Math 4527 (Number Theory 2)

Lecture #12 of 38 $\sim$ February 17, 2021

Cubic Curves and Elliptic Curves

- Cubic Curves and Weierstrass Form
- The Addition Law

This material represents §7.1.1-7.1.2 from the course notes.

Overview of Chapter, I

We now move into our next chapter $\sim$ §7: Elliptic Curves.

- Elliptic curves have a long and interesting history, and their study involves elements from most of the major disciplines of mathematics: algebra, geometry, analysis, number theory, topology, and even logic.
- Particularly, elliptic curves appear in the proofs of many deep results in mathematics. As I mentioned last class, they are a central ingredient in Wiles's proof of Fermat's Last Theorem.

Overview of Chapter, II

Our goals are fairly modest in comparison, but here is the plan:

- In §7.1 we will begin by outlining the basic algebraic and geometric properties of elliptic curves and motivate the group law, which establishes that the rational points on an elliptic curve have the structure of an abelian group, and study elliptic curves modulo p.
- Then in §7.2 we will explore these analogies and then to use them to convert certain cryptosystems and factorization algorithms that rely on modular arithmetic to ones that rely on elliptic curves.
- Finally, in §7.3 we will discuss some more advanced results about rational and integral points on elliptic curves and apply them to some applications of elliptic curves to Diophantine equations, such as the famous congruent number problem.

Cubic Curves, I

In elementary coordinate geometry, one begins by studying the behavior of lines in the plane, which have the general equation $ax + by + c = 0$, and then afterwards studies more general quadratic curves (the conic sections) having the general equation $ax^2 + bxy + cy^2 + dx + ey + f = 0$.

- In each case, we can do simple manipulations and changes of variable to put the equations into a more standard form.
- For example, if $b \neq 0$, we can rewrite the equation $ax + by + c = 0$ as $y = (-a/b)x + (-c/b)$, which for $m = -a/b$ and $b' = -c/b$ is the familiar $y = mx + b'$.

Cubic Curves, II

Similarly, if we have a quadratic relation $ax^2 + bxy + cy^2 + dx + ey + f = 0$ with $a \neq 0$, we can make a change of variable $x_1 = y + (b/(2a))x$ to remove the term bxy.

- This will yield an equation of the form $ax_1^2 + c_1y^2 + d_1x_1 + e_1y + f_1 = 0$ for new coefficients c_1, d_1, e_1, f_1.
- If $a, c_1 \neq 0$, we can complete the square in both x_1 and y by setting $x_2 = x_1 + d_1/(2a_1)$ and $y_2 = y + e_1/(2c_1)$, which eventually yields an equation having the much simpler form $ax_2^2 + c_1y_2^2 + f_2 = 0$. Otherwise, if a or c_1 is zero, by swapping variables if necessary and completing the square, we get a parabola $y_2 = a_2x_2^2 + f_2$.
- We conclude that every conic can be put into the form $ax^2 + cy^2 + f = 0$ or $y = ax^2 + f$ after changing coordinates.

Cubic Curves, II

The next step is to study cubic curves:
$ax^3 + bx^2y + cxy^2 + dy^3 + ex^2 + fxy + gy^2 + hx + iy + j = 0.$

- Like in the case of quadratic curves above, we can perform a series of changes of variable to reduce the general form to a simpler one.
- We will not give the full details of the procedure, as it is rather complicated.
- Instead, we will summarize matters by saying that as long as the equation is actually cubic (i.e., it is not the case that all of a, b, c, d are zero), then the general equation above can always be transformed using rational changes of variable into one of the form $y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$, for appropriate coefficients a_1, a_2, a_3, a_4, a_6.

Cubic Curves, III

We can illustrate the kind of procedure involved with a simple example: consider the cubic curve $x^3 + y^3 = 1$.

- There are various fairly natural ways to try to lower the degree in y, such as taking $x' = x + y$, but none of the obvious ones will give a y^2 term with coefficient 1.
- Here is one way to do it...

Cubic Curves, III

We can illustrate the kind of procedure involved with a simple example: consider the cubic curve $x^3 + y^3 = 1$.

- There are various fairly natural ways to try to lower the degree in y, such as taking $x' = x + y$, but none of the obvious ones will give a y^2 term with coefficient 1.
- Here is one way to do it...
  Set $a = 12/(x + y)$ and $b = 36(x - y)/(x + y)$.
- Then $x + y = 12/a$ and $x - y = b/(3a)$, so
  $x = (b + 36)/(6a)$ and $y = (b - 36)/(6a)$.
- Then $x^3 + y^3 = 1$ becomes $\left[\frac{b+36}{6a}\right]^3 + \left[\frac{b-36}{6a}\right]^3 = 1$,
  which after expanding simplifies to $b^2/a^3 + 432/a^3 = 1$.
- This last equation is equivalent to $b^2 = a^3 - 432$.

Elliptic Curves, I

Definition

An elliptic curve E over a field K is a curve having an equation of the form

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

for appropriate coefficients a_1, a_2, a_3, a_4, a_6 in K. This expression is called the Weierstrass form of E.

We will generally restrict our attention to the situation where K is one of the rational numbers $\mathbb{Q}$, the real numbers $\mathbb{R}$, the complex numbers $\mathbb{C}$, or the field $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ of integers modulo p.

Elliptic Curves, II

This expression $y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$ is not the simplest possible one in most cases.

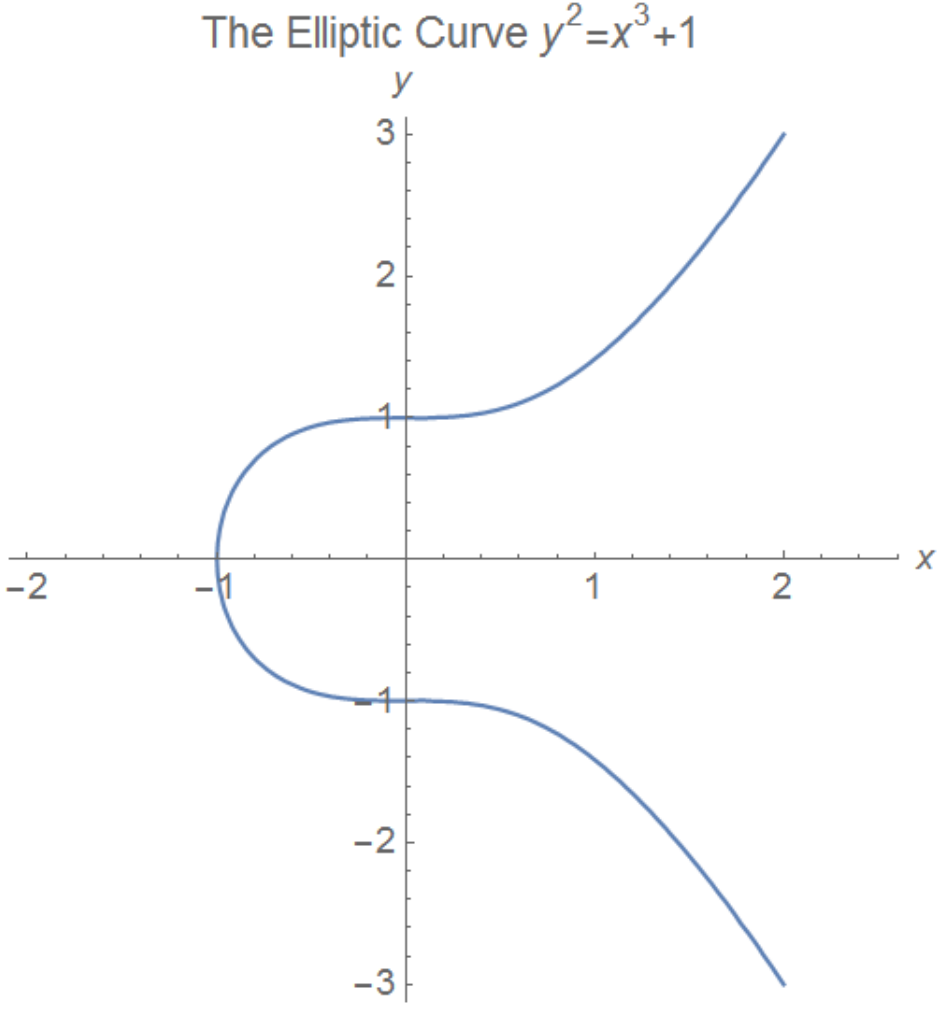
- As long as the characteristic of K is not 2 or 3, we can simplify it by completing the square in y and completing the cube in x.
- Explicitly, if we set $y' = y + (a_1/2)x + (a_3/2)$ and $x' = x + (a_2/3)$, we can reduce the Weierstrass equation above to one of the form $(y')^2 = (x')^3 + A(x') + B$.

An elliptic curve having an equation of the form $y^2 = x^3 + Ax + B$ is said to be in reduced Weierstrass form.

- This reduced form is much nicer to use. It is also nearly unique: the only change of variables that preserves it is one of the form $x = u^2x'$, $y = u^3y'$ for some nonzero u, from which we see that $A = u^4A'$ and $B = u^6B'$.

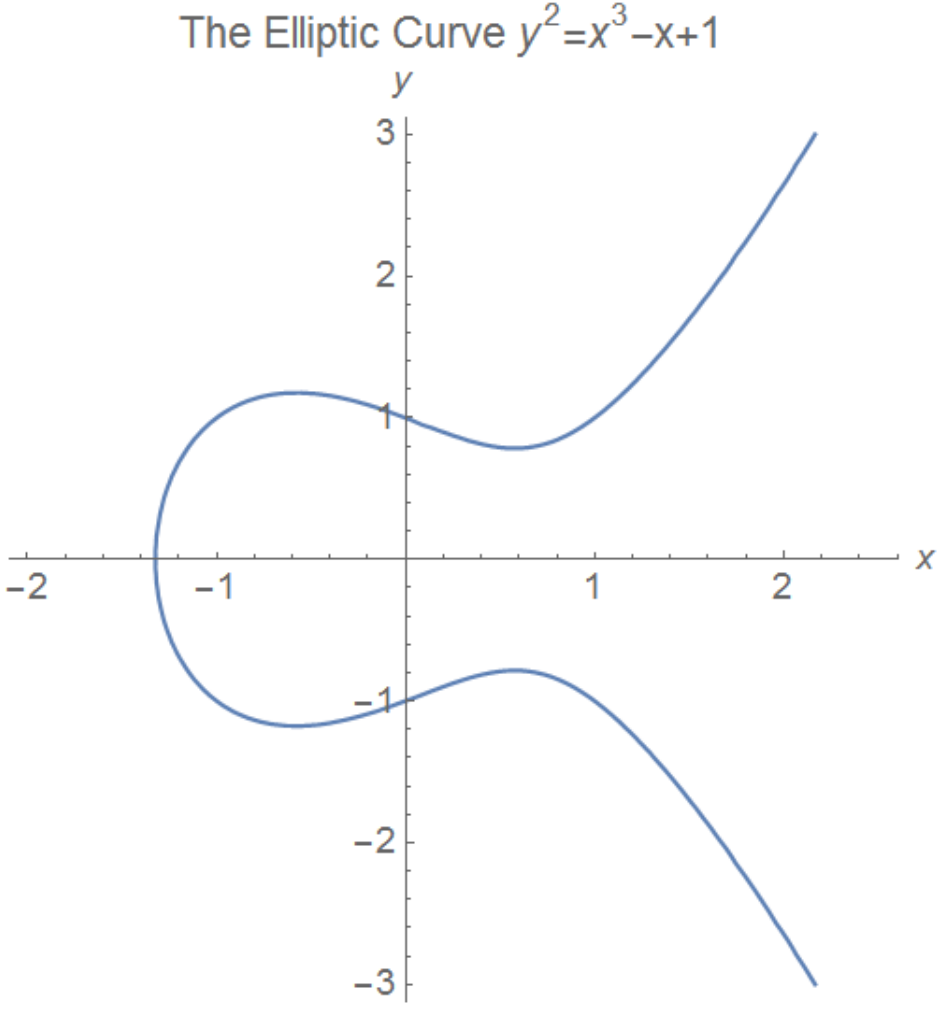
Elliptic Curves, III

When $K = \mathbb{R}$, we can draw graphs to visualize elliptic curves. Here is the graph of $y^2 = x^3 + 1$:



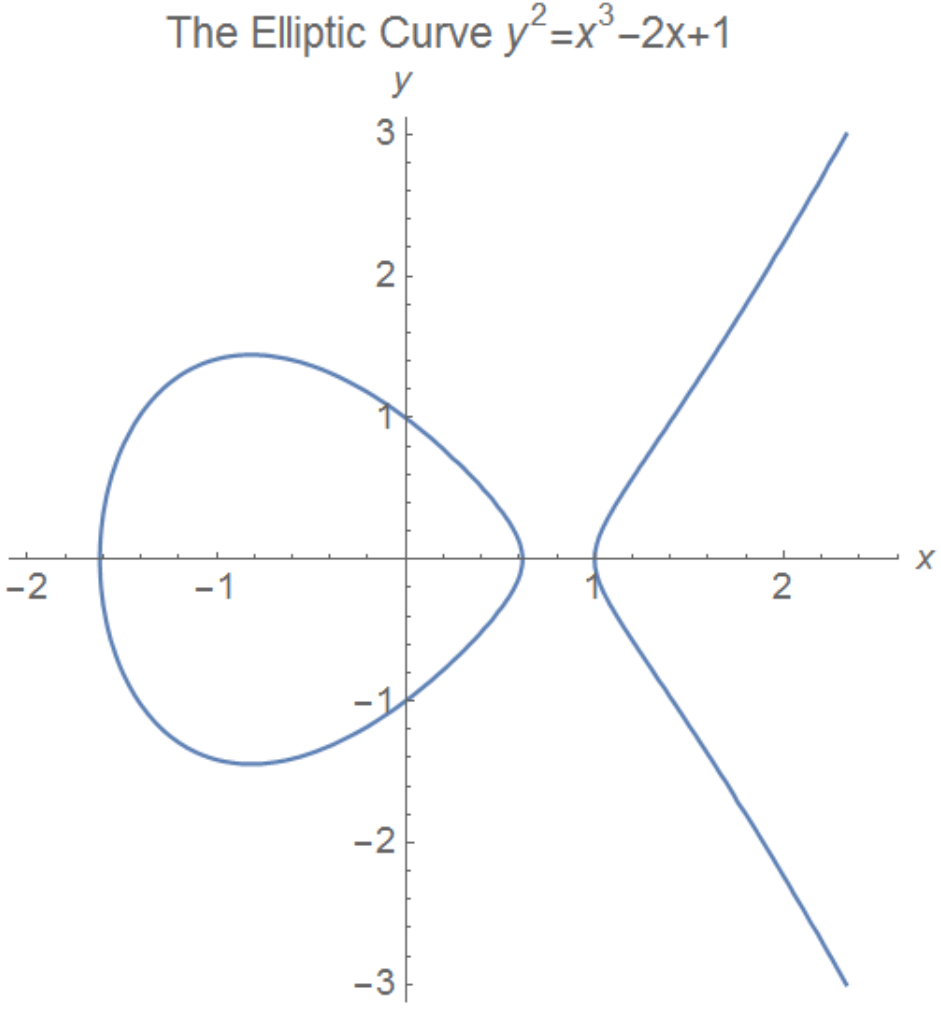
Elliptic Curves, IV

When $K = \mathbb{R}$, we can draw graphs to visualize elliptic curves. Here is the graph of $y^2 = x^3 - x + 1$:



Elliptic Curves, V

When $K = \mathbb{R}$, we can draw graphs to visualize elliptic curves. Here is the graph of $y^2 = x^3 - 2x + 1$:



Elliptic Curves, V

Observation #1: Elliptic curves are not ellipses!

- The reason for the similar name is that if one wants to compute the arclength of an ellipse (an elliptic integral), a few changes of variable will transform the resulting integral into one of the general form $\int \frac{1}{\sqrt{x^3 + Ax + B}}\, dx$.
- Upon setting $y = \sqrt{x^3 + Ax + B}$, we see that this elliptic integral is rather naturally related to the curve $y^2 = x^3 + Ax + B$.
- Sadly, except in very special circumstances (e.g., if $A = B = 0$), the elliptic integral given above is non-elementary.

Elliptic Curves, VI

Observation #2: The graph of an elliptic curve $y^2 = x^3 + Ax + B$ will always be symmetric about the x-axis.

- This is easy to see because since if (x, y) satisfies the equation then so does $(x, -y)$.
- By using this observation and invoking the implicit function theorem, we can see that the graph of an elliptic curve will have either one or two components depending on the values of the coefficients.
- Specifically, it will have two components when the polynomial $x^3 + Ax + B$ has three distinct real roots, and it will have one component otherwise.

Elliptic Curves, VII

Observation #3: The tangent line at each crossing of the x-axis is always vertical.

- Using implicit differentiation, we can compute $y' = \dfrac{3x^2 + A}{2y}$.
- Thus, we see that $y' = \infty$ when y is zero, provided that $3x^2 + A$ is not also zero.
- This behavior can only occur when $x^3 + Ax + B$ has a root in common with its derivative $3x^2 + A$, which is in turn equivalent to saying that $x^3 + Ax + B$ has a double root.

Elliptic Curves, VIII

Definition

If the polynomial $x^3 + Ax + B$ has a repeated root, we say that the elliptic curve $y^2 = x^3 + Ax + B$ is singular. Otherwise (if the roots are distinct) we say the elliptic curve is nonsingular.

We can give a simple way to identify whether a given elliptic curve is nonsingular:

Proposition (Singular Curves and the Discriminant)

The elliptic curve $y^2 = x^3 + Ax + B$ is singular if and only if its discriminant $\Delta = -16(4A^3 + 27B^2)$ is zero.

Remark: The -16 is superfluous here, but there is also a definition of Δ in terms of the original coefficients a_1, a_2, a_3, a_4, a_6 for a general Weierstrass form. To avoid having denominators in that expression, we do end up needing the factor of -16.

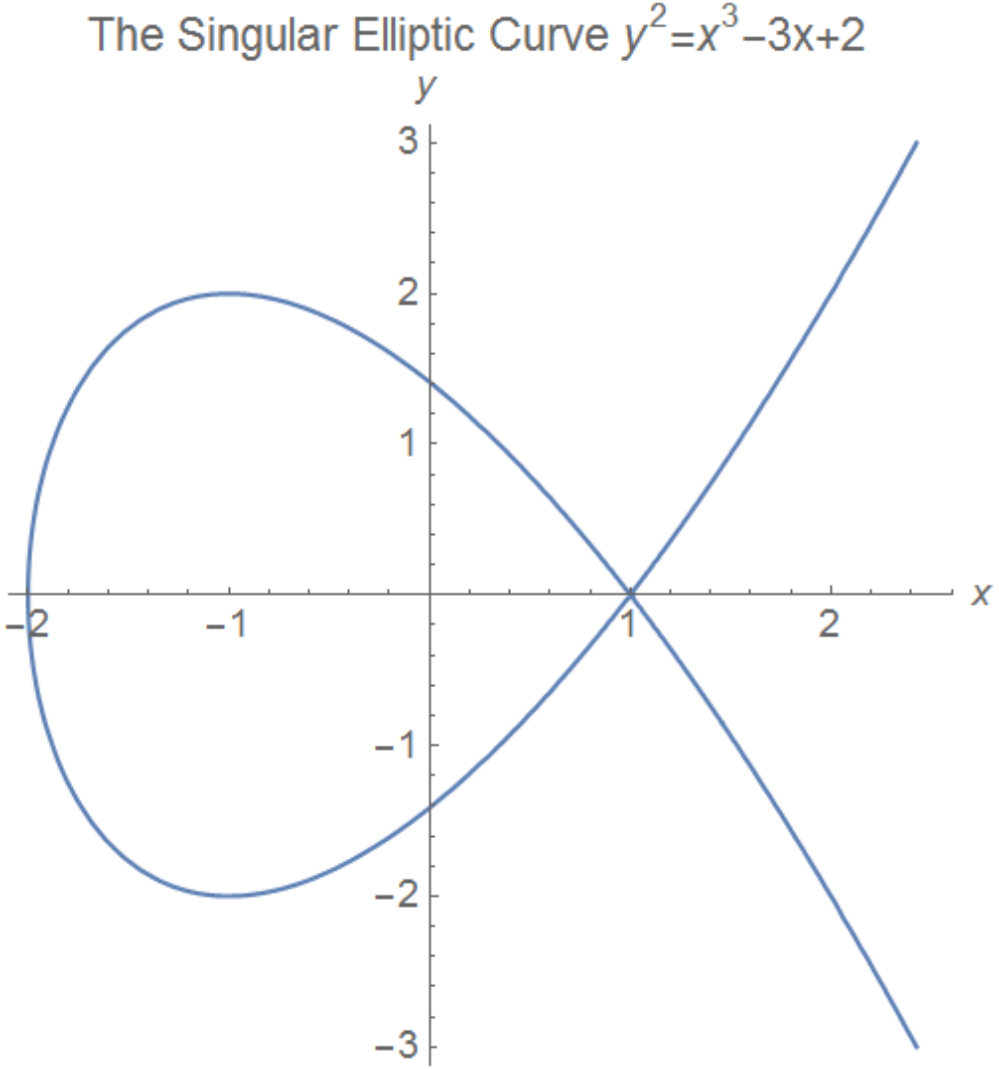
Elliptic Curves, IX

Proof:

- We first remark that if p is an arbitrary polynomial, then p has a repeated root r if and only if $p(r) = p'(r) = 0$.
- Explicitly, if $(x-r)^2$ divides $p(x)$, then $p(x) = (x-r)^2 q(x)$ for some q. Then $p'(x) = 2(x-r)q(x) + (x-r)^2 q'(x)$ and so $p'(r) = 0$.
- Conversely, if $p(r) = p'(r) = 0$ then $p(x) = (x-r)s(x)$ for some s, and then $p'(x) = s(x) + (x-r)s'(x)$ so $p'(r) = s(r)$. Thus s is also divisible by $x-r$ so $p(x)$ is divisible by $(x-r)^2$.
- In particular, $p(x) = x^3 + Ax + B$ has a repeated root if and only if it has a root in common with $p'(x) = 3x^2 + A$.
- This occurs iff $x^2 = -A/3$ which requires $x(2A/3) + B = 0$ so $x = -3B/(2A)$. Clearing denominators in $[-3B/(2A)]^2 = -A/3$ yields $\Delta = 0$ as claimed.

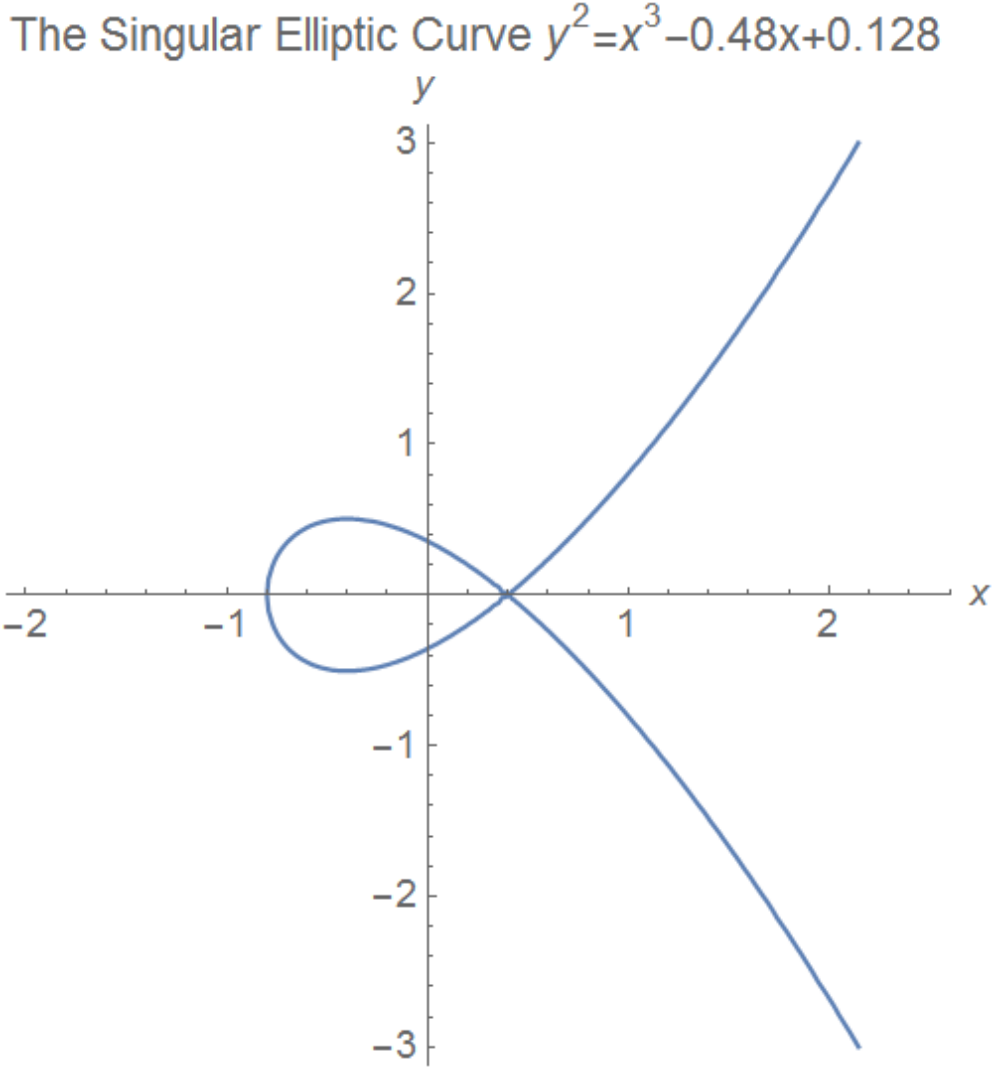
Elliptic Curves, X

Here is the graph of the singular elliptic curve $y^2 = x^3 - 3x + 2$:



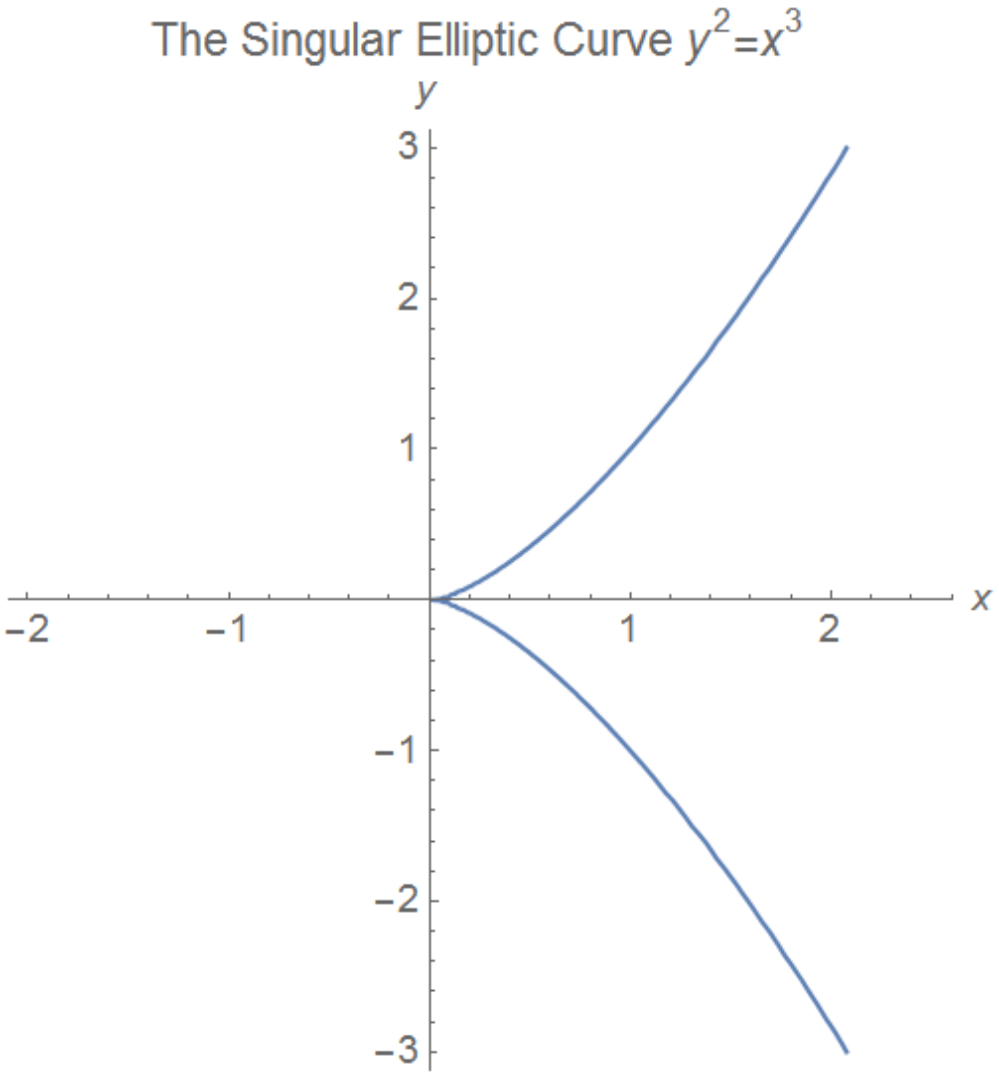
Elliptic Curves, XI

Here is the graph of the singular curve $y^2 = x^3 - 0.48x + 0.128$:



Elliptic Curves, XII

Here is the graph of the singular curve $y^2 = x^3$:



Elliptic Curves, XIII

Each of these three curves has one singular point (i.e., a point where the curve is nondifferentiable).

- On the first two curves, the singularity is where the curve crosses itself. This type of singularity is known as a node, and will occur when the polynomial $x^3 + Ax + B$ has a double root.
- The singular point on the third curve is the cusp at the origin $(0, 0)$. This type of singularity will occur when the polynomial $x^3 + Ax + B$ has a triple root, which can only happen when $A = B = 0$.

In general, singular elliptic curves tend to have unusual properties relative to nonsingular curves. We will therefore exclude singular elliptic curves and speak only of nonsingular elliptic curves from this point onward.

The Addition Law on an Elliptic Curve, I

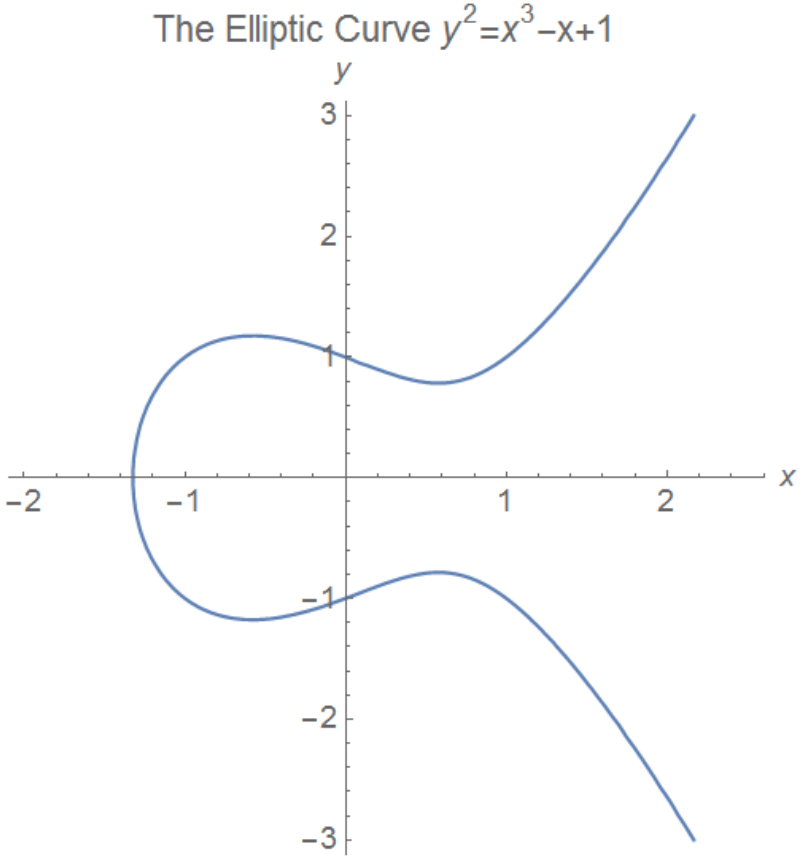
The key property of elliptic curves that make them so interesting (and useful!) is the following algebraic and/or geometric observation:

Observation

If we have two points that lie on an elliptic curve, we can use them to construct a third point on the curve.

The Addition Law on an Elliptic Curve, II

Here is an interactive "proof"[1] by picture (you pick two points and I'll give you a third one):



[1] This proof technique is not valid in mathematics. Your experience in other disciplines (physics, philosophy) may vary.

The Addition Law on an Elliptic Curve, III

Here is an actual argument: suppose $P_1 = (x_1, y_1)$, $P_2 = (x_2, y_2)$ are two distinct points on the elliptic curve E: $y^2 = x^3 + Ax + B$.

- Draw the line through P_1 and P_2: we claim that this line L must intersect E in a third point Q.
- To see this, suppose the line through P_1 and P_2 has equation $y = mx + b$. (Ignore the vertical-line case for now.)
- Then the intersection points of L with E are the solutions to the system $y = mx + b$ and $y^2 = x^3 + Ax + B$.
- Equivalently, we must solve $(mx + b)^2 = x^3 + Ax + B$, or $x^3 + (-m^2)x^2 + (A - 2mb)x + (B - b^2) = 0$.
- However, we already know that this cubic has the two roots $x = x_1$ and $x = x_2$, so it must have a third root: this gives us the third point Q we wanted.

The Addition Law on an Elliptic Curve, IV

Example: For the elliptic curve $E : y^2 = x^3 - 7x + 10$, find the third intersection point of E with the line through the two points $P_1 = (-3, 2)$ and $P_2 = (1, -2)$ on E.

The Addition Law on an Elliptic Curve, IV

Example: For the elliptic curve $E : y^2 = x^3 - 7x + 10$, find the third intersection point of E with the line through the two points $P_1 = (-3, 2)$ and $P_2 = (1, -2)$ on E.

- The equation of the line is $y = -x - 1$.
- Plugging this into the equation for E yields $(-x - 1)^2 = x^3 - 7x + 10$, or $x^3 - x^2 - 9x + 9 = 0$.
- We know this cubic has roots $x = -3, 1$ so we can easily find the factorization $(x + 3)(x - 1)(x - 3) = 0$, so the third root is $x = 3$, yielding $y = -4$.
- Thus, the other intersection point is $(3, -4)$.

The Addition Law on an Elliptic Curve, V

Once we construct a third point on an elliptic curve this way, we might try to find more points.

- If we try this procedure directly using our points P_1, P_2, and Q, however, we will not get anywhere: the line through any of these two points intersects the elliptic curve at the other point.
- However, we can also exploit the vertical symmetry of the curve to make new points: if $P = (x, y)$ lies on the curve, then the point $-P = (x, -y)$ also lies on the curve.
- We can then take lines through one of our starting points and this point $-P$ to find even more points on the curve.

The Addition Law on an Elliptic Curve, VI

Example: For the elliptic curve $E : y^2 = x^3 - 7x + 10$, with $P_1 = (-3, 2)$ and $P_2 = (1, -2)$ on E, we calculated a third point $Q = (3, -4)$. Find the third intersection point of the line through P_1 and $-Q$ with E.

The Addition Law on an Elliptic Curve, VI

Example: For the elliptic curve $E : y^2 = x^3 - 7x + 10$, with $P_1 = (-3, 2)$ and $P_2 = (1, -2)$ on E, we calculated a third point $Q = (3, -4)$. Find the third intersection point of the line through P_1 and $-Q$ with E.

- We have $-Q = (3, 4)$, so the line through P_1 and $-Q$ is $y = x/3 + 3$.
- Plugging this into the equation for E yields $(x/3 + 3)^2 = x^3 - 7x + 10$, or $x^3 - \frac{1}{9}x^2 - 9x + 1 = 0$.
- As before, we have two roots $x = -3$ and $x = 3$, so we can easily get the factorization $(x - 1/9)(x - 3)(x + 3) = 0$, so the third root is $x = 1/9$.
- Thus, the third intersection point is $(1/9, 82/27)$.

The Addition Law on an Elliptic Curve, VII

If we combine these two procedures (taking the third point on the line through two given points and then reflecting this point vertically), we can often generate many points on the curve starting from just two.

Definition (Group Law I)

If P_1 and P_2 are two distinct points on the elliptic curve $E : y^2 = x^3 + Ax + B$, let $Q = (x', y')$ be the third intersection point of E with the line L joining P_1 and P_2. We define the sum $P_1 + P_2$ to be the point $-Q = (x', -y')$.

- It is clear from the definition, but just to emphasize, the sum $P_1 + P_2$ is *not* the pointwise coordinate sum of P_1 and P_2!
- It is also not immediately clear why we define the sum of two points to be the reflection of Q rather than Q itself (though the name of the definition should give you a hint!).

The Addition Law on an Elliptic Curve, VIII

There is one other important issue we need to address now, however, which is the situation of having a vertical line that I ignored earlier.

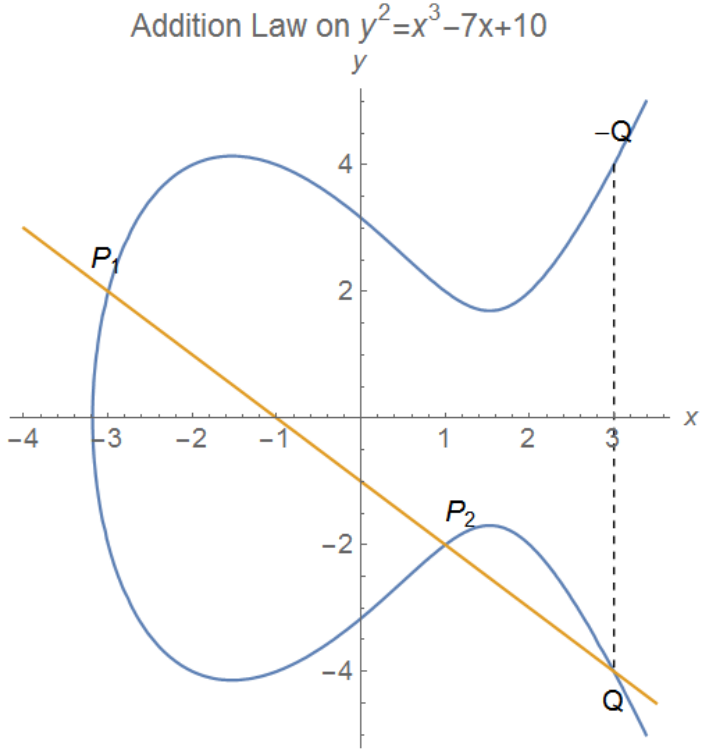
- Specifically, if we attempt to add two points which are vertical reflections of one another on the graph of $y^2 = x^3 + Ax + B$, the resulting line will not intersect the curve again.
- One option would simply be to declare that this operation is invalid. However, there is a much better approach: we will simply declare the curve E also includes a point at ∞ (which we denote simply as ∞) that we consider as lying on any vertical line.
- It will become clear very soon why this is the right convention.

The Addition Law on an Elliptic Curve, IX

Example: Given the points $P_1 = (-3, 2)$ and $P_2 = (1, -2)$ on the elliptic curve $y^2 = x^3 - 7x + 10$, find the following:

1. The sum $P_1 + P_2$.
2. The sum $(P_1 + P_2) + P_2$.

- Here is a plot of the curve and the line through the two points:



The Addition Law on an Elliptic Curve, X

Example: Given the points $P_1 = (-3, 2)$ and $P_2 = (1, -2)$ on the elliptic curve $y^2 = x^3 - 7x + 10$, find the following:

1. The sum $P_1 + P_2$.
2. The sum $(P_1 + P_2) + P_2$.

- We already found $P_1 + P_2 = (3, 4)$ earlier.
- To find the sum $(P_1 + P_2) + P_2$ we perform a similar procedure: the line through $P_1 + P_2 = (3, 4)$ and $P_2 = (1, -2)$ has equation $y = 3x - 5$.
- This yields $(3x - 5)^2 = x^3 - 7x + 10$ so that $x^3 - 9x^2 + 23x - 15 = 0$ which factors as $(x - 1)(x - 3)(x - 5) = 0$, and so the x-coordinate of $(P_1 + P_2) + P_2$ is 5.
- Thus, remembering to negate, we get $(P_1 + P_2) + P_2 = (5, -10)$.

The Addition Law on an Elliptic Curve, XI

We have essentially defined addition of points on an elliptic curve, except for one case: can we add a point to itself?

- Obviously, our approach of using the line through two points P and Q does not work correctly when $P = Q$.
- However, if P and Q are distinct points, then at least over the real numbers, $P + Q$ is a continuous function of the coordinates of the points.
- If we are working over $\mathbb{R}$, we could therefore define the addition $P + P$ to be the limit as $Q \to P$ of sums $P + Q$.
- Geometrically, the lines used in the construction also have a limit as $Q \to P$: they approach the tangent line to the curve E at the point P.
- Thus, a natural way to define $P + P$ is to let L be the tangent line to E at P, and then take Q to be the third point of intersection of L with E.

The Addition Law on an Elliptic Curve, XI

Here is our formal definition of this "doubling" law:

Definition (Group Law II)

If P is any point on the elliptic curve $E : y^2 = x^3 + Ax + B$, let $Q = (x', y')$ be the third intersection point of E with the tangent line L to E at P. We define the sum $P + P$ to be the point $-Q = (x', -y')$.

We can compute the slope of the tangent line to E at P using implicit differentiation[2].

[2] Finally, a useful application of implicit differentiation!

The Addition Law on an Elliptic Curve, XII

Example: Given the points $P_1 = (-3, 2)$ and $P_2 = (1, -2)$ on the elliptic curve $y^2 = x^3 - 7x + 10$, find $P_2 + P_2$ and $P_1 + (P_2 + P_2)$.

The Addition Law on an Elliptic Curve, XII

Example: Given the points $P_1 = (-3, 2)$ and $P_2 = (1, -2)$ on the elliptic curve $y^2 = x^3 - 7x + 10$, find $P_2 + P_2$ and $P_1 + (P_2 + P_2)$.

- Differentiating implicitly yields $2yy' = 3x^2 - 7$ so that $y' = (3x^2 - 7)/(2y)$. Thus, the tangent line to E at P_2 has slope 1 and its equation is $y = x - 3$.
- The point Q lies on the intersection of $y = x - 3$ and $y^2 = x^3 - 7x + 10$, so $(x - 3)^2 = x^3 - 7x + 10$ which yields $x^3 - x^2 - x + 1 = 0$. Factoring gives $(x + 1)(x - 1)^2 = 0$, so the third root has $x = -1$ and then $y = -4$.
- Remembering to negate, we see $P_2 + P_2 = (-1, 4)$.
- Using the regular addition procedure, we see the line through $P_1 = (-3, 2)$ and $P_2 + P_2 = (-1, 4)$ is $y = x + 5$ and so solving $(x + 5)^2 = x^3 - 7x + 10$ yields $(x + 1)(x + 3)(x - 5) = 0$ so $P_1 + (P_2 + P_2) = (5, -10)$.

The Addition Law on an Elliptic Curve, XIII

We just computed $(P_1 + P_2) + P_2 = (5, -10) = P_1 + (P_2 + P_2)$, and so here the addition law is actually associative. More is true:

Theorem (The Group Law)

If K is any field and E is any elliptic curve defined over K, then for any points P, P_1, P_2, and P_3 on E, the following are true:

1. *The addition law is commutative: $P_1 + P_2 = P_2 + P_1$.*
2. *The addition law is associative:* $(P_1 + P_2) + P_3 = P_1 + (P_2 + P_3)$.
3. *The point at ∞ is a two-sided identity: $P + \infty = P = \infty + P$.*
4. *The point P has a two-sided inverse $-P$:* $P + (-P) = \infty = (-P) + P$.

A concise way of phrasing this statement is to say that the set of points on E (including the point at ∞) forms an abelian group.

Summary

We introduced cubic curves and elliptic curves in Weierstrass form.

We discussed the addition law on elliptic curves.

Next lecture: Elliptic curves modulo p.