

ELLIPTIC CURVES OVER FINITE FIELDS

JONATHAN YU

ABSTRACT. Determining the set of all K -rational point on an arbitrary curve is a very difficult problem in number theory. While there is no known way to do this in general, it can be done for simple curves like lines and conics. For elliptic curves over finite fields, Hasse's theorem can be used to put a bound on the number of possible K -rational points. For more general curves over finite fields, the Hasse-Weil bound can be used.

1. INTRODUCTION

The theory of Diophantine equations, named after the Greek mathematician Diophantus of Alexandria, is a branch of mathematics that deals with integer and sometimes rational solutions to polynomial equations. A famous example of a Diophantine equation is the one that appears in Fermat's last theorem. The theorem states that for any integer $n > 2$, there are no positive integer solutions to the equation

$$a^n + b^n = c^n.$$

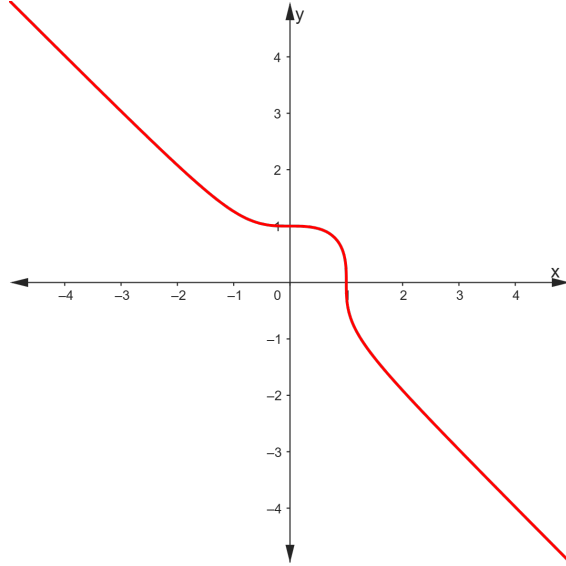
Pierre de Fermat first proposed this theorem in the margin of his copy of the book *Arithmetica* around 1637, stating, "It is impossible to separate a cube into two cubes, or a fourth power into two fourth powers, or in general, any power higher than the second, into two like powers. I have discovered a truly marvelous proof of this, which this margin is too narrow to contain." Although it is quite unlikely that he had a valid proof of this fact, he did manage to prove it for $n = 4$. The problem remained unsolved for more than 350 years. Then in 1995, Andrew Wiles, with some help from Richard Taylor, published in a pair of groundbreaking papers a proof of the theorem. This proof relied on a connection between a special type of cubic curve known as an elliptic curve and modular forms.

Elliptic curves, contrary what their name suggests, are not really that closely related to ellipses. Ellipses are a type of conic section, which is a degree 2 curve. The term arose when people were studying how to compute the arc length of an ellipse. If one makes a certain elementary substitution into the integral for the arc length of an ellipse, the integrand will involve the form $y = \sqrt{f(x)}$, where $f(x)$ is either a cubic or quartic polynomial. The solutions to such an integral can be written in terms of functions related to the curve $y^2 = f(x)$.

Fermat's last theorem holds for all nonzero integers a , b and c . To see this, note if n is even, we get $a^n + b^n = c^n$, regardless of whether or not we negative any of the values, and if n is odd, we can rearrange the equation to get an equation resembling the one in the theorem. If we divide through by c^n , we get that the only rational solutions to the equation

$$x^n + y^n = 1$$

are the ones with either $x = 0$ or $y = 0$.

FIGURE 1. $x^3 + y^3 = 1$

Consider the curve $x^3 + y^3 = 1$ (see figure 1). One natural question we might ask is what rational points lie on this curve. While it may seem like there could be infinitely many, Fermat's last theorem tells us that the only ones are $(0, 1)$ and $(1, 0)$.

In general, classifying the rational points on curves is not this easy. In this paper, we will begin by studying the degree 1 and 2 versions of this problem. Next, we will introduce a special class of degree 3 polynomials known as elliptic curves. After building up a bit of theory, we will present a theorem that puts bounds on the number of rational points that can exist on a curve.

2. PRELIMINARIES

Definition 2.1 (Group). Let G be a set equipped with a binary operation $\cdot : G^2 \rightarrow G$ such that

1. there exists an identity element 1 such that for all $g \in G$, $1 \cdot g = g \cdot 1 = g$
2. for every $g \in G$, there exists an inverse element $g^{-1} \in G$ such that $g \cdot (-g) = (-g) \cdot g = 1$
3. for all $g, h, i \in G$, $(g \cdot h) \cdot i = g \cdot (h \cdot i)$.

We call G a *group*. The pair (G, \cdot) denotes the group G under \cdot .

Definition 2.2 (Abelian group). Let (G, \cdot) be a group. If G also has the property that for all g, h in G , $g \cdot h = h \cdot g$, then G is called an *Abelian group*.

Definition 2.3 (Field). Let K be a set equipped with two binary operations $+: K^2 \rightarrow K$ and $\cdot: K^2 \rightarrow K$ such that

1. $(K, +)$ forms an Abelian group
2. $(K \setminus \{0\}, \cdot)$ forms an Abelian group, where 0 is the identity element of the group $(K, +)$
3. for all $a, b, c \in K$, $a \cdot (b + c) = a \cdot b + a \cdot c$.

We call F a *field*. We denote the field equipped with $+$ and \cdot by $(F, +, \cdot)$.

Definition 2.4 (Finite field). Let F be a field. If F has a finite number of elements, we call F a *finite field*.

Definition 2.5 (Subfield). Let L be a field. A *subfield* of L is a field $K \subseteq L$ whose operations are inherited from L .

Definition 2.6 (Extension field). Let L be a field and K a subfield of L . We call L the *extension field* of K .

Definition 2.7. Let K be a field. An algebraic object is said to be *defined over K* if its coefficients lie in K .

Definition 2.8 (Splitting field). Let K be a field and $f(x)$ a nonzero polynomial over K . If L is the smallest field extension of K such that $f(x)$ can be completely decomposed into a product of linear factors, then L is called the *splitting field*.

Definition 2.9 (Seperable). Let K be a field and $f(x)$ a nonzero polynomial over K . If every root of $f(x)$ is distinct in its splitting field, we say that $f(x)$ is *seperable*.

Proposition 2.10. Let $f(x)$ a polynomial over a field K and α be an arbitrary root. The polynomial $f(x)$ is *seperable* if and only if $f'(\alpha) \neq 0$.

Proof. We start by proving that if $f(x)$ is seperable, $f'(\alpha) \neq 0$. Assume that $f(x)$ is a degree n polynomial. We have

$$f(x) = c \prod_{i=1}^n (x - \alpha_i),$$

where $c \in K \setminus \{0\}$ and each α_i is a distinct root of $f(x)$ in its splitting field. Taking the formal derivative, we get

$$f'(x) = c \sum_{j=1}^n \prod_{\substack{i=1 \\ i \neq j}}^n (x - \alpha_i).$$

Let $\alpha = \alpha_k$ for some $1 \leq k \leq n$. We claim that

$$f'(\alpha) = c \sum_{j=1}^n \prod_{\substack{i=1 \\ i \neq j}}^n (\alpha - \alpha_i) \neq 0.$$

For every $j \neq k$, $\prod_{\substack{i=1 \\ i \neq j}}^n (\alpha - \alpha_i) = 0$, since it contains the factor $\alpha - \alpha_k$. When $j = k$, $\prod_{\substack{i=1 \\ i \neq j}}^n (\alpha - \alpha_i)$

will not contain $\alpha - \alpha_k$. Since every α_i is distinct, $\prod_{\substack{i=1 \\ i \neq j}}^n (\alpha - \alpha_i) \neq 0$, so $f'(\alpha) \neq 0$. Therefore,

if α is a root of $f(x)$ and $f(x)$ is seperable, $f'(\alpha) \neq 0$.

Now we show that if $f'(\alpha) \neq 0$, $f(x)$ is seperable. We proceed by contraposition. We need to show that if $f(x)$ is not seperable, $f'(\alpha) = 0$. If $f(x)$ is not seperable, it must contain a repeated root α , so we can write

$$f(x) = (x - \alpha)^2 g(x).$$

Taking the formal derivative gives us

$$f'(x) = 2(x - \alpha)g(x) + (x - \alpha)^2 g'(x).$$

We have

$$f'(\alpha) = 2(\alpha - \alpha)g(x) + (\alpha - \alpha)^2g'(x) = 0.$$

Therefore, $f(x)$ is separable if and only if for an arbitrary root α , $f'(\alpha) \neq 0$. \square

Definition 2.11 (Homomorphism). Let $(G, +)$ and (H, \cdot) be groups. If $\varphi: G \rightarrow H$ is a function such that for all $g, h \in G$,

$$\varphi(g + h) = \varphi(g) \cdot \varphi(h),$$

we call φ a *homomorphism*.

Definition 2.12 (Endomorphism). Let $\varphi: E(\overline{K}) \rightarrow E(\overline{K})$ be a group homomorphism given by rational functions. We call φ an *endomorphism* of E .

Definition 2.13 (Degree). Let $\varphi = \frac{p(x)}{q(x)}$ be an endomorphism. The *degree* of φ , denoted $\deg(\varphi)$, is

$$\deg(\varphi) = \max\{\deg(p(x)), \deg(q(x))\}.$$

Definition 2.14 (Characteristic). The *characteristic* of a field K , denoted $\text{char}(K)$, is the smallest positive integer n such that $\underbrace{1 + 1 + \cdots + 1}_{n \text{ times}} = 0$. If no such n exists, the characteristic is defined to be 0.

Definition 2.15 (Kernel). Let G and H be groups. The *kernel* of a homomorphism $\varphi: G \rightarrow H$, denoted $\ker(\varphi)$, is

$$\ker(\varphi) = \{g \in G \mid \varphi(g) = e_H\},$$

where e_H is the identity element of H .

Definition 2.16. The set of common solutions to a finite set of polynomial equations over a field K is called an *algebraic variety*.

Definition 2.17. Let $L \supseteq K$ be a field. The set of *L -rational points* on an algebraic variety X defined over K is the set of points on X with coordinates in L . This is denoted by $X(L)$.

Definition 2.18. Let X and Y be algebraic varieties over an arbitrary field. Let $f: X \dashrightarrow Y$ be a rational map (a map given by a ratio of polynomials). If there exists a rational map $g: Y \dashrightarrow X$ such that both $f \circ g$ and $g \circ f$ are the identity map wherever they are defined, we say that f and g are *birational maps* and that X and Y are *birationally equivalent*.

Definition 2.19 (Algebraic extension). Let K and L be fields such that $K \subseteq L$. If there exists a nonconstant polynomial $f(x)$ over K such that $f(a) = 0$ for every $a \in L$, we call L an *algebraic extension* of K .

Definition 2.20 (Algebraic closure). Let \overline{K} be a field containing K . If \overline{K} is an algebraic extension of K and every nonconstant polynomial over \overline{K} has a root in \overline{K} , then \overline{K} is an *algebraic closure* of K .

Proposition 2.21. Let \mathbb{F}_q be the field with $q = p^n$ for some prime p and positive integer n . Then

$$\overline{\mathbb{F}}_q = \bigcup_{i=1}^{\infty} \mathbb{F}_{q^i}$$

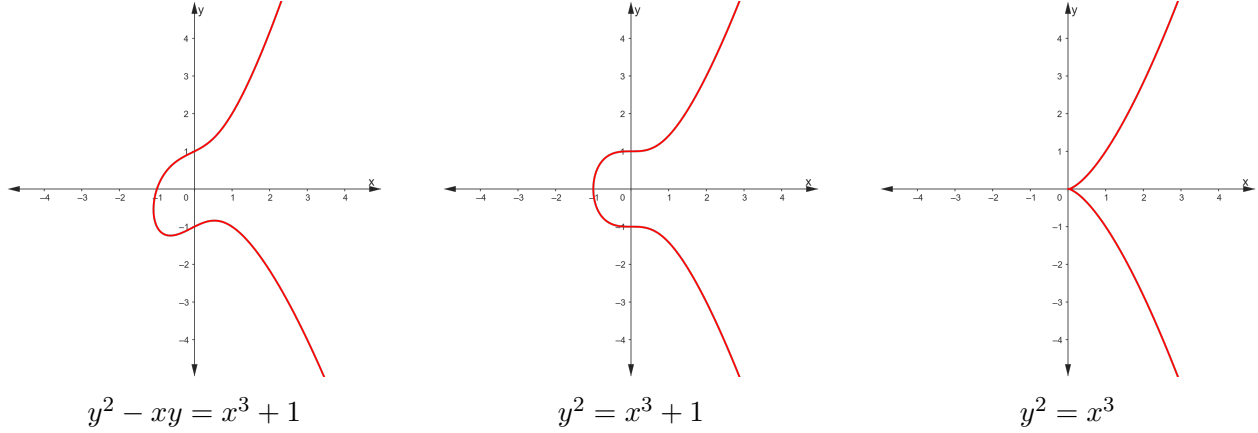


FIGURE 2. General Weierstrass Curves

Proposition 2.22. *Let K be a field. Then \overline{K} has an infinite number of elements.*

Proof. If K has a finite number of elements, we can write it as \mathbb{F}_q , where $q = p^n$ for some prime p and positive integer n . By proposition 2.21, $\overline{\mathbb{F}_q} = \bigcup_{i=1}^{\infty} \mathbb{F}_{q^i}$. We have $\mathbb{F}_{q^m} \subseteq \mathbb{F}_{q^n}$ if and only if $m \mid n$. Since there are an infinite number of primes, we have an infinite number of unions that are not subsets of each other. Therefore, $\overline{\mathbb{F}_q}$ has an infinite number of elements.

If K has an infinite number of elements, its algebraic closure trivially has an infinite number of elements, since by definition, its algebraic closure must contain K . Therefore, for any field K , \overline{K} contains an infinite number of elements. \square

Definition 2.23 (General Weierstrass Equation). A *general Weierstrass equation* over a field K is

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

where $a_1, a_2, a_3, a_4, a_6 \in K$ (see figure 2).

Definition 2.24 (Elliptic Curve). Let E be the set of solutions to the general Weierstrass equation and Δ the discriminant of the equation. If $\Delta \neq 0$, we call E an *elliptic curve* (see figure 3).

Definition 2.25 (Singular Weierstrass Curve). Let E be the set of solutions to

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

such that the discriminant $\Delta = 0$. We call E a *singular Weierstrass curve* (see figure 4). Points where the curve has multiple roots are called *singular points*.

The general Weierstrass equation tends to lead to overly complicated equations. We can often rewrite it in a simpler form.

Proposition 2.26. *Let K be a field such that $\text{char}(K) \neq 2, 3$. The general Weierstrass equation can be rewritten in the form*

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

for $a, b \in K$. This is known as a *short Weierstrass equation*.

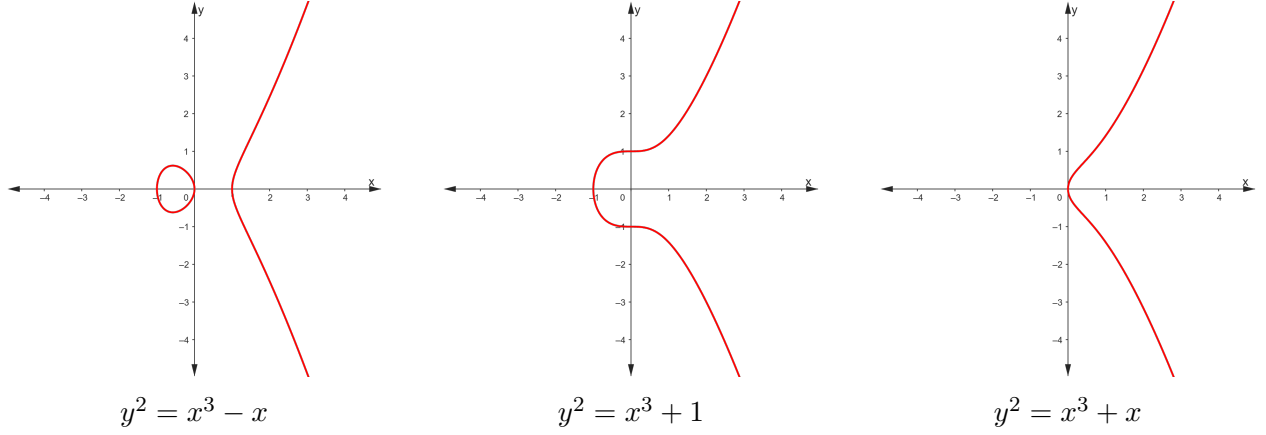


FIGURE 3. Elliptic Curves

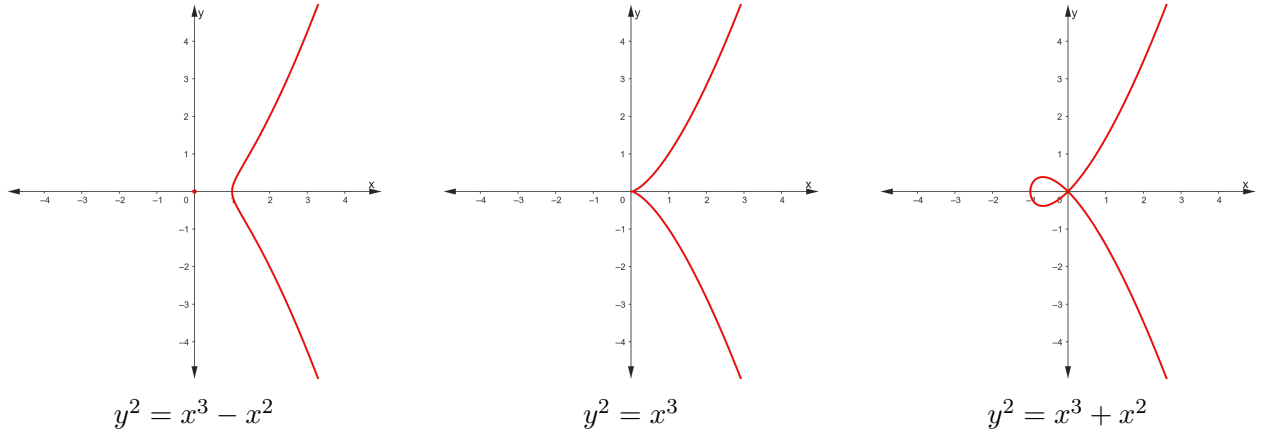


FIGURE 4. Singular Weierstrass Curves

Proof. Let $y = y' - \frac{a_1x + a_3}{2}$. This is valid if $\text{char}(K) \neq 2$. We have

$$\begin{aligned}
 y^2 + a_1xy + a_3y &= \left(y' - \frac{a_1x + a_3}{2}\right)^2 + a_1x \left(y' - \frac{a_1x + a_3}{2}\right) + a_3 \left(y' - \frac{a_1x + a_3}{2}\right) \\
 &= y'^2 - a_1xy' - a_3y' + \frac{a_1^2}{4}x^2 + \frac{a_1a_3}{2}x + \frac{a_3^2}{4} + a_1xy' - \\
 &\quad \frac{a_1^2}{2}x^2 - \frac{a_1a_3}{2}x + a_3y' - \frac{a_1a_3}{2}x - \frac{a_3^2}{2} \\
 &= y'^2 - \frac{a_1^2}{4}x^2 - \frac{a_1a_3}{2}x - \frac{a_3^2}{4}
 \end{aligned}$$

so

$$y'^2 = x^3 + \frac{a_1^2 + 4a_2}{4}x^2 + \frac{a_1a_3 + 2a_4}{2}x + \frac{a_3^2 + 4a_6}{4}.$$

Let $a'_2 = \frac{a_1^2 + 4a_2}{4}$, $a'_4 = \frac{a_1a_3 + 2a_4}{2}$, and $a'_6 = \frac{a_3^2 + 4a_6}{4}$, so that we get

$$y'^2 = x^3 + a'_2x^2 + a'_4x + a'_6.$$

If we define $x' = x - \frac{a'_2}{3}$, which is valid for $\text{char}(K) \neq 3$, we get

$$\begin{aligned} y'^2 &= \left(x - \frac{a'_2}{3}\right)^3 + a_2 \left(x - \frac{a'_2}{3}\right)^2 + a'_4 \left(x - \frac{a'_2}{3}\right) + a'_6 \\ &= x^3 - a'_2 x^2 + \frac{a_2'^2}{3} x - \frac{a_2'^3}{27} + a'_2 x^2 - \frac{2a_2'^2}{3} x + \frac{a_2'^3}{9} + a'_4 x - \frac{a'_2 a'_4}{3} + a'_6 \\ &= x^3 - \frac{2a_2'^2 - 3a'_4 + a_2'^2}{3} x + \frac{2a_2'^3 - 9a'_2 a'_4 + 27a'_6}{27} \end{aligned}$$

Letting $a = -\frac{2a_2'^2 - 3a'_4 + a_2'^2}{3}$ and $b = \frac{2a_2'^3 - 9a'_2 a'_4 + 27a'_6}{27}$ gives us the equation $y'^2 = x^3 + ax + b$. \square

Theorem 2.27 (Bézout's theorem). *Let K be an algebraically closed field and C_1 and C_2 curves in $\mathbb{P}^2(K)$ such that C_1 and C_2 do not have any irreducible factors in common. Define C_1 and C_2 by homogeneous polynomials of degrees m and n respectively. Then C_1 and C_2 intersect at exactly mn points, counting multiplicity.*

3. RATIONAL POINTS ON CURVES

We begin our study of elliptic curves with degree 1 and 2 polynomials.

Consider a polynomial of degree 1 over \mathbb{Q} . This is simply a line with rational coefficients. We know that the general form of such a curve is $ax + by + c = 0$ for $a, b, c \in \mathbb{Q}$ such that at least one of a and b is nonzero.

Proposition 3.1. *Two nonparallel rational lines intersect at a rational point (a point with rational coordinates).*

Proof. Let $a_1x + b_1y + c_1 = 0$ and $a_2x + b_2y + c_2 = 0$ be the two nonparallel rational lines. Since they are not parallel, their slopes $-\frac{b_1}{a_1}$ and $-\frac{b_2}{a_2}$ are not equal. This tells us that $a_1b_2 - a_2b_1 \neq 0$. Solving the system gives us $(x, y) = \left(\frac{b_1c_2 - b_2c_1}{a_1b_2 - a_2b_1}, \frac{a_2c_1 - a_1c_2}{a_1b_2 - a_2b_1}\right)$. Since $a_1, b_1, c_1, a_2, b_2, c_2 \in \mathbb{Q}$, (x, y) must be a rational point. \square

Proposition 3.2. *There exist an infinite number of rational points on the rational line $ax + by + c = 0$.*

Proof. Let $b \neq 0$ in $ax + by + c = 0$. Solving for y gives us $y = -\frac{ax + c}{b}$. Let $x_0 \in \mathbb{Q}$ and $y_0 = -\frac{ax_0 + c}{b}$. Since $a, b, c \in \mathbb{Q}$, we must have $y_0 \in \mathbb{Q}$. Therefore, (x_0, y_0) is a rational point. Because it also satisfies the equation $ax_0 + by_0 + c = 0$, the point must lie on the line. Since there are an infinite number of choices for the rational number x_0 , there must be an infinite number of rational points on the rational line.

If $b = 0$, we have the line $x = -\frac{c}{a}$. Since $a, c \in \mathbb{Q}$, $x \in \mathbb{Q}$. Let $x_0 = -\frac{c}{a}$ and $y_0 \in \mathbb{Q}$. Then (x_0, y_0) is a rational point that satisfies the equation $x_0 = -\frac{c}{a}$, so it lies on the line. Since there are an infinite number of rational choices for y_0 , the rational line also has an infinite number of rational points in this case. Therefore, any rational line has an infinite number of rational points. \square

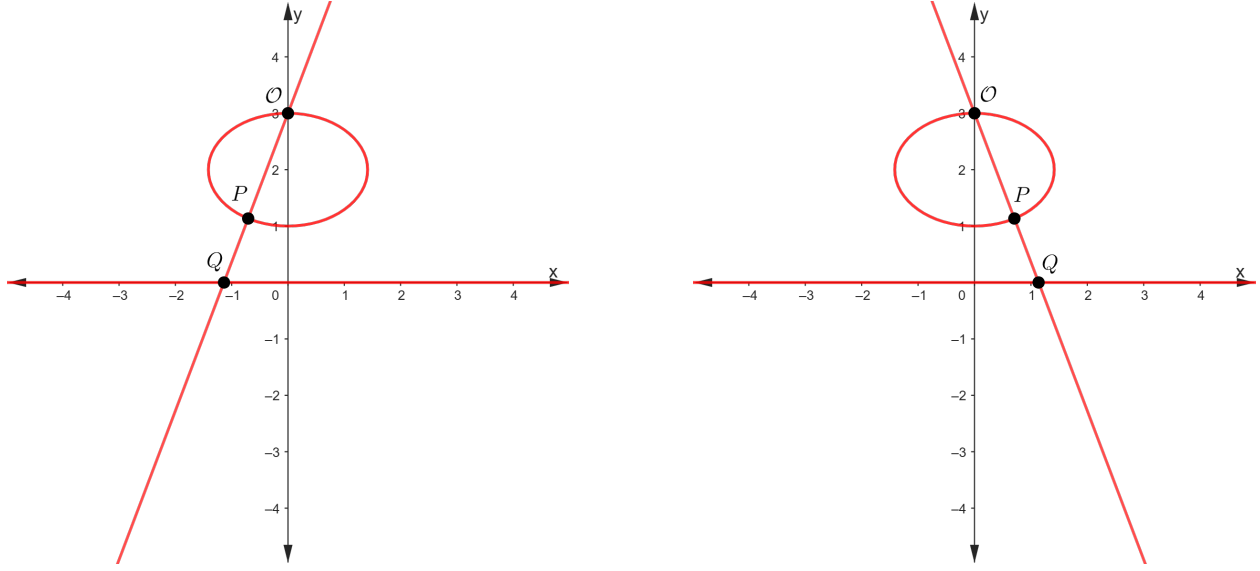


FIGURE 5. Line-Conic Correspondence

Things get quite a bit more interesting for polynomials of degree 2. These are our usual conic sections, which can be written in the form $ax^2 + bxy + cy^2 + dx + ey + f = 0$ for constants $a, b, c, d, e, f \in \mathbb{Q}$ such that at least one of a, b , and c is nonzero. We will assume that such a conic is nondegenerate.

Proposition 3.3. *There is a one-to-one correspondence between the points on a rational line and a rational conic for all but one point (see figure 5).*

Lemma 3.4. *Any line can be transformed into the rational line $y = 0$.*

Proof. Let $ax + by + c = 0$ be our line. Assume that $b \neq 0$. Dividing through by b gives us $\frac{a}{b}x + y + \frac{c}{b} = 0$. If we let $x' = x$ and $y' = \frac{a}{b}x + y + \frac{c}{b}$, we get $y' = 0$. Therefore, the transformation $(x', y') = \left(x, \frac{a}{b}x + y + \frac{c}{b}\right)$ transforms our line into the line $y = 0$ if $b \neq 0$.

If $b = 0$, we have the vertical line $ax + c = 0$. Let $x' = x$ and $y' = ax + c$. Then we get $y' = 0$, so the transformation $(x', y') = (x, ax + c)$ transforms our line into the line $y = 0$ if $b = 0$. Therefore, any line can be transformed into the line $y = 0$. \square

Lemma 3.5. *The transformation given in lemma 3.4 takes a rational conic to a rational conic.*

Proof. Let $ax^2 + bxy + cy^2 + dx + ey + f = 0$ be a conic. First, we consider the transformation $(x', y') = \left(x, \frac{a}{b}x + y + \frac{c}{b}\right)$. We have $x' = x$ and $y' = \frac{a}{b}x + y + \frac{c}{b}$, so $x = x'$ and $y = -\frac{a}{b}x' + y' - \frac{c}{b}$. Substituting these into the conic gives us

$$ax'^2 + bx' \left(\frac{a}{b}x' + y' + \frac{c}{b}\right) + c \left(\frac{a}{b}x' + y' + \frac{c}{b}\right)^2 + dx' + e \left(\frac{a}{b}x' + y' + \frac{c}{b}\right) + f = 0$$

Expanding and simplifying yields

$$\begin{aligned} \left(2a + \frac{a^2c}{b^2}\right) x'^2 + \left(b + \frac{2ac}{b}\right) x'y' + cy'^2 + \left(c + \frac{2ac^2}{b^2} + d + \frac{ae}{b}\right) x' + \\ \left(\frac{2c^2}{b} + e\right) y' + \frac{c^3}{b^2} + \frac{ce}{b} + f = 0. \end{aligned}$$

Since $a, b, c, d, e, f \in \mathbb{Q}$, this transformed conic is also a rational conic. \square

Proof of proposition 3.3. By applying the transformation given in lemma 3.4 to both the line and the conic, we can assume without loss of generality that the line is $y = 0$ and that the conic is $ax^2 + bxy + cy^2 + dx + ey + f = 0$. We can write the conic in this form because of lemma 3.5.

Let $\mathcal{O} = (x_1, y_1)$ be a fixed point and $P = (x_2, y_2)$ be another point on the transformed conic (see ??). The line through these two points is given by $y - y_1 = \frac{y_2 - y_1}{x_2 - x_1}(x - x_1)$. Given a point P , we can find a corresponding point Q on the line $y = 0$ by taking the intersection of the two lines. Some basic algebra yields $Q = \left(\frac{x_1y_2 - x_2y_1}{y_2 - y_1}, 0\right)$.

Given the point Q , we can also find a corresponding point P on the conic. The line through \mathcal{O} and Q has equation

$$\begin{aligned} y - y_1 &= \frac{0 - y_1}{\frac{x_1y_2 - x_2y_1}{y_2 - y_1} - x_1}(x - x_1) \\ y - y_1 &= \frac{y_2 - y_1}{x_2 - x_1}(x - x_1) \end{aligned}$$

The point $P = (x_2, y_2)$ is easily seen to satisfy that equation. The only point that this method will not work for is when $P = \mathcal{O}$. Therefore, there is a one-to-one correspondence between the points on a rational line and a rational conic for all but one point. \square

We know that there are an infinite number of rational points on a rational line by proposition 3.2. Therefore, we know that there are an infinite number of rational points on a rational conic. We have a nice way to generate all of these rational points too, simply by fixing a point on the conic and choosing a rational point on the line.

One might wonder why we needed to distinguish between singular and nonsingular curves. It turns out that the two types of curves behave very differently. For example, determining all the rational points on the singular curve is relatively easy.

Proposition 3.6. *Let E be a singular Weistrass curve. If E has a double root, then there is a one-to-one correspondence between every point in E and a rational line except for the singular point (see figure 6).*

Proof. Let \mathcal{O} be at the double root. Then any line will only intersect the singular Weistrass curve at one other point by Bézout's theorem. By reasoning similar to that of proposition 3.3, we see that every points on the curve has a corresponding point on the line. \square

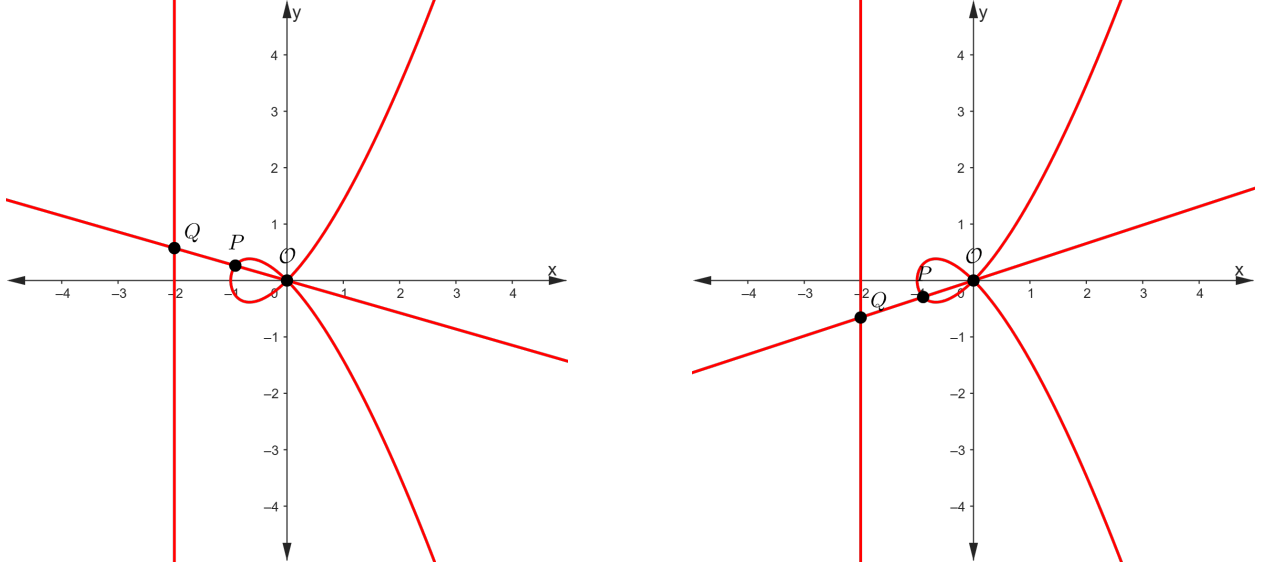


FIGURE 6. Line-Singular Weierstrass Curve Correspondence

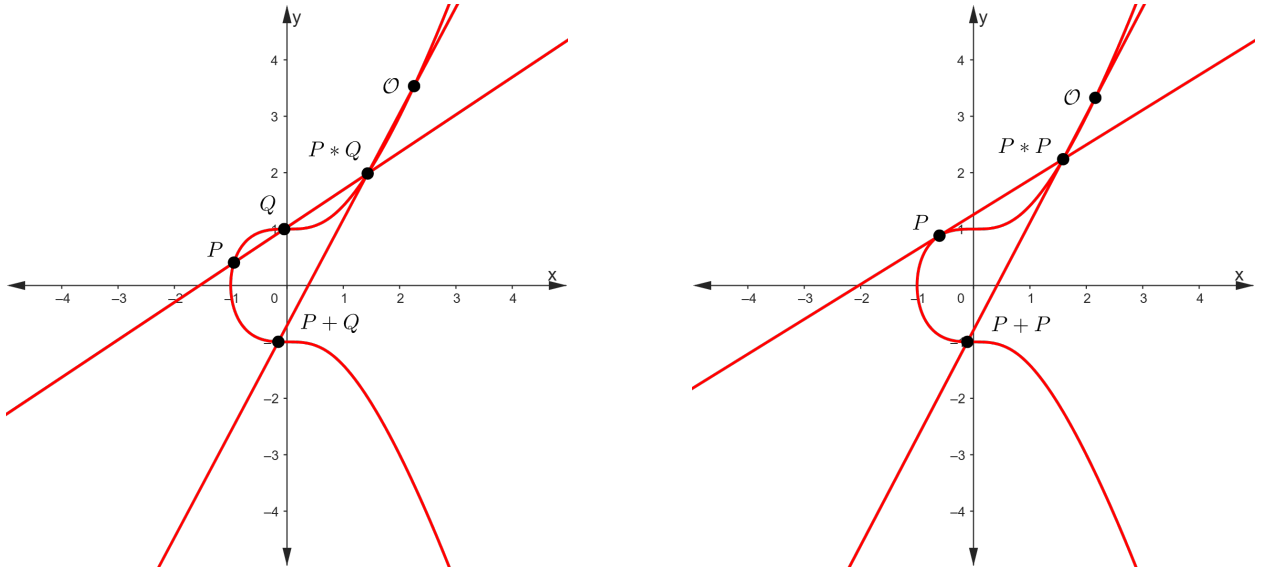


FIGURE 7. Group Law

4. THE GROUP LAW

Let \mathcal{O} , P , and Q be points on an elliptic curve E such that \mathcal{O} is fixed. We add points in the following way. Assuming that none of \mathcal{O} , P , and Q are the same point. First, we construct the line through points P and Q . By Bézout's theorem, this line will intersect E at exactly one other point, which we will call $P * Q$. Next, we construct the line through \mathcal{O} and $P * Q$, which will also intersect E in exactly one other point. We will call this point $P + Q$. If any of the points are the same, we use the tangent line instead (see figure 7).

While we could technically let \mathcal{O} be any point on the elliptic curve, it is often simplest to let \mathcal{O} be the point at infinity. This turns the last step of our addition into simply reflecting $P * Q$ over the x -axis (see figure 8 and lemma 4.3).

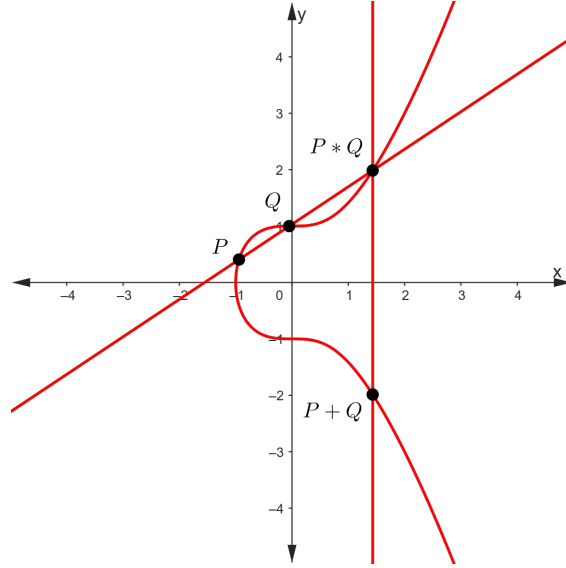


FIGURE 8. Group Law With Identity at Infinity

Let $P = (x_1, y_1)$ and $Q = (x_2, y_2)$ be points on an elliptic curve E written in short Weierstrass form, and let $P * Q = (x_3, -y_3)$ and $P + Q = (x_3, y_3)$. We wish to find a formula for (x_3, y_3) . The line through P and Q can be written as

$$y - y_1 = m(x - x_1)$$

or

$$y - y_2 = m(x - x_2),$$

where $m = \frac{y_2 - y_1}{x_2 - x_1}$. Solving for y gives us

$$y = mx - mx_1 + y_1$$

and

$$y = mx - mx_2 + y_2.$$

Let $c = -mx_1 + y_1 = -mx_2 + y_2$, so that both of our equations become $y = mx + c$. We want to find the intersection of $y = mx + c$ and $y^2 = x^3 + ax + b$. Squaring $y = mx + c$ gives us

$$y^2 = m^2x^2 + 2cmx + c^2,$$

so we can write

$$m^2x^2 + 2cmx + c^2 = x^3 + ax + b,$$

or

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

The roots of this cubic are precisely x_1 , x_2 , and x_3 . This means that we can write it as

$$x^3 - m^2x^2 + (a - 2cm)x + b - c^2 = (x - x_1)(x - x_2)(x - x_3).$$

The right side of the equation becomes $x^3 - (x_1 + x_2 + x_3)x^2 + (x_1x_2 + x_1x_3 + x_2x_3)x - x_1x_2x_3$. Equating the x^2 terms gives us

$$m^2 = x_1 + x_2 + x_3,$$

so

$$x_3 = m^2 - x_1 - x_2.$$

We know that $(x_3, -y_3)$ lies on the line $y = mx + c$. Substituting x_3 into $y = mx + c$ gives us

$$-y_3 = mx_3 + c,$$

so

$$y_3 = -mx_3 - c.$$

Proposition 4.1 (Group law). *The set of points on an elliptic curve under the defined addition form an Abelian group.*

Lemma 4.2. *Let $y^2 = x^3 + ax + b$ be an elliptic curve. Then the point at infinity \mathcal{O} has homogeneous coordinates $(0 : 1 : 0)$.*

Proof. Homogenizing $y^2 = x^3 + ax + b$ gives us

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

Let $(X : Y : Z)$ be a point on the homogeneous polynomial. We get the point at infinity when $Z = 0$. At $Z = 0$, the equation becomes

$$X^3 = 0,$$

so

$$X = 0.$$

Therefore, our point at infinity is $(0 : Y : 0) = (0 : 1 : 0)$. □

Lemma 4.3. *Let E be an elliptic curve in short Weierstrass form. Let \mathcal{O} be the fixed point at infinity and P and Q two points on an elliptic curve. The line through \mathcal{O} and $P = (x, y)$ intersects E a third time at $Q = (x, -y)$. If $P = \mathcal{O}$, the line through \mathcal{O} and P intersect E a third time at \mathcal{O} .*

Proof. Let $P = (x : y : 1)$, and let the line going through \mathcal{O} and P be $aX + bY + cZ = 0$. By lemma 4.2, we know that the point at infinity is $(0 : 1 : 0)$. Substituting this into our equation gives us

$$b = 0.$$

We know that the line also goes through $(x : -y : 1)$, so

$$ax - by + c = 0.$$

Since $b = 0$,

$$ax + c = 0,$$

so

$$c = -ax.$$

Substituting this into $aX + bY + cZ = 0$ tells us that

$$aX - axZ = 0$$

or

$$X = xZ.$$

We want to find the intersection of this line with $Y^2Z = X^3 + aXZ^2 + bZ^3$. Substituting $X = xZ$ yields

$$Y^2Z = (xZ)^3 + a(xZ)Z^2 + bZ^3 = x^3Z^3 + axZ^3 + bZ^3 = (x^3 + ax + b)Z^3,$$

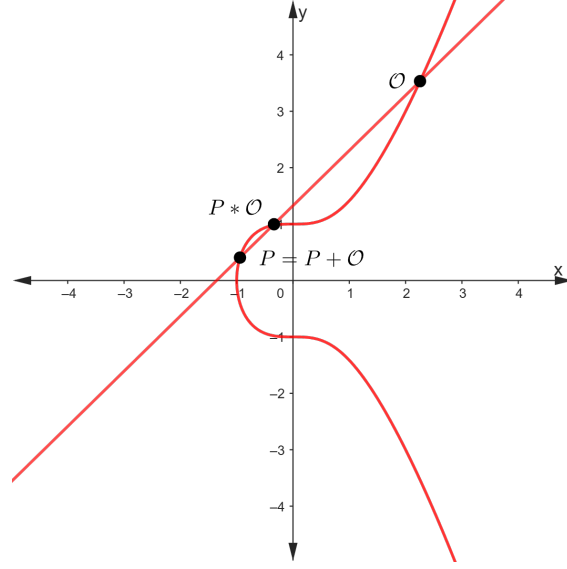


FIGURE 9. Group Law Identity

so

$$Y^2 = (x^3 + ax + b)Z^2$$

if $Z \neq 0$. Solving for Y gives us

$$Y = \pm \sqrt{x^3 + ax + b}Z$$

so the two points of intersection are $(xZ : \pm \sqrt{x^3 + ax + b}Z : Z) = (x : \pm y : 1)$. If $Z = 0$, we get the point at infinity.

If $P \neq \mathcal{O}$, the line going through P and \mathcal{O} intersects E at $(x : y : 1)$, since we already let $P = (x : -y : 1)$. This corresponds to the affine point (x, y) . If $P = \mathcal{O} = (0 : 1 : 0)$, the intersection happens at $(0 : -1 : 0) = (0 : 1 : 0) = \mathcal{O}$. \square

First, we verify that \mathcal{O} acts as the identity element (see figure 9).

Lemma 4.4 (Identity). *Let P be a point on an elliptic curve and \mathcal{O} the fixed point at infinity. Then $\mathcal{O} + P = P + \mathcal{O} = P$.*

Proof. Assume that $P \neq \mathcal{O}$. Let $P = (x, y)$. By lemma 4.3, we have $\mathcal{O} * P = (x, -y)$. Applying the lemma again gives us $\mathcal{O} + P = (x, y) = P$. By the same reasoning, $P * \mathcal{O} = (x, -y)$ and $P + \mathcal{O} = (x, y) = P$. Therefore, $\mathcal{O} + P = P + \mathcal{O} = P$.

Now assume that $P = \mathcal{O}$. Lemma 4.3 tells us that $\mathcal{O} + \mathcal{O} = \mathcal{O}$, $\mathcal{O} + P = P + \mathcal{O} = P$ still holds. \square

We construct the inverse geometrically by first constructing the tangent line at \mathcal{O} . This will intersect the elliptic curve at a third point we will call R . Next, we construct the line going through P and R . This second line will intersect the curve at another point which we will denote $-P$ (see figure 10).

Lemma 4.5 (Inverse). *Let P be a point on an elliptic curve and \mathcal{O} the fixed point at infinity. Then there exists a $-P$ such that $P + (-P) = (-P) + P = \mathcal{O}$.*

Proof. Since \mathcal{O} is the point at infinity and the short Weierstrass equation is symmetric about the x -axis, if $P = (x : y : 1)$, $-P = (x : -y : 1)$. From lemma 4.3, we get the equation

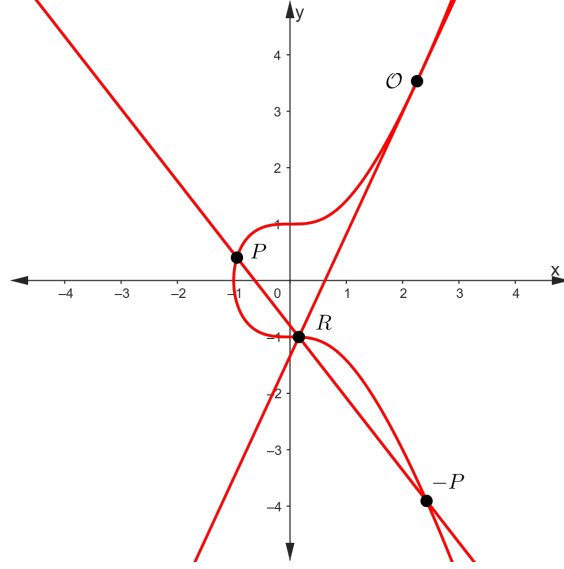


FIGURE 10. Group Law Inverse

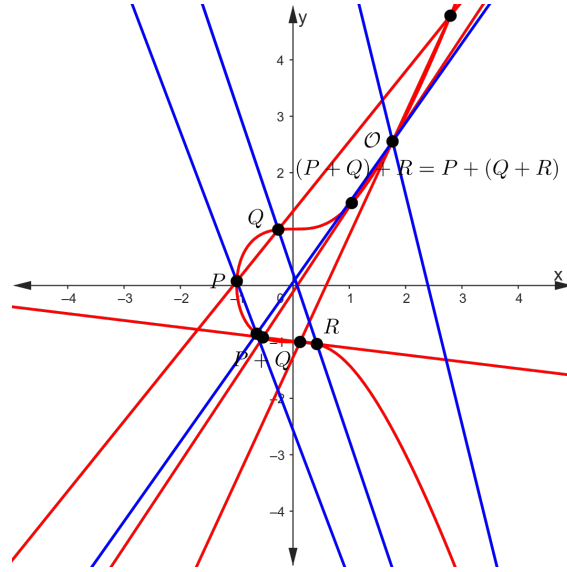


FIGURE 11. Group Law Associativity

$Y = \pm yZ$. We see that the point at infinity, $(0 : 1 : 0)$, satisfies this equation. Bézout's theorem tells us that there are exactly three intersection points. Since we already have the points $(x : y : 1)$ and $(x : -y : 1)$, the only other intersection point is $(0 : 1 : 0)$. This tells us that $P + (-P) = \mathcal{O}$. By similar reasoning, $(-P) + P = \mathcal{O}$, so $P + (-P) = (-P) + P = \mathcal{O}$. \square

Now we prove associativity (see figure 11).

Lemma 4.6 (Associativity). *Let P , Q , and R be points on an elliptic curve and \mathcal{O} the fixed point at infinity. Then $(P + Q) + R = P + (Q + R)$.*

A proof can be found in [2].

Now we show commutativity, making the group Abelian (see figure 12).

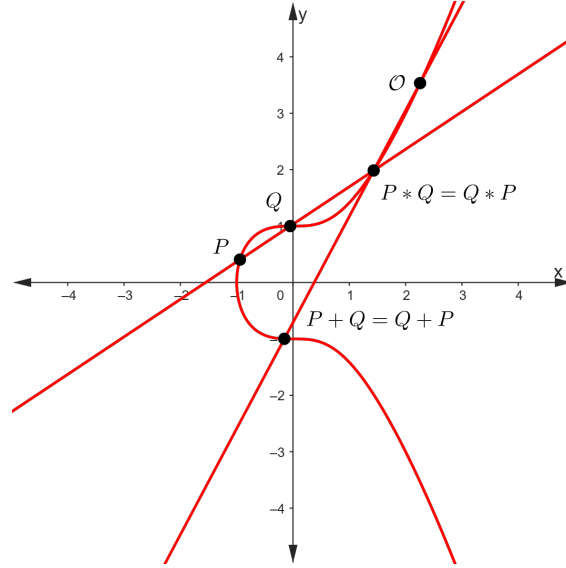


FIGURE 12. Group Law Commutativity

Lemma 4.7 (Commutativity). *Let P and Q be points on an elliptic curve and \mathcal{O} the fixed point at infinity. Then $P + Q = Q + P$.*

Proof. The line through P and Q is the same as the line through Q and P . Both lines give us the same $P * Q$ and $P + Q$, so the order of the points does not matter. Therefore, $P + Q = Q + P$. \square

Proof of proposition 4.1. First, we show that adding points is a binary operation. Obviously, the operation takes two points and gives us back one point, so all we need to show is that we always get another point on the curve. Let \mathcal{O} be a fixed point and P and Q be two points on the elliptic curve. By Bézout's theorem, the line going through P and Q intersects the curve at exactly one other point. By the same reasoning, the line going through \mathcal{O} and this new point must intersect the curve at exactly one other point. By definition, our addition gives us this point, so it is a binary operation.

Let P be a point on the elliptic curve and \mathcal{O} the fixed point at infinity. By lemma 4.4, $\mathcal{O} + P = P + \mathcal{O} = P$. By lemma 4.5, there exists an inverse $-P$ for every point P such that $P + (-P) = (-P) + P = \mathcal{O}$. By lemma 4.6, $(P + Q) + R = P + (Q + R)$. By lemma 4.7, $P + Q = Q + P$, so adding points on an elliptic curve forms an Abelian group. \square

5. HASSE'S THEOREM AND GENERALIZATIONS

The following proof has been adapted from [2].

Theorem 5.1 (Hasse's theorem). *Let E be an elliptic curve and \mathbb{F}_q a finite field of order $q = p^n$ for some prime p and positive integer n . Then*

$$|\#E(\mathbb{F}_q) - (q + 1)| \leq 2\sqrt{q}.$$

Hasse's theorem bounds the number of possible \mathbb{F}_q -rational points. It tells us that the expected number of such points is close to $q + 1$. The plus is because we are including the point at infinity.

Lemma 5.2. *In \mathbb{F}_q , $(x + y)^q = x^q + y^q$.*

Proof. We prove this by induction. Consider the field \mathbb{F}_p . By the binomial theorem,

$$(x + y)^p = \sum_{i=0}^p \binom{p}{i} x^{p-i} y^i.$$

We know that $\binom{p}{i} = \frac{p!}{i!(p-i)!}$. Since p is prime, it follows that $p \nmid i!(p-i)!$ for $1 \leq i \leq p-1$.

Therefore, $\frac{\binom{p}{i}}{p} \in \mathbb{Z}$, so $\binom{p}{i} = p \cdot \frac{\binom{p}{i}}{p} \equiv 0 \pmod{p}$. Then

$$\sum_{i=0}^p \binom{p}{i} x^i y^{p-i} \equiv \binom{p}{0} x^p y^0 + \binom{p}{p} x^0 y^p = x^p + y^p \pmod{p},$$

so $(x + y)^p = x^p + y^p$ in \mathbb{F}_p .

Now we generalize to \mathbb{F}_q . We have $(x + y)^p = x^p + y^p$ in \mathbb{F}_p as our base case. We need to show that if $(x + y)^{p^k} = x^{p^k} + y^{p^k}$, $(x + y)^{p^{k+1}} = x^{p^{k+1}} + y^{p^{k+1}}$. We have

$$\begin{aligned} \left((x + y)^{p^k}\right)^p &= \left(x^{p^k} + y^{p^k}\right)^p \\ (x + y)^{p^{k+1}} &= \left(x^{p^k}\right)^p + \left(y^{p^k}\right)^p \\ (x + y)^{p^{k+1}} &= x^{p^{k+1}} + y^{p^{k+1}}, \end{aligned}$$

since $x^{p^k}, y^{p^k} \in \mathbb{F}_q$, so our inductive step is complete. Therefore, $(x + y)^q = x^q + y^q$ in \mathbb{F}_q . \square

Lemma 5.3. *We have $a \in \mathbb{F}_q$ if and only if $a^q = a$.*

Proof. First we show that if $a^q = a$, $a \in \mathbb{F}_q$. Consider the polynomial $f(t) = t^q - t$. Since we are assuming that $a^q = a$, $a^q - a = 0$, so $f(a) = 0$ for every $a \in \mathbb{F}_q$. Because $\deg f(t) = q$, we know that $f(t)$ has at most q roots in \mathbb{F}_q . Since each of the q elements of \mathbb{F}_q is a root of $f(t)$ and $f(t)$ has at most q roots, the roots of $f(t)$ must be the elements of \mathbb{F}_q . This tells us $a^q = a$ implies that $a \in \mathbb{F}_q$. Therefore, $a \in \mathbb{F}_q$ if $a^q = a$.

Now we prove that if $a \in \mathbb{F}_q$, $a^q = a$. We know that $0^q = 0$, so assume that $a \neq 0$. Consider the multiplicative group $\mathbb{F}_q^\times = \mathbb{F}_q \setminus \{0\}$. By Lagrange's theorem, $\text{ord } a \mid \#(\mathbb{F}_q^\times)$, so $\text{ord}(a) \cdot k = q - 1$ for some positive $k \in \mathbb{Z}$. By definition, $a^{\text{ord } a} = 1$, so

$$\begin{aligned} (a^{\text{ord } a})^k &= 1^k \\ a^{q-1} &= 1. \end{aligned}$$

Multiplying through by a gives us $a^q = a$, so $a^q = a$ for all $a \in \mathbb{F}_q$. Therefore, $a \in \mathbb{F}_q$ if and only if $a^q = a$. \square

Lemma 5.4. *Let $\phi_q: E(\overline{\mathbb{F}_q}) \rightarrow E(\overline{\mathbb{F}_q})$ be the Frobenius endomorphism defined by $\phi_q(x, y) = (x^q, y^q)$ and $\phi_q(\mathcal{O}) = \mathcal{O}$. Then $E(\mathbb{F}_q) = \ker(\phi_q - 1)$.*

Proof. Let $P \in E(\mathbb{F}_q)$. By lemma 5.3, $\phi_q(P) = P$. Then $(\phi_q - 1)(P) = 0$, so $P \in \ker(\phi_q - 1)$. Since this holds for all P , we have

$$E(\mathbb{F}_q) \subseteq \ker(\phi_q - 1).$$

Now suppose that $P \in \ker(\phi_q - 1)$. By definition, we have $(\phi_q - 1)(P) = 0$, or $\phi_q(P) = P$. lemma 5.3 tells us that $P \in \mathbb{F}_q$. The endomorphism ϕ_q is only defined for points that satisfy

the elliptic curve equation. We know that the endomorphism -1 is also on the curve. Since $\phi_q - 1$ uses the group law, it maps points onto the elliptic curve. Then by definition,

$$\ker(\phi_q - 1) \subseteq E(\mathbb{F}_q).$$

Since $E(\mathbb{F}_q) \subseteq \ker(\phi_q - 1)$ and $\ker(\phi_q - 1) \subseteq E(\mathbb{F}_q)$, we conclude that $E(\mathbb{F}_q) = \ker(\phi_q - 1)$. \square

Lemma 5.5. *Let $f(x, y)$ be a rational function on E . Then there exist rational functions $f_1(x)$ and $f_2(x)$ such that $f(x, y) = f_1(x) + f_2(x)y$.*

Proof. Consider a polynomial $p(x, y)$ on E . Since it must satisfy the equation $y^2 = x^3 + ax + b$, we can replace any even power of y by a polynomial in x and any odd power by y times a polynomial in x . Therefore, we can write $p(x, y) = p_1(x) + yp_2(x)$ for polynomials p_1 and p_2 .

Let $f(x, y) = \frac{g(x, y)}{h(x, y)} = \frac{g_1(x) + g_2(x)y}{h_1(x) + h_2(x)y}$. We have

$$\begin{aligned} \frac{g_1 + g_2y}{h_1 + h_2y} \cdot \frac{h_1 - h_2y}{h_1 - h_2y} &= \frac{g_1h_1 - g_1h_2y + g_2h_1y - g_2h_2y^2}{h_1^2 - h_2^2y^2} \\ &= \frac{g_1h_1 - g_2h_2(x^3 + ax + b) + y(g_1h_2 + g_2h_1)}{h_1^2 - h_2^2(x^3 + ax + b)} \\ &= \frac{g_1h_1 - g_2h_2(x^3 + ax + b)}{h_1^2 - h_2^2(x^3 + ax + b)} + \frac{(g_1h_2 + g_2h_1)}{h_1^2 - (x^3 + ax + b)h_2^2}y \end{aligned}$$

We see that both $\frac{g_1h_1 - g_2h_2(x^3 + ax + b)}{h_1^2 - h_2^2(x^3 + ax + b)}$ and $\frac{(g_1h_2 + g_2h_1)}{h_1^2 - h_2^2(x^3 + ax + b)}$ are rational functions in x , so we can let $f_1 = \frac{g_1h_1 - g_2h_2(x^3 + ax + b)}{h_1^2 - h_2^2(x^3 + ax + b)}$ and $f_2 = \frac{(g_1h_2 + g_2h_1)}{h_1^2 - h_2^2(x^3 + ax + b)}$ for rational functions f_1 and f_2 . This gives us $f(x, y) = f_1(x) + f_2(x)y$ for any rational function on E . \square

Lemma 5.6. *Let $\varphi \neq 0$ be an endomorphism. If φ is separable,*

$$\#\ker(\varphi) = \deg(\varphi),$$

and if φ is not separable,

$$\#\ker(\varphi) < \deg(\varphi).$$

Proof. Assume that φ is separable. By lemma 5.5, we can let $\varphi(x, y) = (\varphi_1(x), \varphi_2(x)y)$, where $\varphi_1(x) = \frac{p(x)}{q(x)}$ for polynomials $p(x)$ and $q(x)$. Since φ is separable, $\varphi'_1 \neq 0$ by proposition 2.10.

This tells us that $\left(\frac{p}{q}\right)' = \frac{p'q - pq'}{q^2} \neq 0$, so $p'q - pq' \neq 0$.

Let $S = (p'q - pq')(x)q(x) = 0$ for $x \in \overline{K}$ be the set of critical points of φ_1 . Let $(a, b) \in E(\overline{K})$ satisfy the following properties:

1. $a \neq 0, b \neq 0, (a, b) \neq \mathcal{O}$, where \mathcal{O} is the point at infinity
2. $\deg(p(x) - aq(x)) = \max(\deg(p), \deg(q)) = \deg(\varphi)$
3. $a \notin \varphi_1(S)$
4. $(a, b) \in \varphi(E(\overline{K}))$

We first need to show that such an (a, b) can exist. By proposition 2.22, \overline{K} has an infinite number of elements. Therefore, we can always choose a point (a, b) such that $a \neq 0$, $b \neq 0$, and $(a, b) \neq \mathcal{O}$.

Let $p(x) = p_m x^m + p_{m-1} x^{m-1} + \cdots + p_0$ and $q(x) = q_n x^n + q_{n-1} x^{n-1} + \cdots + q_0$. The only way we could have $\deg(p(x) - aq(x)) \neq \max(\deg(p), \deg(q))$ is if $m = n$ and $p_m = aq_n$. There is only one value of a such that this is the case, namely $a = \frac{p_m}{q_n}$. Since \overline{K} is infinite, we can always choose an $a \neq \frac{p_m}{q_n}$, so we can have $\deg(p(x) - aq(x)) = \max(\deg(p), \deg(q)) = \deg(\varphi)$.

We know that $p'q - pq'$ is not the zero polynomial, so S must be finite. There are an infinite number of $\varphi_1(x)$, since there are an infinite number of $x \in \overline{K}$. Therefore, we can always choose an $a \notin \varphi_1(S)$.

The last property holds because φ is an endomorphism, so it must map $E(\overline{K})$ to $E(\overline{K})$.

We claim that there are exactly $\deg(\alpha)$ points $(x_1, y_1) \in E(\overline{K})$ such that $\alpha(x_1, y_1) = (a, b)$. For such an (x_1, y_1) , we have

$$\varphi(x_1, y_1) = (\varphi_1(x_1), \varphi_2(x_1)y_1) = \left(\frac{p(x_1)}{q(x_1)}, \varphi_2(x_1)y_1 \right) \in \varphi(E(\overline{K})).$$

Since $(a, b) \in \varphi(E(\overline{K}))$, we get

$$\begin{aligned} a &= \frac{p(x_1)}{q(x_1)} \\ b &= y_1 \varphi_2(x_1). \end{aligned}$$

Because $(a, b) \neq \mathcal{O}$, $q(x_1) \neq 0$. Since $b \neq 0$, $y_1 = \frac{b}{\varphi_2(x_1)}$. This tells us that y_1 depends entirely on x_1 , so counting the number of x_1 s that satisfy the constraints is sufficient.

We assumed that $\deg(p(x) - aq(x)) = \deg(\varphi)$, counting multiplicities. We will show that the roots of $p - aq$ must all be distinct. Suppose that x_0 is a multiple root. Then

$$p(x_0) - aq(x_0) = 0$$

and

$$p'(x_0) - aq'(x_0) = 0.$$

Multiplying the equations $aq(x_0) = p(x_0)$ and $p'(x_0) = aq'(x_0)$ gives us

$$ap'(x_0)q(x_0) = ap(x_0)q'(x_0),$$

which implies that

$$p'(x_0)q(x_0) - p(x_0)q'(x_0) = 0,$$

since $a \neq 0$. However, this tells us that $x_0 \in S$, so $a = \varphi_1(x_0) \in \varphi_1(S)$, contrary to our assumption that $a \notin \varphi_1(S)$. Therefore, $p - aq$ has no multiple roots and thus has $\deg(\varphi)$ roots.

We have shown that there are exactly $\deg(\alpha)$ points (x_1, y_1) such that $\varphi(x_1, y_1) = (a, b)$ for some point (a, b) . This tells us that $\ker(\varphi)$ also has $\deg(\varphi)$ distinct points. Therefore, $\#\ker(\varphi) = \deg(\varphi)$ if φ is separable.

If φ is not separable, $p' - aq'$ is the zero polynomial, so $p(x) - aq(x)$ has multiple roots and therefore has less than $\deg(\alpha)$ solutions. Therefore, $\#\ker(\varphi) < \deg(\varphi)$ if φ is not separable. \square

Lemma 5.7. *We have $\#E(\mathbb{F}_q) = \# \ker(\varphi_q - 1) = \deg(\varphi_q - 1)$.*

Proof. By lemma 5.4, $E(\mathbb{F}_q) = \ker(\varphi_q - 1)$, so

$$\#E(\mathbb{F}_q) = \# \ker(\varphi_q - 1).$$

Since $q = 0$ in \mathbb{F}_q , we have $(\varphi_q - 1)'(x) = (x^q - x)' = qx^{q-1} - 1 = -1$. Since this can never be 0, $\varphi_q - 1$ is separable. By lemma 5.6, $\# \ker(\varphi_q - 1) = \deg(\varphi_q - 1)$. Therefore, $\#E(\mathbb{F}_q) = \# \ker(\varphi_q - 1) = \deg(\varphi_q - 1)$. \square

Lemma 5.8. *We have $\deg(a\alpha + b\beta) = a^2 \deg \alpha + b^2 \deg \beta + ab(\deg(\alpha + \beta) - \deg \alpha - \deg \beta)$.*

Lemma 5.9. *Let $a = \#E(\mathbb{F}_q) - (q + 1)$, and let $r, s \in \mathbb{Z}$ such that $\gcd(s, q) = 1$. Then $\deg(r\varphi_q - s) = r^2q + s^2 - rsa$.*

Proof. By lemma 5.8,

$$\deg(r\varphi_q - s) = r^2 \deg(\varphi_q) + s^2 \deg(-1) + rs(\deg(\varphi_q - 1) - \deg(\varphi_q) - \deg(-1)).$$

Since $\deg(\varphi_q) = q$ and $\deg(-1) = 1$, we have

$$\deg(r\varphi_q - s) = r^2q + s^2 + rs(\deg(\varphi_q - 1) - (q + 1)).$$

By lemma 5.7, $\#E(\mathbb{F}_q) = \deg(\varphi_q - 1)$, so and $a = \#E(\mathbb{F}_q) - (q + 1) = \deg(\varphi_q - 1) - (q + 1)$. This gives us

$$\deg(r\varphi_q - s) = r^2q + s^2 + rsa. \quad \square$$

Proof of theorem 5.1. By lemma 5.8, we know that $\deg(r\varphi_q - s) = r^2q + s^2 + rsa$, where $a = \#E(\mathbb{F}_q) - (q + 1)$. Since $\deg(r\varphi_q - s) \geq 0$,

$$r^2q + s^2 + rsa \geq 0.$$

Dividing through by s^2 gives us

$$q \left(\frac{r}{s}\right)^2 + a \left(\frac{r}{s}\right) + 1 \geq 0.$$

Since the rationals are dense in \mathbb{R} , we can replace $\frac{r}{s}$ with x . This gives us

$$qx^2 + ax + 1 \geq 0.$$

This quadratic must either have 1 real solution or 2 imaginary solutions, so the discriminant

$$a^2 - 4q \leq 0.$$

This gives us

$$|a| \leq 2\sqrt{q},$$

so

$$|\#(E(\mathbb{F}_q)) - (q + 1)| \leq 2\sqrt{q}. \quad \square$$

The following generalization of theorem 5.1 was can be found in [1].

Theorem 5.10 (Hasse-Weil bound). *Let C be an algebraic curve of genus g and \mathbb{F}_q a finite field of order $q = p^n$ for some prime p and positive integer n . Then*

$$|\#C(\mathbb{F}_q) - (q + 1)| \leq 2g\sqrt{q}.$$

ACKNOWLEDGMENTS

The author would like to thank Simon Rubinstein-Salzedo and Rajiv Nelakanti for their guidance on this paper.

REFERENCES

- [1] Joseph H. Silverman, *Rational Points on Elliptic Curves*, 2nd Edition, Springer, 2009.
- [2] Lawrence C. Washington, *Elliptic Curves: Number Theory and Cryptography*, 2nd Edition, Chapman and Hall/CRC, 2008.