CCS Discrete Math I | Professor: Padraic Bartlett

Lecture 9: Elliptic Curves

Week 9 | *UCSB 2014*

> It is possible to write endlessly on elliptic curves. (This is not a threat.)
>
> Serge Lang, *Elliptic curves: Diophantine analysis.*

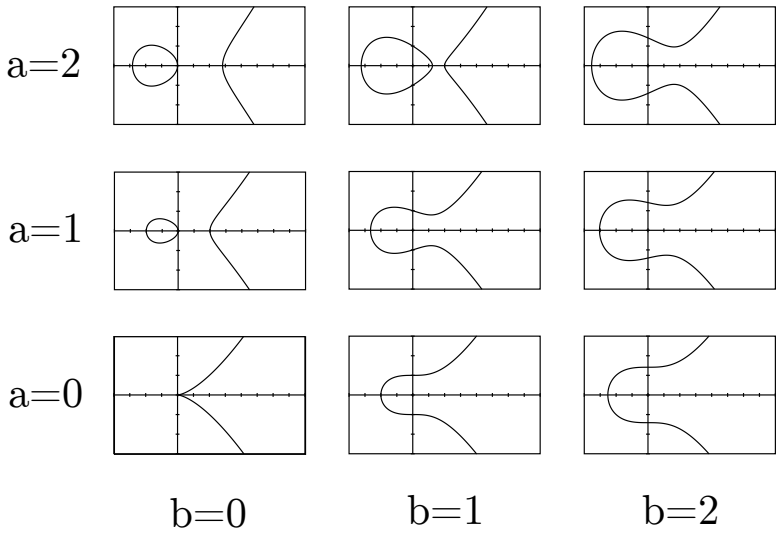
1 Elliptic Curves

1.1 Basic definitions and observations.

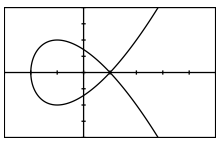
Definition. An **elliptic curve** over $\mathbb{R}$ with coefficients $a, b \neq 0 \in \mathbb{R}$ is the collection of all points $(x, y) \in \mathbb{R}^2$ satisfying the equation

$$y^2 = x^3 - ax + b.$$

We sketch some sample elliptic curves below:



We want to restrict our notion of elliptic curves to those with coefficients a, b such that $27b^2 - 4a^3 \neq 0$. This condition will insure that our curves have a well-defined notion of "tangent" at every point on the curve, unlike the $a = 0, b = 0$ curve above (which has no tangent at $(x, y) = (0, 0)$,) or the $a = 3, b = 2$ curve below (which has no well-defined tangent at $(1, 0)$.)



We justify this claim here:

Proposition. Suppose that E is a curve of the form

$$y^2 = x^3 - ax + b,$$

with $27b^2 - 4a^3 \neq 0$. Then E has a well-defined and unique tangent slope at every point.

Proof. To do this, we should first try to determine what the tangent slopes to this curve even look like. So: how do we determine the tangent line to an expression in two variables? You could try writing E as the graph of the two curves $y = \pm\sqrt{x^3 - ax + b}$; in this case then we can express the slope of the tangent to E at any point other than when $y = 0$ as

$$\begin{aligned}\pm\frac{d}{dx}\left(\sqrt{x^3 - ax + b}\right) &= \pm\frac{1}{2\sqrt{x^3 - ax + b}} \cdot \frac{d}{dx}\left(x^3 - ax + b\right) \\ &= \pm\frac{3x^2 - a}{2\sqrt{x^3 - ax + b}}\end{aligned}$$

with the choice of sign above corresponding to whether we are looking for points with $y > 0$ or $y < 0$.

The above definition gives us a defined tangent slope whenever $x^3 - ax + b > 0$. If $x^3 - ax + b < 0$, there is no value of y such that $y^2 = x^3 - ax + b$, so no such values of x lie on our curve; this leaves us only with the case where $y^2 = x^3 - ax + b = 0$.

However, this case is not one that the above methods can help us with. To do things in this situation, we can use the technique of **implicit differentiation**: that is, simply apply the operator $\frac{d}{dx}$ to the equation $y^2 = x^3 - ax + b$ to get

$$\frac{d}{dx}(y^2) = 2y \cdot \frac{dy}{dx} = 3x^2 - a \Rightarrow \frac{dy}{dx} = \frac{3x^2 - a}{2y}.$$

The LHS, the change in y as we vary x slightly, is what we want! The RHS is an expression in terms of this object; accordingly, we now have an expression for the derivative!

You might object here that this formula runs into the same pitfalls that the old formula encountered: that is, it is undefined when $y = 0$! I claim that with some thought, however, we can still interpret this in many more situations than those earlier. Specifically:

- Notice that if the top and bottom are zero simultaneously, then we have $3x^2 - a = 0$ and $y = 0$. This implies that $x = \pm\sqrt{a/3}$, and thus that because $0 = y = x^3 - ax + b$, we have for $x = \sqrt{a/3}$

$$\begin{aligned} 0 &= \left(\sqrt{a/3}\right)^3 - a(\sqrt{a/3}) + b \\ \Rightarrow 0 &= \frac{a^{3/2}}{3\sqrt{3}} - \frac{3a^{3/2}}{3\sqrt{3}} + b \\ \Rightarrow \frac{2a^{3/2}}{3\sqrt{3}} &= b \\ \Rightarrow \frac{4a^3}{27} &= b^2 \\ \Rightarrow 0 &= 27b^2 - 4a^3. \end{aligned}$$

But we said this is impossible in our definition of elliptic curves from before! So this case does not happen.

- In the other case, suppose that the top is nonzero while the bottom is zero. How can we interpret this situation? Well: let's flip x and y in our graph! This flips our curve over the line $x = y$, and transforms our derivative here into one where the top is zero but the bottom is nonzero — in other words, the tangent here is a horizontal line!

  Therefore, if we return to our "normal" world where x, y are not flipped, we actually have a **vertical** tangent line, as horizontal lines become vertical lines when flipped over $y = x$ (and remain tangent lines as well, as their slopes still agree with the function!) So we have a well-defined tangent at this point.

This covers all of our cases; therefore elliptic curves have a well-defined tangents at all of their points.

□

This proposition is particularly useful in proving the following two results, which characterize some of the most useful properties we will refer to in these talks:

Proposition. Take any two distinct points P, Q on an elliptic curve E given by the equation $y^2 = x^3 - ax + b$. Draw a straight line L through these two points. There are two possibilities: either L intersects no other points of our elliptic curve, or L intersects our elliptic curve at exactly one other point R.

Proof. This is not too difficult to see. Simply take our two points P, Q, and let L be the line through P, Q. There are two possibilities:

- The line L through P, Q has a well-defined slope. In this case, we can find constants m, c such that L is graphed by the linear equation $y = mx + c$. The x-coördinates of intersections of L with E occur precisely when

  $$(mx + c)^2 = x^3 - ax + b;$$

  in other words, at roots of the polynomial

  $$p(x) = x^3 - m^2x^2 - (2mc + a)x + (b - c^2).$$

  This polynomial is degree 3, and therefore has at most three distinct roots. If $P \neq Q$ and these are both points on the line L given by $y = mx + c$, then these two points occur at different x-coördinates; so P, Q account for two of the possible roots. Call these two roots r_1, r_2, and factor them out of $p(x)$; this leaves

  $$p(x) = (x - r_1)(x - r_2)q(x),$$

  for some linear polynomial $q(x)$. But all degree-1 polynomials have exactly one root; therefore there is exactly one more x-coördinate r_3 such that $p(r_3) = 0$, and therefore (because any x-coördinate uniquely identifies a single point on L) exactly one other point R at which E, L intersect. If P or Q are equal to R, then this point is one that we've already intersected; otherwise it is a new point of intersection. In either case, there was at most one new point of intersection, which proves our claim!

- The line L through P, Q does not have a well-defined slope, which can only happen when P, Q share the same x-coördinate and have different y-coördinates. In this situation, if α denotes this shared x-coördinate, we have that the line through X, Y is the line L given by the graph of $x = \alpha$, and thus that any intersections of L, E happen when

$$y^2 = \alpha^3 - a\alpha + b.$$

  There are at most two roots to this equation above, and we've already found both of them in P, Q! So there cannot be any third point that is not equal to either P, Q.

□

If we look at what we've proven here, we can actually strengthen our result as follows:

Proposition. Suppose that P, Q are two points on an elliptic curve E given by $y^2 = x^3 - ax + b$, such that the line L through P, Q is not tangent to E at either P or Q. Also assume that P, Q have distinct x-coördinates. Then the line L intersects E at a third point $R \neq P, Q$.

Proof. Simply revisit our proof above. Because P, Q have distinct x-coördinates, we know that we are not in the second case from our earlier proof. So we can assume that we are in the first case; as before, let L be graphed by $y = mx + c$, and again find the three roots r_1, r_2, r_3 given by intersecting our elliptic curve E with L.

We know by assumption that $r_1 \neq r_2$. If $r_3 \neq r_1, r_2$, our proof is done. So we simply need to show that it is impossible for r_3 to equal r_1, r_2 given the assumptions of our problem.

Why is this impossible? Well: let's try a proof by contradiction. Consider what would happen if we had a repeated root; let's say r_1 is repeated, without losing any generality. This would imply that the polynomial

$$p(x) = x^3 - ax + b - (mx + c)^2$$

can be factored as

$$p(x) = (x - r_1)^2(x - r_2),$$

because it factors into three roots, and we know that $r_1 \neq r_2$ are both roots because P, Q are places of intersection of L with E.

However: this implies that the derivative of $p(x)$,

$$\frac{d}{dx}p(x) = 2(x - r_1)(x - r_2) + (x - r_1)^2$$

also has a root at r_1, as plugging this in above gives us a zero. If we use the fact that $p(x) = (x^3 - ax + b) - (mx + c)^2$, this gives us

$$3r_1^2 - a = \frac{d}{dx}\left(x^3 - ax + b\right)\Big|_{x=r_1} = \frac{d}{dx}\left((mx + c)^2\right)\Big|_{x=r_1} = 2m(mr_1 + c).$$

Also, notice that at $x = r_1$, y $= mr_1 + c$ by definition; so we actually have that $3x^2 - a = 2my$. What does this tell us? Well: recall from earlier that our elliptic curve has a well-defined notion of derivative, given by

$$\frac{d}{dx}(y^2) = 2y \cdot \frac{dy}{dx} = 3x^2 - a \Rightarrow \frac{dy}{dx} = \frac{3x^2 - a}{2y}.$$

Consequently, if we plug in our observations from above here, we get

$$\begin{aligned} \frac{dy}{dx}\Big|_{x=r_1} &= \frac{3x^2 - a}{2y}\Big|_{x=r_1} \\ &= \frac{2my}{2y} \\ &= m, \end{aligned}$$

provided that we weren't looking at a value of r_1 at which $y = 0$. But in this case we have that the line L is tangent to our curve, which we said was impossible in our proposition's claim! Therefore this situation is impossible.

This leaves us with only the situation where we have a repeated root r_1 for $p(x)$, at which $y = 0$; we want to show that is impossible. This is not hard: because $y = 0$, we know that at $x = r_1$, we have

$$x^3 - ax + b = 0$$

by definiton, and also

$$0 = 2my = 2m(mx + c) = 3x^2 - a$$

by our earlier calculations. But in our proof where we showed that elliptic curves have well-defined tangent lines, we showed that this combination of conditions forced $27b^2 - 4a^3 = 0$! So we know that this, too, is impossible. Consequently we have proven that the only possible situation is when we have three distinct roots, as claimed. □

If you let $P = Q$, and interpret the "straight line" through those two points to be the tangent line to our curve at that point, you can interpret its claims as follows:

Proposition. Suppose that P is a point with nonzero y-coördinate on an elliptic curve E given by $y^2 = x^3 - ax + b$. Take the tangent line L to E at P. There are two possibilities:

1. L intersects E at exactly one other point on the curve. If we graph L by $y = mx + b$, which we can do because at points with nonzero y-coördinate we have shown that the slopes of tangent lines exist and are finite, we have that $p(x) = (x^3 - ax + b) - (mx + b)^2$ can be factored into something of the form $(x - r_1)^2(x - r_2)$, where r_1 is the $x-$coördinate of P, and r_2 is the x-coördinate of the unique other point on the curve we cross.

2. L never intersects E at any other points on our curve. If we graph L by $y = mx + b$, we have that $p(x) = (x^3 - ax + b) - (mx + b)^2$ can be factored into something of the form $(x - r_1)^3$, where r_1 is the $x-$coördinate of P.

Proof. On the HW! Essentially, it works just like the earlier proofs we've done. □

Finally, we can extend this to the last remaining case, where P has zero y-coördinate:

Proposition. Suppose that P is a point of the form $(c, 0)$ on an elliptic curve E given by $y^2 = x^3 - ax + b$. Then the tangent line L to P is of the form $x = c$, and this line has no other intersections with our elliptic curve E.

Proof. We proved the first part of this claim (that the tangent line has the form $x = c$) earlier, when we showed that any elliptic curve has a well-defined tangent everywhere; so it suffices to prove the second part. This is immediate; we know that at $(c, 0)$ we have

$$0^2 = y^2 = c^3 - ac + b.$$

Any other intersection of the line would have to have the form (c, d), where $d^2 = c^3 - ac + b$; but this forces $d = 0$, because $c^3 - ac + b = 0$. So no other intersection exists! □

1.2 Creating the group structure.

The main reason we care about these curve properties is because they let us do something very strange to our elliptic curves: we can make them into a group! Consider the following operation on elliptic curves:

Definition. Take some fixed elliptic curve E $y^2 = x^3 + ax + b$. Add to the collection of the points in E a "point at infinity," $\mathcal{O}$. Given this collection of points, consider the following binary operation:

1. Take any two points P, Q.
2. Draw a line L as follows:
   - If P, Q are distinct points, neither of which are the point at infinity, draw the line L through them.
   - If $P = Q$ and neither are the point at infinity, draw the tangent line L to P.
   - If one of P, Q is the point at infinity $\mathcal{O}$, and the other is not, then let (c, d) be the coördinates of the non-$\mathcal{O}$ point, and draw the line L given by $x = c$.
   - Finally, if both $P, Q = \mathcal{O}$, draw the "line at infinity" L that consists only of the point $\mathcal{O}$.
3. Using this line, define a third new point R as follows:
   - If L intersects our curve at a point that is neither P nor Q, it does so at exactly one such new point, by our earlier work. Call this point R.
   - Otherwise, we must be in one of the cases where we did not get a new point.
   - Suppose that we are in the case where $P \neq Q$ and $P = (c, d)$, $Q = (c, f)$; i.e. the line L is of the form $x = c$. Then we know that this line has no other intersections with our curve! Take this to mean that the line L's third point of intersection happens "at infinity," and set $R = O$.

- Otherwise, if we are still in the case where $P \neq Q$, $P, Q \neq \mathcal{O}$, then this can only happen when L is tangent to one of P, Q. Think of this tangency as representing the idea of "repeating" one of P, Q, and in fact of that "repeated" point as occurring twice on our line! Set R equal to this "repeated" point.
- Similarly, if $P = Q$ and we didn't get a third point, then our tangency at P was in some sense of "order 3", in that we could factor out P's x-coordinate as a root three times. Think of this as meaning that P was repeated 3 times on our line, and set $R = P$.
- If one of P, Q was the point at infinity, the other had coördinates (c, d), and we didn't get a second point of crossing for the line $x = c$, then this means that the d-coordinate of this point is 0, because otherwise $(c, -d)$ would be another distinct point on both our curve E and line L. This means that the line L is tangent to our non-O point, as proven before; so that non-O point is again repeated twice in a sense! Set R equal to that repeated point.
- Finally, we have $P = Q = \mathcal{O}$. In this sense our line only contains the point at infinity; so let's set $R = O$ here.

4. If $R = \mathcal{O}$, set $-R = \mathcal{O}$. Otherwise, if $R = (e, f)$, set $-R = (e, -f)$, which we know is on our curve.

5. Finally, define our binary operation $+$ as follows: set $P + Q = -R$. By all of our earlier propositions, we know that this binary operation is well-defined.

Pictures of points being added on an elliptic curve. Again, pictures stolen from Wikipedia.

I claim to you that this defines a group! We prove this here:

Theorem. Take some fixed elliptic curve E $y^2 = x^3 + ax + b$, and add to E a "point at infinity," $\mathcal{O}$. This curve, along with the binary operation defined above, forms an abelian (that is, commutative) group.

Proof. To prove this, we simply check the properties of a commutative group.

Inverses: We claim that for any point P in our curve, there is some $-P$ such that $P + (-P) = \mathcal{O}$ (and will show later that $\mathcal{O}$ is indeed the identity.) This falls out by definition: we said that the "third point" through $P, -P$ was $\mathcal{O}$ and that $-\mathcal{O} = \mathcal{O}$, therefore we have $P + -P = \mathcal{O}$.

Identity: We claim that O is the identity of this group. This is not hard to see: take any point P, and consider the sum $P + \mathcal{O}$. To calculate this sum, we draw a vertical line L through P. If $P = (x, y)$ with $y \neq 0$, then this line's third point of intersection is at $(x, -y) = -P$; therefore we have $P + \mathcal{O} = -(-P) = P$, as claimed.

Commutativity: We claim that for any two points P, Q in our curve, we have $P + Q = Q + P$. This is immediate; our operation above was centered around drawing the line through P, Q to find $P + Q$. This line-drawing operation creates the same line no matter in what order we specify our points; therefore we have commutativity.

Associativity: We claim that for any three points P, Q, R in our curve, we have $(P + Q) + R = P + (Q + R)$. This, surprisingly, is a **nightmare** to prove. Well, not a nightmare;

it's incredibly gorgeous! It however would take us at least 2-3 more weeks to prove, which is more time than we have. Talk to me if you'd like a proof sketch! Otherwise, take this on faith for now. □

Instead, we choose to focus here on an example of this operation:

Example. Consider the elliptic curve $\mathrm{E} = y^2 = x^3 - 43x + 166$. Suppose you look at the point $(3, 8)$, which is on this curve. What points can you create by repeatedly adding this point to itself?

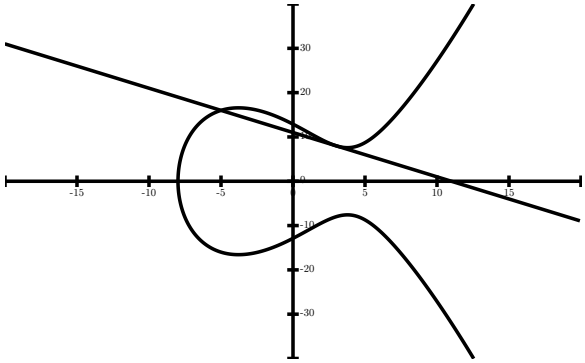
Proof. We start by first verifying that this point is even on our curve:

$$8^2 = 64, 3^3 - 43 \cdot 3 + 166 = 64.$$

With this done, we add our point to itself! We do this by finding the tangent to our curve at $(3, 8)$. The slope of the tangent, as discussed before, is $\frac{dy}{dx} = \frac{3x^2-a}{2y} = \frac{3x^2-43}{2y}$, which at $(3, 8)$ is $\frac{-16}{16} = -1$. So our tangent line is $y = 11 - x$, and therefore our points of intersection are all of the values of (x, y) such that

$$(11 - x)^2 = x^3 - 43x + 166 \Rightarrow \qquad 0 = x^3 - x^2 - 21x + 45 = (x - 3)^2(x + 5).$$

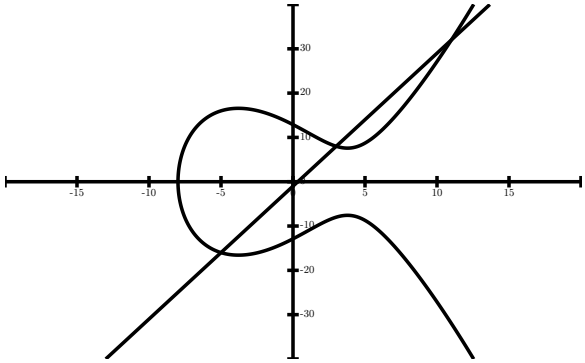
So the third point of intersection is at $x = -5$, at which value $y = 11 - (-5) = 16$. So the third point on our curve is $(-5, 16)$, and therefore we have $(3, 8) + (3, 8) = (-5, -16)$.



We now add $(3, 8)$ to this point again. To do this, we construct the line through both $(3, 8)$ and $(-5, -16)$, which is just $y = 3x - 1$. Solving again yields

$$(3x - 1)^2 = x^3 - 43x + 166 \Rightarrow \quad 0 = x^3 - 9x^2 - 37x + 165 = (x - 3)(x + 5)(x - 11).$$

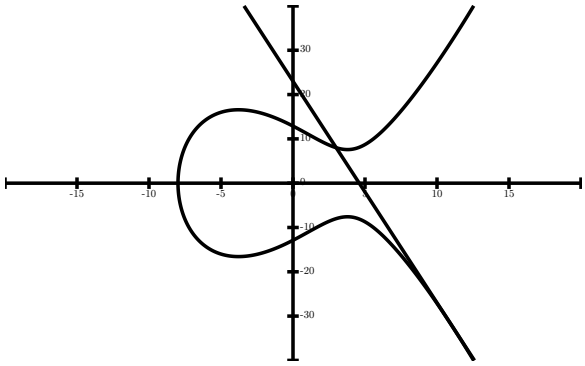
Our third point of intersection therefore occurs at $x = 11$, $y = 3 \cdot 11 - 1 = 32$; consequently, the sum $(3, 8) + (-5, -16) = (11, -32)$.



We add $(3, 8)$ to this point again. To do this, we construct the line through both $(3, 8)$ and $(11, -32)$, which is just $y = 23 - 5x$. Solving again yields

$$(23 - 5x)^2 = x^3 - 43x + 166 \Rightarrow \qquad 0 = x^3 - 25x^2 + 187x - 363 = (x - 11)^2(x - 3).$$

So our third point of intersection occurs at $x = 11$ again! In other words; this line L is tangent to our curve at $(11, -32)$, so we have that the third point is $(11, -32)$, and therefore that $(3, 8) + (11, -32) = (11, 32)$.



We do this again! The line through $(3, 8)$ and $(11, 32)$ is $y = 3x - 1$ again, because we've already dealt with this case in our earlier work! This line went through $(-5, -16)$ as its third point, so we have that $(3, 8) + (11, 32) = (-5, 16)$.

As well, we already know that the line through $(3, 8)$ and $(-5, 16)$ is tangent to $(3, 8)$; therefore the third point on this curve is $(3, 8)$ again, and we have $(3, 8)+(-5, 16) = (3, -8)$.

Finally, we now want to add $(3, 8)$ to $(3, -8)$; this is done by drawing the vertical line connecting these two points, which goes through $\mathcal{O}$ as its third point. So we have $(3, 8) + (3, -8) = \mathcal{O}$.

We could go further, but we'd loop back, because $(3, 8) + \prime = (3, 8)$, and we'd be back to where we started! So we've actually found **all** of the multiples of $(3, 8)$: that is, we've

shown that

$$\begin{aligned}
1 \cdot (3,8) &= (3,8) \\
2 \cdot (3,8) &= (-5,-16) \\
3 \cdot (3,8) &= (11,-32) \\
4 \cdot (3,8) &= (11,32)) \\
5 \cdot (3,8) &= (-5,16) \\
6 \cdot (3,8) &= (3,-8) \\
7 \cdot (3,8) &= \mathcal{O}.
\end{aligned}$$

(When we write $n \cdot (a,b)$ here, we mean $\overbrace{(a,b) + \ldots + (a,b)}^{n \text{ times.}}$.) □

1.3 Elliptic Curves over Finite Fields

So: the reason that we're interested in these objects is not because of their structure over $\mathbb{R}^2$, as gorgeous as it is. Instead, I want to look at these objects as curves in **finite fields**! Specifically, we make the following definitions:

Definition. An **elliptic curve** over $\mathbb{F}_q$, for any finite field $\mathbb{F}_q$ for $q \neq 2, 3$ consists of all of the points $(x, y) \in \mathbb{F}_q^2$ satisfying the equation

$$y^2 = x^3 - ax + b.$$

(For $q = 2, 3$ the story is a bit more complicated. We'll explore this next week!)

In the past, we showed that any line L that intersects an elliptic curve E at one place intersects E in exactly three places (given appropriate notions of "three.") This still holds here:

Theorem. If E is an elliptic curve over some finite field $F_q, q \neq 2, 3$, and L is any line in $\mathbb{F}_q^2$, then L intersects E in three places (provided we count tangencies as multiple intersections, and add a point at infinity that is in every vertical line.)

We leave the proof of this claim for the homework! It goes through in exactly the same way as our earlier proof did.

Instead, we calculate an example to illustrate how elliptic curves work over a finite field:

Example. Let E denote the collection of all points in $\mathbb{F}_5^2 = (\mathbb{Z}/5\mathbb{Z})^2$ such that $y^2 = x^3 + 2x + 1$. Draw E, and create the group table corresponding to the group given by E and our elliptic curve group law.

Proof. We start by finding all of the points (x, y) that are in our curve. We do this by first noticing that in $\mathbb{Z}/5\mathbb{Z}$, we have

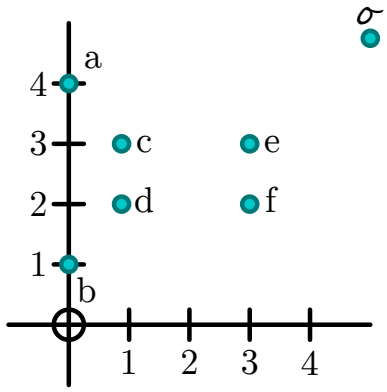
$$0^2 = 0, 1^2 = 1, 2^2 = 4, 3^3 = 4, 4^2 = 1$$

and therefore that the only numbers that have "square roots" are $0, 1$, and 4.

So: from here, to calculate points on our curve, it suffices to look at all possible values of x. Notice that

- $x = 0$ forces $x^3 + 2x + 1 = 1 = y^2$. Consequently, we have $y = 1$ or 4; so the two points $(0, 1), (0, 4)$ are in our curve.
- $x = 1$ forces $x^3 + 2x + 1 = 4 = y^2$. Consequently, we have $y = 2$ or 3; so the two points $(1, 2), (1, 3)$ are in our curve.
- $x = 2$ forces $x^3 + 2x + 1 = 13 \equiv 3 \mod 5$. There are no values of y that square to 3, so there are no points with x-coördinate 2 on our curve.
- $x = 3$ forces $x^3 + 2x + 1 = 34 \equiv 4 \mod 5$. Consequently, we have $y = 2$ or 3; so the two points $(3, 2), (3, 3)$ are in our curve.
- $x = 4$ forces $x^3 + 2x + 1 = 73 \equiv 3 \mod 5$. There are no values of y that square to 3, so there are no points with x-coördinate 2 on our curve.

We graph our points here, and give them the labels $a, b, c, d, e, f, \mathcal{O}$:



We start to make a group table for our curve here. We use the labelings above:

$+$	$\mathcal{O}$	a	b	c	d	e	f
$\mathcal{O}$							
a							
b							
c							
d							
e							
f							

We can use our knowledge of how groups work to get several values in our table "for free." $\mathcal{O}$, for example, we know to be an identity; so adding it to any point in our group doesn't change that element! Similarly, we know that any vertical line through our curve contains two non-$\mathcal{O}$ points and one $\mathcal{O}$ point. If we call the two points on that line α, β, then we have that $\alpha + \beta = \mathcal{O}$; in other words, $\alpha = -\beta$! This is really useful; it tells us that $b = -a$,

$d = -c$, and $f = -e$.

$+$	$\mathcal{O}$	a	$-a$	c	$-c$	e	$-e$
$\mathcal{O}$	$\mathcal{O}$	a	$-a$	c	$-c$	e	$-e$
a	a		$\mathcal{O}$				
$-a$	$-a$	$\mathcal{O}$					
c	c				$\mathcal{O}$		
$-c$	$-c$			$\mathcal{O}$			
e	e						$\mathcal{O}$
$-e$	$-e$					$\mathcal{O}$	

Let's fill in a few entries in this table!

1. To start, let's look at $a + a$. To do this, we want to find the tangent line to our curve at $a = (0, 4)$. The slope of the tangent line to our curve, as discussed before, is $\frac{\frac{d}{dx}(x^3+2x+1)}{2y} = \frac{3x^2+2}{2y}$ whenever the denominator is nonzero, and is infinity (i.e. a vertical line) otherwise. At $(0, 4)$ this is just $\frac{2}{8} = \frac{2}{3} = 2 \cdot \frac{1}{3} = 2 \cdot 2 = 4$. (If you're bothered by these calculations, remember that we're working in $\mathbb{Z}/5\mathbb{Z}$; therefore $1/3$, the inverse of 3, is just whatever number we multiply 3 by to get to 1. In particular, this is 2, as $3 \cdot 2 = 6 \equiv 1 \mod 5$.)

   So, we have a line of slope 4 through $(0, 4)$; this means our line is just $y = 4x + 4$. We want all intersections of $y = 4x + 4$ with our curve $y^2 = x^3 + 2x + 1$; to solve this, we just plug in our line into our curve to get

$$\begin{aligned}(4x+4)^2 &= x^3 + 2x + 1 \\ \Rightarrow 0 &= x^3 + 2x + 1 - 16x^2 - 32x - 16 \\ &\equiv x^3 - x^2 \mod 5 \\ &= x^2(x-1).\end{aligned}$$

   This has three roots; two of them are $x = 0$, which we expected because we took the tangent at $x = 0$, and the third is at $x = 1$. At $x = 1$ our curve is $y = 4 \cdot 1 + 4 = 3 \mod 5$, and thus the third point on our curve is $(1, 3) = c$.

   So we have $a + a = -c$!

2. We can also use the fact that a, a, c are all on the same line to get for "free" that $a + c = -a$, as we've already done by the above all of the hard work! As well, commutativity tells us that $c + a = a + c = -a$. That's nice.

3. We can use a similar process to calculate $c + c$, for $c = (1, 3)$; the slope here is $\frac{3x^2+2}{2y} = 0$, and therefore our line is the horizontal line $y = 3$. We then want to find

all points on this line that satisfy

$$\begin{aligned}(3)^2 &= x^3 + 2x + 1\\ \Rightarrow 0 &= x^3 + 2x + 1 - 9\\ &\equiv x^3 - 2x - 3 \mod 5\\ &\equiv x^3 - -5x^2 + 7x - 3 \mod 5\\ &= (x-1)^2(x-3).\end{aligned}$$

One trick I used to factor the above: we already know that $x-1$ is a doubly repeated root of the RHS! So we're just looking for what the third root should be; we can find this by just looking at the product of the constant terms, and then verify it by multiplying terms out. (So, in fact, writing the RHS as $x^3 - 5x^2 + 7x - 3$ is something I did after the fact; my first step was to guess that the third root is $x = 3$ because $-1 \cdot -1 \cdot -3 = -3$ gave me the desired constant term.)

This tells us that our third root occurs at $x = 3$, and in particular is the point $(3, 3) = e$. So we have $c + c = -e$, and (just like before) $c + e = e + c = -c$.

4. We do the same trick to add $e + e$, for $e = (3, 3)$; we have that the slope is $\frac{3x^2+2}{2y} = \frac{29}{6} = 4$, and thus that this line is $y = 4x + 1$.

   We then solve

$$\begin{aligned}(4x+1)^2 &= x^3 + 2x + 1\\ \Rightarrow 0 &= x^3 + 2x + 1 - 16x^2 - 8x - 1\\ &\equiv x^3 + 4x^2 + 4x \mod 5\\ &= x(x-3)^2\end{aligned}$$

   to see that our third point is at $x = 0$, and in particular is $(0, 1) = -a$. So we have $e + e = a$, and also $e + -a = -a + e = -e$.

We could solve for the rest of our points in this fashion, but we don't actually need to! We know that our elliptic curve's points form a group; therefore, if we simply look at the points we've placed so far and use our knowledge of groups (there are no repetitions in any row or column + we have associativity) we can add to our currently-known values! In particular, look at what we know so far:

$+$	$\mathcal{O}$	a	$-a$	c	$-c$	e	$-e$
$\mathcal{O}$	$\mathcal{O}$	a	$-a$	c	$-c$	e	$-e$
a	a	$-c$	$\mathcal{O}$	$-a$			
$-a$	$-a$	$\mathcal{O}$				$-e$	
c	c	$-a$		$-e$	$\mathcal{O}$	$-c$	
$-c$	$-c$			$\mathcal{O}$			
e	e		$-e$	$-c$		$-a$	$\mathcal{O}$
$-e$	$-e$					$\mathcal{O}$	

The cell for $e + a$ must be c and the cell for $e - c$ must be a; this is because the two remaining cells in e's row are c, a, and a cannot occur in the column corresponding to a!

$+$	$\mathcal{O}$	a	$-a$	c	$-c$	e	$-e$
$\mathcal{O}$	$\mathcal{O}$	a	$-a$	c	$-c$	e	$-e$
a	a	$-c$	$\mathcal{O}$	$-a$		c	
$-a$	$-a$	$\mathcal{O}$				$-e$	
c	c	$-a$		$-e$	$\mathcal{O}$	$-c$	
$-c$	$-c$			$\mathcal{O}$		a	
e	e	c	$-e$	$-c$	a	$-a$	$\mathcal{O}$
$-e$	$-e$					$\mathcal{O}$	

By repeatedly using logic like the above, you can fill in the rest of this table (yay, group table sudoku!) to get

$+$	$\mathcal{O}$	a	$-a$	c	$-c$	e	$-e$
$\mathcal{O}$	$\mathcal{O}$	a	$-a$	c	$-c$	e	$-e$
a	a	$-c$	$\mathcal{O}$	$-a$	$-e$	c	e
$-a$	$-a$	$\mathcal{O}$	c	e	c	$-e$	$-c$
c	c	$-a$	e	$-e$	$\mathcal{O}$	$-c$	a
$-c$	$-c$	$-e$	c	$\mathcal{O}$	e	a	$-a$
e	e	c	$-e$	$-c$	a	$-a$	$\mathcal{O}$
$-e$	$-e$	e	$-c$	a	$-a$	$\mathcal{O}$	c

Yay! □