23MAC260 Elliptic Curves: Week 3

Last updated: February 20, 2024

1 Weierstrass form of a cubic

Last time we saw how to **add together** points on an elliptic curve.

This week we'll introduce a convenient form for elliptic curve equations, to make calculations easier.

Definition 1.1. *A cubic curve is in* **Weierstrass form** *if it is given by an equation*

$$Y^2Z = X^3 + aXZ^2 + bZ^3$$

for some coefficients a, b.

In affine form, this means the curve is given by an equation

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

Remark: The form above is sometimes called "short" Weierstrass form. In references, you may also see "medium" and "long" Weierstrass forms, but they are not relevant for us.

Examples: In these examples we stick to equations in affine form, for simplicity.

1. The equation
$$y^2 = 2x^3 + 3x + 9$$
is not in Weierstrass form.

2. The equation
$$y^2 = x^3 + 3x^2 + 4x$$
is not in Weierstrass form.

3. The equations
$$y^2 = x^3 + 28$$
$$y^2 = x^3 - 12x$$
are both in Weierstrass form.

The key theoretical result is the following:

Theorem 1.2. *Every elliptic curve can be transformed to Weierstrass form by an affine linear change of coordinates.*

Proof. Suppose our curve is given in affine form by an equation

$$y^2 = AX^3 + Bx^2 + Cx + D \tag{1}$$

for some coefficients A, B, C, D with $A \neq 0$.

Step 1: Eliminate A.

To do this let

$$x' = Ax \Leftrightarrow x = \frac{x'}{A}$$
$$y' = Ay \Leftrightarrow y = \frac{y'}{A}$$

Making these substitutions in Equation (1) we get

$$\left(\frac{y'}{A}\right)^2 = A\left(\frac{x'}{A}\right)^3 + B\left(\frac{x'}{A}\right)^2 + C\left(\frac{x'}{A}\right) + D$$

and multiplying across by A^2 gives

$$(y')^2 = (x')^3 + B(x')^2 + AC(x') + A^2D. \tag{2}$$

To simplify, let's write x instead of x' and y instead of y', and rename the coefficients in Equation 2 as follows:

$$B = \beta,\ AC = \gamma,\ A^2D = \delta.$$

Then Equation (2) becomes

$$y^2 = x^2 + \beta x^2 + \gamma x + \delta. \tag{3}$$

Step 2: Eliminate the x^2 term.

To do this, make the substitution

$$x' = x + \frac{\beta}{3} \Leftrightarrow x = x' - \frac{\beta}{3}$$

in Equation (3) to get

$$y^2 = \left(x' - \frac{\beta}{3}\right)^3 + \beta\left(x' - \frac{\beta}{3}\right)^2 + \gamma\left(x' - \frac{\beta}{3}\right) + \delta.$$

Expanding out and simplifying this becomes (check!)

$$y^2 = (x')^3 + \gamma' x' + \delta' \tag{4}$$

where

$$\gamma' = \gamma - \frac{1}{3}\beta^2$$
$$\delta' = \delta - \frac{1}{3}\beta\gamma + \frac{2}{27}\beta^3.$$

Again writing x instead of x' in Equation (4), we get the equation

$$y^2 = x^3 + \gamma' x + \delta'$$

which is in Weierstrass form as required. □

Remark: The coordinate changes we made in the course of the proof:

$$x \to \frac{x'}{A},\ y \to \frac{y'}{A}$$
$$x \to x' - \frac{\beta}{3}$$

are invertible affine linear transformations of the plane. In particular

(a) They transform a nonsingular curve to a nonsingular curve. So if we start with an elliptic curve, its Weierstrass form is still an elliptic curve.

(b) They map lines to lines, which implies that they give a **homomorphism** of the sets of points on the curve (with respect to the addition operation defined last time).

(c) They are bijective, which implies that the homomorphism in (b) above is an **isomorphism**.

The moral is that transforming a curve to Weierstrass form does not change any of its basic properties.

Example: Let's transform the curve given by the following equation to Weierstrass form:

$$y^2 = 2x^3 + 5x^2 + 4x + 2 \tag{5}$$

Step 1: This equation has $A = 2$, so we start by setting

$$x = \frac{x'}{A} = \frac{x'}{2}$$
$$y = \frac{y'}{A} = \frac{y'}{2}.$$

Making these substitutions in Equation (5) and clearing denominators we get

$$(y')^2 = (x')^3 + 5(x')^2 + 8x' + 8$$

or more simply

$$y^2 = x^3 + 5x^2 + 8x + 8. \tag{6}$$

So we have new coefficients $\beta = 5$, $\gamma = 8$, $\delta = 8$.

Step 2: Now put $x = x' - \frac{\beta}{3} = x' - \frac{5}{3}$ in Equation (6) to get

$$y^2 = \left(x' - \frac{5}{3}\right)^3 + 5\left(x' - \frac{5}{3}\right)^2 + 8\left(x' - \frac{5}{3}\right) + 8.$$

Our transformation was chosen so that the coefficient of $(x')^2$ turns out to be 0. Working out the other coefficients, the equation simplifies to (check!)

$$y^2 = x^3 - \frac{1}{3}x + \frac{106}{27}.$$

2 The discriminant of a cubic

Not every cubic in Weierstrass form defines an elliptic curve: for example, the equation $y^2 = x^3$ is in Weierstrass form, but the curve it defines has a singular point at the origin.

However, there is a simple criterion to check which Weierstrass forms give elliptic curves.

Definition 2.1. *Let C be a curve defined by an equation in Weierstrass form*

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

The quantity

$$\Delta = -4a^3 - 27b^2$$

is called the **discriminant** *of this curve.*

Theorem 2.2. *A Weierstrass equation*

$$y^2 = x^3 + ax + b \tag{7}$$

defines an elliptic curve if and only if

$$\Delta = -4a^3 - 27b^2 \neq 0.$$

In the module "Rings & Polynomials" we saw more generally that a polynomial in 1 variable has a multiple root if and only if its discriminant equals zero. But we can give a simpler proof for cubics in Weierstrass form as follows.

Proof. By definition, Equation (7) defines an elliptic curve if and only if the cubic $x^3 + ax + b$ has 3 distinct roots, or equivalently no multiple roots.

We saw in the module "Rings & Polynomials" that a multiple root of a polynomial $f \in \mathbb{C}[x]$ is the same thing as a common root of f and its derivative f'. Here

$$\begin{aligned} f &= x^3 + ax + b, \\ f' &= 3x^2 + a. \end{aligned}$$

Denote the roots of f' by $\pm\rho$, so $3\rho^2 + a = 0$, hence $\rho^2 = -\frac{a}{3}$.

Let us check if $\pm\rho$ is a root of $x^3 + ax + b$: this is the case if and only if

$$\begin{aligned} \rho^3 + a\rho + b = 0 \quad &\text{or} \\ -\rho^3 - a\rho + b = 0 & \end{aligned}$$

Using $\rho^2 = -\frac{a}{3}$ these conditions become

$$\begin{aligned} \frac{2a}{3}\rho + b = 0 \quad &\text{or} \\ -\frac{2a}{3}\rho + b = 0 & \end{aligned}$$

Multiplying the two equations together, this is equivalent to the single condition

$$-\frac{4a^2}{9}\rho^2 + b^2 = 0$$

and using $\rho^2 = -\frac{a}{3}$ again this becomes

$$\begin{aligned} \frac{4a^3}{27} + b^2 = 0; \quad &\text{that is,} \\ 4a^3 + 27b^2 = 0. & \end{aligned}$$

□

Example: Does the following Weierstrass equation define an elliptic curve?

$$y^2 = x^3 - 3x + 2 \tag{8}$$

We compute

$$\begin{aligned} -\Delta &= 4a^3 + 27b^2 \\ &= 4(-3)^3 + 27(2)^2 \\ &= 0 \end{aligned}$$

so this is not an elliptic curve.

We could also verify this as follows:

$$x^3 - 3x + 2 = (x-1)^2(x+2)$$

so there is a multiple root at $x = 1$.

Exercise: Draw the graph of Equation (8) to see geometrically why it does not define an elliptic curve.

3 Adding Points in Weierstrass Form

Last week we saw how to add points on an elliptic curve C. If P and Q are points on C, then:

- $P * Q$ is the 3rd point of the intersection $C \cap \overline{PQ}$;
- $P \oplus Q = O * (P * Q)$.

If C is given by a Weierstrass equation, then we can write down formulas for the coordinates of $P \oplus Q$ in terms of those of P and Q:

Theorem 3.1. *For an elliptic curve C defined by an equation in Weierstrass form:*

$$y^2 = x^3 + ax + b \tag{9}$$

addition of points on C is given by the following formulas:

(A) If $P, Q \in C$ are points with

$$P = (x_0, y_0), \quad Q = (x_0, -y_0)$$

then

$$P \oplus Q = O.$$

(B) If $P, Q \in C$ are points with

$$P = (x_0, y_0), \quad Q = (x_1, y_1) \quad x_0 \neq x_1$$

let

$$m = \frac{y_1 - y_0}{x_1 - x_0}, \quad x_2 = m^2 - x_0 - x_1, \quad y_2 = y_0 + m(x_2 - x_0);$$

then

$$P \oplus Q = (x_2, -y_2).$$

(C) If $P \in C$ is a point with $P = (x_0, y_0)$ where $y_0 \neq 0$, let

$$m' = \frac{3x_0^2 + a}{2y_0}, \quad x_1 = (m')^2 - 2x_0, \quad y_1 = y_0 + m'(x_1 - x_0);$$

then

$$P \oplus P = (x_1, -y_1).$$

Proof. (A): If $P \neq Q$, then we have seen that the line $\overline{PQ}$, which is the vertical line $x = x_0$, passes through the point at infinity O. So we get

$$\begin{aligned} P * Q &= O \quad \text{hence} \\ P \oplus Q &= O * O \\ &= O. \end{aligned}$$

Now suppose $P = Q = (x_0, 0)$. Let $f = y^2 - x^3 - ax - b$. The tangent line to C at P is given by the equation

$$\frac{\partial f}{\partial x}(P) \cdot (x - x_0) + \frac{\partial f}{\partial y}(P) \cdot (y - 0) = 0.$$

Since $\frac{\partial f}{\partial x}(P) \neq 0$ but $\frac{\partial f}{\partial y}(P) = 0$, this equation simplifies to $x = x_0$. So again the tangent line to C at P is a vertical line, and so again we get $P * P = O$ and $P \oplus P = O$,

(B): To find $P * Q$ we intersect the line $\overline{PQ}$ with C. The equation of the line $\overline{PQ}$ is

$$\begin{aligned} y &= mx + c \quad \text{where} \\ m = \frac{y_1 - y_0}{x_1 - x_0}, &\quad c = y_0 - mx_0. \end{aligned}$$

Substitute $y = mx + c$ into Equation (9) to get

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

which we can rearrange as

$$x^3 - m^2x^2 + (a - 2mc)x + (b - c^2) = 0.$$

The roots of this equation are the x-coordinates of P, Q, and $P * Q$, which we denote as x_0, x_1, and x_2.

For any cubic, the sum of the roots equal the negative of the coefficient of x^2. So here we get

$$x_0 + x_1 + x_2 = m^2 \quad \text{hence } x_2 = m^2 - x_0 - x_1.$$

If $P * Q = (x_2, y_2)$, then since $P * Q$ lies on the line $y = mx + c$, we get

$$\begin{aligned} y_2 &= mx_2 + c \\ &= mx_2 + (y_0 - mx_0) \\ &= m(x_2 - x_0) + y_0. \end{aligned}$$

Finally we get $P \oplus Q$ by changing the sign of the y-coordinate in $P * Q$: this gives

$$P \oplus Q = (x_2, -y_2)$$

as claimed.

(C): The proof of this part is almost the same as (B), except the slope of the tangent line to C at P is given by

$$m' = \frac{dy}{dx}(P).$$

From the equation $y^2 = x^3 + ax + b$ we get

$$2y\frac{dy}{dx} = 3x^2 + a.$$

Hence

$$m' = \frac{dy}{dx}(P) = \frac{3x_0^2 + a}{2y_0}.$$

□

Example: Let C be given by

$$y^2 = x^3 + 73.$$

This contains the points $P = (2, 9)$ and $Q = (3, 10)$. Let's compute $P \oplus Q$.

We have coordinates

$$\begin{aligned} x_0 &= 2, \quad y_0 = 9, \\ x_1 &= 3, \quad y_1 = 10. \end{aligned}$$

Then we use the formulas from Theorem 3.1 (B) to compute

$$\begin{aligned} m = \frac{10-9}{3-2} &= 1 \\ x_2 &= m^2 - x_0 - x_1 \\ &= -4 \\ y_2 &= y_0 + m(x_2 - x_0) \\ &= 3 \end{aligned}$$

Hence

$$\begin{aligned} P \oplus Q &= (x_2, -y_2) \\ &= (-4, -3). \end{aligned}$$

Exercise: Compute $2P = P \oplus P$.

3.1 Subfields of $\mathbb{C}$

Now let C be an elliptic curve and let K be a subfield of the complex numbers $\mathbb{C}$ — for example, the rational numbers $\mathbb{Q}$ or the real numbers $\mathbb{R}$. Then we can consider the points of C whose coordinates lie in K. Later in the module we will see that the case $K = \mathbb{Q}$ leads to a rich theory. For now, we will prove that this gives a subgroup of C.

Definition 3.2. *Let* K *be a subfield of* $\mathbb{C}$*. If* C *is an elliptic curve defined by an equation*

$$Y^2 = G(X, Z)$$

where all the coefficients of the cubic G *are in the field* K*, we say that the curve* C *is* **defined over** K*.*

If C *is defined over* K*, we define* **the set of** K**-points** *of* C *to be*

$$C(K) = \{P \in C \mid P \text{ has coordinates in } K\}.$$

We can use the Theorem 3.1 to show that the subset $C(K)$ is in fact a **subgroup** of C, with respect to the addition operation $\oplus$.

Corollary 3.3. *Let* K *be a subfield of* $\mathbb{C}$*, and let* C *be an elliptic curve defined over* K*. If* $P, Q \in C(K)$*, then* $P \oplus Q \in C(K)$*.*

Proof. The proof of Theorem 1.2 shows that if C is defined over K, then so is its Weierstrass form. If $P, Q \in C(K)$, then their images under the coordinate changes also have coordinates in K; since the coordinate changes give a homomorphism with respect to $\oplus$, the image of $P \oplus Q$ is the sum of the images of P and Q. So it suffices to prove the claim for C in Weierstrass form.

If P or Q or both equal the idenity O, there is nothing to prove. So we may assume that P and Q are different from O, hence they are both affine points on C.

First assume $P \neq Q$. Let $P = (x_0, y_0)$ and $Q = (x_1, y_1)$. If x_0, y_0, x_1, y_1 are all in K, then the formulas of Theorem 3.1 show that $m \in K$ and hence x_2 and y_2 are also in K. Hence $P \oplus Q$ has coordinates in K too.

Similarly if P has coordinates in K, then $m' \in K$, hence $P \oplus P$ has coordinates in K also. □