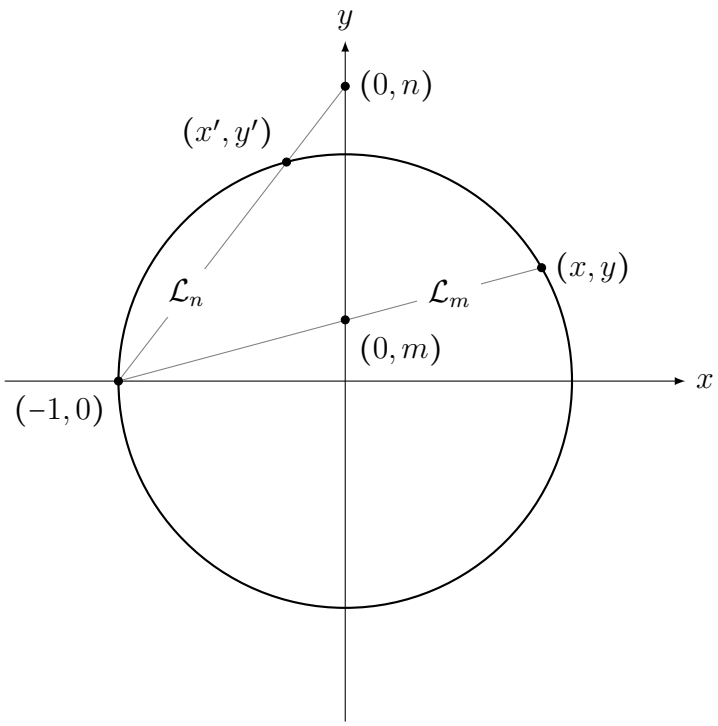
Last time we took up the problem of finding all pythagorean triples (which we abbreviate PTs). In fact, all we really need to find are all nontrivial *primitive* PTs, which we'll abbreviate as PPTs. We concluded last class with the observation that the problem of finding all PTs is essentially equivalent to the problem of finding all rational points on the unit circle. But how can we find these?

The main difficulty is the shape of the circle; it would be much easier to find all the rational points on a given line, for example. Can we change the problem from being about rational points on the circle to rational points on lines? The first approach is to simply cut the circle somewhere and unfold it, but the problem is that it's not clear what the relationship is between rational points on this flattened out circle and rational points on the original circle. Fortunately, there's a beautiful geometric way to turn a circle into a line that *preserves* rational points: projecting the circle onto the y-axis.

To see how this works, pick any real number m, and consider the line $\mathcal{L}_m$ of slope m that passes through the point $(-1, 0)$. $\mathcal{L}_m$ hits the circle at some point (x, y). Thus, to any given real number m we can associate a point (x, y) on the unit circle. We can also reverse this process: given a point (x, y) on the unit circle, we connect it to $(-1, 0)$ with a straight line, and let m be the slope of that line. Thus we've set up a correspondence between a real number m and a point (x, y) on the unit circle; we write $(x, y) \leftrightarrow m$. Here's a picture:



Visualizing the correspondence:
$(x, y) \longleftrightarrow m$ *and* $(x', y') \longleftrightarrow n$

This *almost* establishes a bijection between all points on the unit circle and $\mathbb{R}$: $f(x,y) := m$, the slope of the line $\mathcal{L}_m$ connecting (x,y) to $(-1,0)$. This is a surjection because for any $m \in \mathbb{R}$, the point $(0,m)$ on the y-axis can be connected via the straight line $\mathcal{L}_m$ to $(-1,0)$, and this line will hit the circle at some point (x,y). And it's an injection because if $f(x,y) = f(x',y')$, then there must be some line $\mathcal{L}$ that passes through $(-1,0)$, (x,y), and (x',y'). But a line can't pass through three different points on a circle! It follows that $(x,y) = (x',y')$, so f is injective.

So why did we say that f is *almost* a bijection? Because we didn't consider one important case: what is $f(-1,0)$? The only reasonable choice of line $\mathcal{L}_m$ is the vertical line tangent to the unit circle at $(-1,0)$, which makes m is undefined. However, this is the *only* issue with our proof that f is a bijection, and it seems a pity to lose such a nice result over a single technicality. Thus, we write $f(-1,0) := \infty$, the "point at infinity".

To write this bijection down more explicitly, we quickly compute the slope m of the line connecting $(-1,0)$ and (x,y):

$$m = \frac{y}{x+1}. \tag{1}$$

We can now write down our promised bijection between the set of points on the unit circle $S := \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ and the y-axis union the point at infinity:

$$\begin{aligned} f : S &\longrightarrow \mathbb{R} \cup \{\infty\} \\ (x,y) &\longmapsto \begin{cases} \frac{y}{x+1} & \text{if } (x,y) \neq (-1,0) \\ \infty & \text{otherwise.} \end{cases} \end{aligned}$$

Since f is a bijection, f^{-1} is, too. What is $f^{-1}(m)$? In other words, given the slope of the line $\mathcal{L}_m$, what is the point $(x,y) \in S$ at the intersection of $\mathcal{L}_m$ with the unit circle? To find the intersection of the line $\mathcal{L}_m$ with the circle S we have to solve a system of equations:

$$\begin{aligned} y &= m(x+1) \\ x^2 + y^2 &= 1. \end{aligned}$$

Substituting the first into the second and solving for x yields

$$x = -1 \quad \text{or} \quad \frac{1-m^2}{1+m^2}.$$

The first solution yields the point $(-1,0)$, which we already knew was on the line. The other solution produces

$$y = \frac{2m}{1+m^2}$$

whence

$$f^{-1}(m) = \left(\frac{1-m^2}{1+m^2}, \frac{2m}{1+m^2}\right). \tag{2}$$

Of course this only applies when $m \in \mathbb{R}$; for the point at infinity we simply define

$$f^{-1}(\infty) := (-1, 0).$$

Remark. This is consistent with our formula (2): $\lim_{m\to\infty} f^{-1}(m) = (-1, 0)$.

OK, so now we have a bijection f between the points on the circle S and the y-axis (including the point at infinity). But the purpose of all this was to find rational points on S! The reason we care about this particular bijection is that it preserves rationality:

Proposition: $f(x, y) \in \mathbb{Q} \cup \{\infty\}$ if and only if $(x, y) \in \mathbb{Q}^2$.
Proof: This follows immediately from (1) and (2). □

We've therefore come up with a nice description (the fancy word is *parametrization*) of all the rational points on the unit circle:

$$\left\{\left(\frac{1-m^2}{1+m^2}, \frac{2m}{1+m^2}\right) : m \in \mathbb{Q} \cup \{\infty\}\right\},$$

is a complete list of all the rational points on the circle.

We now return to our original problem of determining all pythagorean triples. Given a nontrivial pythagorean triple (A, B, C), there's a corresponding rational point on the unit circle: $(\frac{A}{C}, \frac{B}{C})$. Since any nontrivial rational point on the unit circle takes the form $(\frac{1-m^2}{1+m^2}, \frac{2m}{1+m^2})$, we deduce that

$$\frac{A}{C} = \frac{1-m^2}{1+m^2} \qquad \text{and} \qquad \frac{B}{C} = \frac{2m}{1+m^2}. \tag{3}$$

Write $m = \frac{a}{b}$. Substituting this in and simplifying yields

$$\frac{A}{C} = \frac{b^2-a^2}{b^2+a^2} \qquad \text{and} \qquad \frac{B}{C} = \frac{2ab}{b^2+a^2}. \tag{4}$$

At this point it's tempting to conclude that our pythagorean triple must be

$$(A, B, C) = (b^2 - a^2, 2ab, b^2 + a^2),$$

but this doesn't follow from the above. The issue is that the fractions above might not be reduced; for example, taking $a = 1$ and $b = 3$ yields $\frac{A}{C} = \frac{8}{10}$, which of course doesn't force $A = 8$ and $C = 10$.

On Problem Set 9 you'll show that not every pythagorean triple can be expressed in the form $(b^2 - a^2, 2ab, b^2 + a^2)$. However, it turns out that every *primitive* triple can be expressed this way:

Theorem. The set $\{(b^2 - a^2, 2ab, b^2 + a^2) : a, b \in \mathbb{Z}\}$ contains all primitive pythagorean triples (up to exchanging the first two entries).

The parenthetical remark simply means that the set might not contain the primitive pythagorean triple $(4, 3, 5)$, but it will contain the equivalent triple $(3, 4, 5)$.

Proof. Given (A, B, C) a nontrivial primitive pythagorean triple. Thus in particular we have

$$(A, B) = 1. \tag{5}$$

But we can say more: that one of A, B is odd and the other is even, often expressed A and B have *opposite parity*. In symbols, this reads

$$A \not\equiv B \pmod 2. \tag{6}$$

Without loss of generality, we may assume that A is odd and B is even. Now pick some $m \in \mathbb{Q}$ and write $m = \frac{a}{b}$ in reduced form. Plugging this into (3) yields (4) as before.

I claim (5) and (6) hold if we replace A and B by a and b:

$$(a, b) = 1 \qquad \text{and} \qquad a \not\equiv b \pmod 2. \tag{7}$$

The first half of this is automatic, since we picked $\frac{a}{b}$ to be reduced. Of course this also means a and b can't both be even. But they also can't both be odd, since in this case we'd have

$$\frac{B}{C} = \frac{ab}{(b^2 + a^2)/2},$$

contradicting our hypothesis that B is even.

As you will show on Problem Set 9, (7) implies that $\frac{b^2-a^2}{b^2+a^2}$ and $\frac{2ab}{b^2+a^2}$ are both reduced. Putting this together with (5) and (6), we deduce that every fraction appearing in (4) is reduced. It immediately follows that

$$A = b^2 - a^2, \qquad B = 2ab, \qquad C = b^2 + a^2.$$

This concludes the proof of the theorem. □

Inspecting the proof, we see that we've actually proved something stronger than what we originally claimed:

Theorem. Let $\mathcal{T}$ denote the collection of all nontrivial primitive pythagorean triples (A, B, C) with B even. Then

$$\mathcal{T} = \{(b^2 - a^2, 2ab, b^2 + a^2) : a, b \in \mathbb{Z}, (a, b) = 1, a \not\equiv b \pmod 2\}.$$

Thus generating nontrivial primitive pythagorean triples is easy: each pair (a, b) of positive coprime integers of opposite parity produces one, and every nontrivial primitive pythagorean triple can be formed in this way. For example, $(1, 2)$ generates $(3, 4, 5)$, $(2, 3)$ generates $(5, 12, 13)$, and $(1, 4)$ generates $(15, 8, 17)$.

Fermat's Last Theorem.

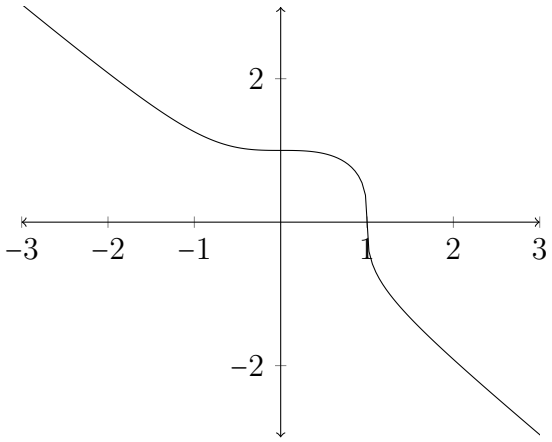
Having found a complete description of the integer solutions to $x^2 + y^2 = z^2$, the next natural question is to find all solutions to

$$x^3 + y^3 = z^3.$$

Some solutions are easy to find: $(0,0,0)$, $(1,-1,0)$, $(1,0,1)$, or in general $(x,0,x)$, $(x,-x,0)$. Notice that these are all trivial.

Fermat's Last Theorem: There are no nontrivial integer solutions to $x^3 + y^3 = z^3$. In fact, for any integer $k \geq 3$, there are no nontrivial integer solutions to $x^k + y^k = z^k$.

After discussing the history of this result, we looked at the cubic case a bit more carefully. Just as in the case of pythagorean triples, we can transform the problem into that of finding all rational points on the curve $x^3 + y^3 = 1$.



Graph of the curve $x^3 + y^3 = 1$

Recall from above that we managed to find a nice bijection of the circle onto the y-axis that preserves rationality. Could we find a similar bijection for this curve? Playing around with various changes of variables that preserve rationality, one might happen to try

$$r := \frac{12}{x+y} \qquad s := \frac{36(x-y)}{x+y};$$

in terms of these new variables, $x^3 + y^3 = 1$ becomes $r^2 = s^3 - 432$. Equations of this type have been well-studied:

Definition: An *elliptic curve* is an equation of the form $y^2 = x^3 + ax + b$.

The rest of our semester will be devoted to discussing rational points on elliptic curves, and the fascinating discoveries that eventually led Wiles to prove Fermat's Last Theorem.