call this line N.



I claim that N is parallel to L. To see why: note that the angle between L and M is $\pi/2$, and the angle between M and N is $\pi/2$; consequently the angle between L and N is $\pi/2 \pm \pi/2 = 0$ or π!

Also: notice how in the proof above we simply used "make a perpendicular line" as a construction we can make. This is because even though we didn't list it as an axiom, we proved it was possible in our theorem earlier! This is a common technique in mathematics; once you've proven you can do something, you can simply just apply that result without having to constantly reprove it. □

Another useful result you might want, based on your own experience with compasses, is a compass that "stays open" after you use it. In other words, suppose that you could place your compass across any two existing points, and transfer that length to anywhere else on your paper! This would be useful. We formalize this claim here:

Theorem. ("Non-Collapsing Compass:") Given any point A and any two other points B, C, we can draw a circle around A with radius equal to the distance between B and C. (In other words, we can "transfer" the circle centered at B through C over to A!)

Proof. On the HW! □

Another result you might want is a way to convert between our notion of constructible "values" (i.e. elements of $\mathbb{R}$) and "points" (i.e. elements in $\mathbb{R}^2$.) This theorem does that for us:

Theorem. Suppose that the two values a, b are constructible. Then we can create the points $(\pm a, \pm b)$ using our compass and straightedge.

Proof. If a is constructible, there is some point P with one coordinate equal to a, and other coordinate equal to some value c. Assume without loss of generality that $P = (a, c)$ (our proof will work equally well in the other case.) Draw the perpendicular line through P and the x-axis: this line's intersection with the x-axis is at the point $(a, 0)$. By drawing a circle centered at the origin through this point, we can now create both $(\pm a, 0)$ and $(0, \pm a)$.

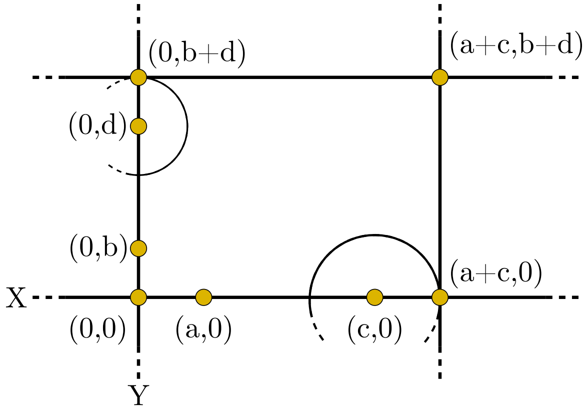
Similarly, if b is constructible, we can make $(\pm b, 0)$ and $(0, \pm b)$ via the same process. By adding points with our theorem earlier, we can therefore make $(\pm a, \pm b)$ as claimed. □

Using these tools, we can do the following:

Theorem. Suppose that we can construct the points (a, b) and (c, d) with a compass and straightedge. Then we can create $(a + c, b + d)$.

Proof. Because a, b, c, d are all constructible, we can make the points $(a, 0), (c, 0), (0, b), (0, d)$. Take the point $(a, 0)$, and (by using our non-collapsing-compass construction) draw a circle of radius given by the distance from $(c, 0)$ to the origin. This circle has radius c by construction, and intersects the x-axis twice, as its center is on the x-axis; the two points of this intersection are $(a \pm c, 0)$. In a similar process, we can construct $(0, b \pm d)$ as well. Do so.

Now, draw a line parallel to the x-axis through $(0, b + d)$; this gives us all points with coördinates of the form $(x, b + d)$. Similarly, draw a line parallel to the y-axis through $(a + c, 0)$; this gives us all points with coördinates of the form $(a + c, y)$. The intersection of these two lines is at $(a + c, b + d)$; so we have constructed the sum of these two points!



□

Theorem. Suppose that we can construct the point (a, b), and also suppose that we can construct the value x, for $x \neq 0$. Then we can construct (xa, xb) and $(a/x, b/x)$.

Proof. First, assume that (a, b) is not on the x-axis. To do this, simply notice that if it is, we can "rotate" it off of the axis by drawing a circle through (a, b) centered at the origin, and taking any intersection of that circle with any other non-x-axis line. If we can scale this new point P by x, then draw a circle through xP centered at the origin, and take its intersection Q with the line L through (a, b) and the origin: this Q is the desired (xa, xb), as it is on the line L (and thus the ratio of x to y-coordinate is correct) and its distance has been scaled from the origin by x.

With this assumption made, consider the following process:

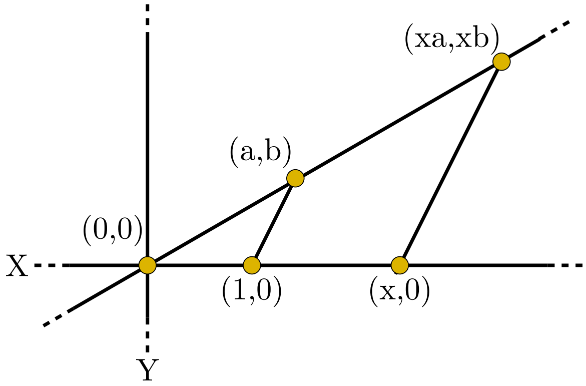
1. Draw the line L connecting (a, b) to $(0, 0)$.

2. Also draw the line M connecting (a, b) to $(1, 0)$.

3. Because x is constructible, by our lemmas above, we can construct $(x, 0)$. Do so.

4. Draw a line M' parallel to M through $(x, 0)$.

5. Label the origin as C, the point (a, b) as D, the point $(1, 0)$ as E, the point of intersection of L, M as F, and the point $(r, 0)$ as G.



I claim that F is (xa, xb). To see why, simply notice that the two triangles $\Delta CDE, CFG$ are similar, as they share the same angles (because M, M' are parallel!) Therefore, their side lengths are proportional: that is, the ratio of the length of (a, b) to $(1, 0)$ is the same as the ratio of F to $(x, 0)$. In other words, F is a point on the line L (and therefore one whose ratio of x to y-coördinate is x/y) that is x times as long as (a, b): i.e. $F = (xa, xb)$, as claimed.

To see how we divide points: suppose instead that we were given the point $F = (xa, xb)$ at the start, and needed to construct $D = (a, b)$! We can do the same process: namely,

1. Draw the line L connecting (xa, xb) to $(0, 0)$.

2. Because x is constructible, by our lemmas above, we can construct $(x, 0)$.

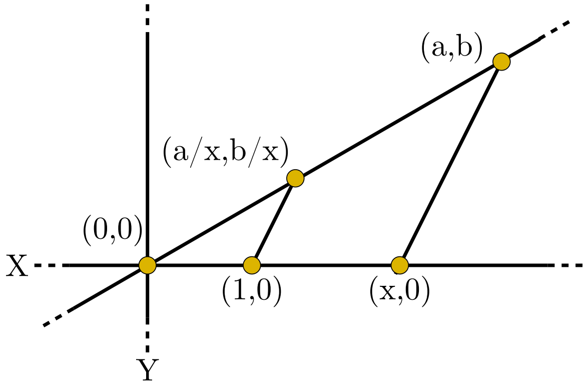
3. Draw the line M' connecting (xa, xb) to $(x, 0)$.

4. Now, draw a line M parallel to M' through $(1, 0)$.

5. Label the intersection of M, L as D.

By the exact same logic as above, D must equal (a, b)!

Therefore, we can also divide by x: if we start with $F = (a, b)$ and proceed with our construction above, connecting F to $(x, 0)$, we will create a point D with coördinates $(a/x, b/x)$, as desired.



□

This last result gives us the following useful pair of corollaries:

Corollary. The collection of all constructible numbers forms a field.

Proof. Because these numbers are a subset of $\mathbb{R}$, have the same addition/multiplication operations, and contain 0,1 by assumption, we get all of the field axioms (associativity $(,\cdot)$, commutativity $(,\cdot)$, distributivity $(,\cdot)$, identity $(+,\cdot)$.) We just need closure and inverses for $+,\cdot$; but the results above have proven that we have both of those as well! So we form a field. □

Corollary. $\mathbb{Q}$ is constructible.

Proof. The rational numbers are the smallest field that contain the integers, as any field containing the integers must contain all of their reciprocals (i.e. all $\frac{1}{n}$'s for $n \neq 0$); being closed under multiplication then gives us that this field must contain $\mathbb{Q}$. □

However, $\mathbb{Q}$ isn't the only thing we can construct! Consider the following result:

Theorem. For any n, we can construct $\sqrt{n}$.

Proof. We prove this by induction. We know that $\sqrt{1} = 1$ is constructible by definition, which takes care of our base case; we are then left with the inductive step.

Suppose that $\sqrt{n}$ is constructible. Perform the following steps:

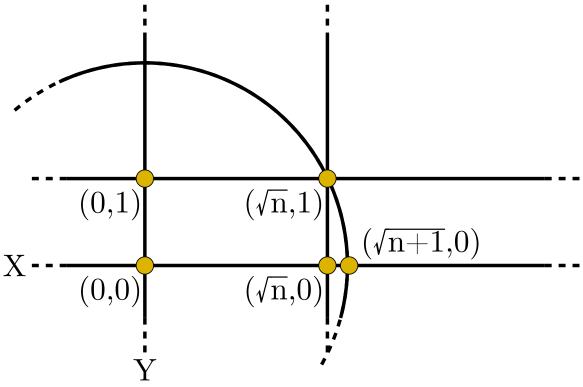
1. Using our earlier lemmas, find the points $(\sqrt{n}, 0)$ and $(0, 1)$.

2. Again using our earlier lemmas, create the point $(\sqrt{n}, 1)$.

3. Note that this point is distance
$$\sqrt{(\sqrt{n})^2 + (1)^2} = \sqrt{n+1}$$
from the origin.

4. Draw a circle centered at the origin through this point. This circle has radius $\sqrt{n+1}$, by construction.

5. Look at the intersection of this circle with the x-axis: this occurs at the point $(\sqrt{n+1}, 0)$ by construction!



□

So. . . is there anything **else** we can make?

At this point in time, we might want to consider changing tactics. We've shown that we can construct tons of numbers; however, our current proof techniques and tricks have all relied on coming up with clever tricks to divide, or scale, or make square roots. These techniques are great for making things — however, they are not particularly well-suited to determining what sorts of things we **cannot** make. This is a general theme in mathematics; often, the techniques used for establishing existence (i.e. working with $\exists$ quantifiers) are very different from those needed to establish nonexistence (i.e. working with $\forall$ quantifiers.)

This happens here as well. If we want to figure out some sorts of restrictions or bounds on what we can make, we cannot simply use clever tricks, or even talk about what sorts of things our specific clever tricks cannot do; we need to talk about **any** possible constructible point!

To do this, we reconsider what sorts of points are constructible. By our properties above, the only sorts of things we could consistently create were the following:

4. Given any two lines l_1, l_2 that intersect at some unique point, we can label that point of intersection.

5. Given any two circles C_1, C_2 that intersect at one or two points, we can label those points of intersection.

6. Given any circle C and line l that intersect at one or two points, we can label those points of intersection.

What kinds of points can these things create? We answer this in the following theorem:

Theorem. Let a be any constructible value. Then there are constructible values b, c such that a is a root of the polynomial

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

Proof. We simply go through cases. First, consider any line we can construct. Lines can only be made by taking two constructible points $(a, b), (c, d)$ and forming the line that goes through both of them: i.e.

$$(y - b)(c - a) = (x - a)(d - b),$$

or equivalently

$$y = \frac{d-b}{c-a}x + \left(\frac{b}{c-a} - \frac{a}{d-b}\right).$$

One useful thing to note here, however, is that because $\frac{d-b}{c-a}, \frac{b}{c-a} - \frac{a}{d-b}$ are both constructible by our earlier work, I don't actually need to care about these specific forms here! I can just instead write our line as

$$y = mx + b,$$

for two constructible numbers m, b. This simplifies our lives considerably.

Similarly, consider any circle we can construct. We can make any such circle by specifying a center (a, b) and a point (c, d) it passes through; consequently, circles all have the form

$$(x-a)^2+(y-b)^2=(c-a)^2+(d-b)^2.$$

Again, we can simply note that the RHS is constructible, and replace it with any arbitrary positive constructible number r:

$$(x-a)^2+(y-b)^2=r.$$

So: what do **intersections** of these things look like? Well; if we intersect any two nonparallel lines of the form

$$y=m_1x+b_1, y=m_2x+b_2,$$

we get that

$$x=\frac{b_2-b_1}{m_1-m_2}, y=m_1\frac{b_2-b_1}{m_1-m_2}+b_1.$$

In other words, our point of intersection is . . . constructible! So intersecting two lines doesn't give us anything new.

Let's try a line and a circle: that is, suppose we have the two equations

$$\begin{gathered} y=mx+c, \\ (x-a)^2+(y-b)^2=r. \end{gathered}$$

Plug in $mx+c$ for y in our second equation, to get

$$\begin{aligned} (x-a)^2+(mx+c-b)^2&=r \\ \Rightarrow x^2+\frac{(2m(c-b)-2a)}{(m^2+1)}x+\frac{(a^2+(c-b)^2-r)}{(m^2+1)}&=0. \end{aligned}$$

Again, notice that $\frac{(2m(c-b)-2a)}{(m^2+1)}, \frac{(a^2+(c-b)^2-r)}{(m^2+1)}$ are both constructible numbers; therefore, we can replace each of them with some appropriate placeholder, and say that all solutions for x are roots of a polynomial of the form

$$x^2+\alpha x+\beta,$$

for some constructible α, β. But this is exactly our claim!

Similarly, suppose we intersect two circles together: then we are trying to solve the two equations

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

This is not hard to do: simply take their difference, and you get the equation

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

which is just the equation of a line with constructible coefficients. We have already dealt with the line and circle case above, and so we're done!

□

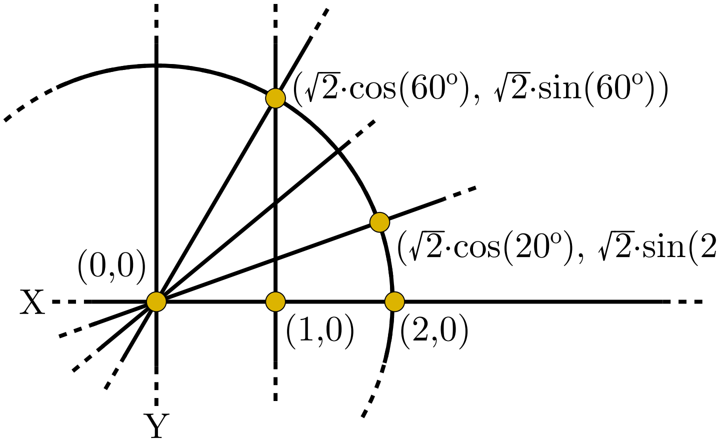
This is a remarkably strong result! In particular, I claim that with some machinery, it will let us answer the following three classical problems that were open for over two millenia, from the Egyptian/Greek eras until the mid-1800's:

- **Doubling the Cube**: Can you construct a number x such that the volume of the cube with side length x is 2?

  In other words; can you construct $\sqrt[3]{2}$?

- **Trisecting the Angle**: Given any two lines L, M that intersect at a unique point P in the plane, can you always draw a third line N through P such that the angle between N, L is a third of that between M, L?

  To give an explicit example: we can make a line that makes an angle of $\pi/3 = 60°$ with the origin by constructing a circle with radius 2 around the origin, drawing a line perpendicular to the x-axis through $(1, 0)$, finding their intersection P, and drawing the line through the origin and P.

  Can you draw a line that makes an angle of $\pi/9 = 20°$ with the origin? Equivalently: can you construct cos(20°), the point of intersection any such line would have with the circle of radius 1 through the origin?

- **Squaring the Circle**: Given a circle C with radius 1, can you construct a point P such that the distance from P to the origin is the same as the circumference of C?

  Basically: can you construct π?

We'll answer these questions next week!