ELLIPTIC CURVES AND CRYPTOGRAPHY

JOHN KOPPER

Abstract. This paper begins by discussing the foundations of the study of elliptic curves and how a the points on an elliptic curve form an additive group. We then explore some of the interesting features of elliptic curves, including the fact that elliptic curves are complex tori. At the end we briefly discuss how elliptic curves can be used in cryptography.

Contents

1. Introduction

The study of elliptic curves can be approached from many perspectives. In this first section, we describe them as solutions to a type of polynomial. Using the geometric properties implied by considering these curves over $\mathbb{R}$, we define a group law. The second section is devoted to showing that there is an isomorphism between complex tori and elliptic curves. In the third section we deviate from the discussion of elliptic curves into one of field theory so that in the final section we can discuss cryptographic applications of elliptic curves.

Definition 1.1. An *elliptic curve E over a field K* is a set of points (x, y) in $K \times K$ satisfying

$$y^2 + Ay = x^3 + Bx^2 + Cx + D \tag{1.2}$$

This is an unwieldy and seldom-used way of writing an elliptic curve. The following proposition characterizes almost all elliptic curves of interest in a simpler form.

Proposition 1.3. *If K is a field with characteristic neither 2 nor 3, then (1.2) can be written as*

$$y^2 = x^3 + bx + c \tag{1.4}$$

Date: August 16, 2010.

Proof. Because the characteristic of K is not 2, Equation 1.2 gives

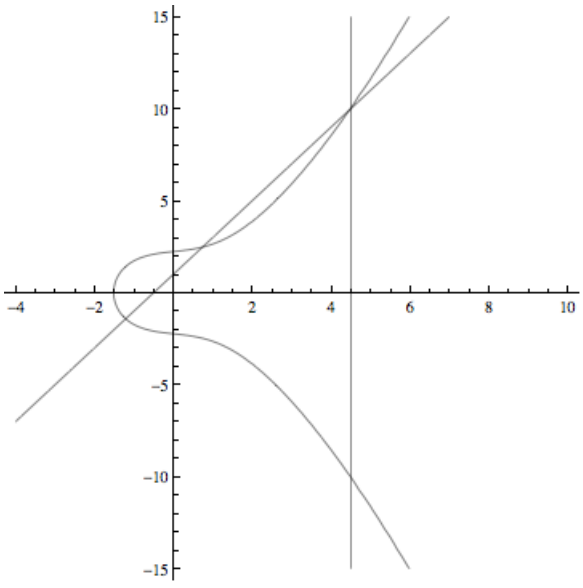
$$\begin{aligned}\left(y + \frac{A}{2}\right)^2 &= x^3 + Bx^2 + Cx + D + \left(\frac{A^2}{4}\right) \\ &= x^3 + Bx^2 + Cx + D'\end{aligned}$$

Let $y' = y + A/2$, and do a similar trick with $x' = x + B/3$ to obtain Equation 1.4. □

An interesting property of elliptic curves, and one of the focuses of this paper, is that an addition can be defined such that the set of points on the curve form an additive group. The procedure for adding two points is purely geometric.

Definition 1.5. Given two distinct points $P_1 = (x_1, y_1)$, $P_2 = (x_2, y_2)$ on an elliptic curve, construct the unique line L through P_1 and P_2. The line L intersects the curve at a third point, $P_3' = (x_3', y_3')$. Define the sum $P_1 + P_2$ to be the reflection of P_3 over the x-axis, $(x_3', -y_3')$.

From the definition alone, it is not clear that this addition is valid, much less useful. In fact it is very nearly valid, but to make it fully so, we need to slightly amend our definition of elliptic curves. To each elliptic curve we add a point "∞," which cannot be given by Cartesian coordinates. The notion of ∞ can be made rigorous with a discussion of projective space, but it is outside the scope of this paper to do so. For our purposes, it suffices to say that the point ∞ lies on every elliptic curve, and every vertical line in the plane passes through it.



Now if two points on an elliptic curve lie also on the same vertical line, their sum is ∞. Note also that $\infty + P = P$, because the line through ∞ and P is the vertical line through P, which intersects the curve a third time at precisely P's reflection over the x-axis. The following definition completes the idea of addition on elliptic curves.

Definition 1.6. Let P be a point on an elliptic curve E. Construct the line L tangent to E at P. L intersects E at one other point, P'. Reflect P' over the x-axis to obtain the sum $P + P$.

Notation 1.7. If P is a point on an elliptic curve and n is a positive integer, then $nP = P + P + \cdots + P$ (with n terms in the sum). If n is a negative integer, then $nP = (-P) + (-P) + \cdots + (-P)$.

What remains unclear from the definitions is whether the lines L in Definitions 1.5 and 1.6 do indeed intersect the elliptic curve in the right number of places. This issue will be addressed in the next section. It can be proven algebraically, but the proof later in this paper is simpler. Also, we are now, in theory, able to prove the following theorem.

Theorem 1.8. *Let E be an elliptic curve. Then the set of points on E with addition defined as in Definitions 1.5 and 1.6 forms an abelian group.*

Proof. Let P_1 and P_2 be two points on E. Then the line through P_1 and P_2 is the same as the line through P_2 and P_1. Addition is thus commutative. Also, from the above remarks it is clear that the point ∞ is the identity. The point P' obtained by reflecting a point P over the x-axis has the property $P + P' = \infty$. Therefore every point has an inverse.

To finish the proof of the theorem we need to prove that addition is associative, but the full proof is long and unnecessary. A roundabout but elegant proof will be given in Section 2. □

Although this group law for elliptic curves has been constructed geometrically, it can be expressed and proven completely algebraically. The following theorem, adapted from Washington [2], is simply a reformulation of Theorem 1.8 by finding the equations for the necessary lines.

Theorem 1.9. *Let E be an elliptic curve over a field K defined by the equation $y^2 = x^3+bx+c$. Let $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$ be points on E, with $P_1, P_2 \neq \infty$. Define $P_1 + P_2 = P_3 = (x_3, y_3)$ by the following:*

(i) *If $x_1 \neq x_2$, then*

$$x_3 = m^2 - x_1 - x_2, \quad y_3 = m(x_1 - x_3) - y_1 \qquad \text{where } m = \frac{y_2 - y_1}{x_2 - x_1}.$$

(ii) *If $x_1 = x_2$ and $y_1 \neq y_2$, then $P_1 + P_2 = \infty$.*

(iii) *If $P_1 = P_2$ and $y_1 \neq 0$, then*

$$x_3 = m^2 - 2x_1, \quad y_3 = m(x_1 - x_3) - y_1 \qquad \text{where } m = \frac{3x_1^2 + b}{2y_1}.$$

(iv) *If $P_1 = P_2$ and $y_1 = 0$, then $P_1 + P_2 = \infty$.*

(v) *For any P on E, define $P + \infty = P$.*

The set of points on E with this addition forms an abelian group.

2. Elliptic Curves Over $\mathbb{C}$

In this section we discuss a wholly different characterization of elliptic curves. The following discussion eventually shows that elliptic curves over $\mathbb{C}$ correspond elegantly to complex tori. To understand and justify this statement, several preliminary theorems are required. For the proofs of theorems 2.1 and 2.4, see, for example, Walter Rudin's *Real And Complex Analysis* [1].

Theorem 2.1. (Liouville's Theorem) *Every bounded entire function is constant.*

Proposition 2.2. *Let f be a meromorphic function on an open set Ω. Then for every a in Ω there exists an integer m, a ball $B(a,r)$, and a holomorphic function g with $g(a) \neq 0$ such that for all z in $B(a,r)$,*

$$f(z) = (z-a)^m g(z)$$

Further, there exists a positive integer k and complex numbers c_n such that

$$f(z) = \sum_{n=-k}^{\infty} c_n (z-a)^n$$

Definition 2.3. Let c_n and m be as in the previous proposition. Then the *order of vanishing of f at a* is $v_a(f) = m$.

Theorem 2.4. (Cauchy Integral Theorem) *Let f be holomorphic on $B(0,r)$, and $\gamma(t) = re^{it}$ for $0 \leq t \leq 2\pi$. Then for all z in $B(0,r)$,*

$$f(z) = \int_\gamma \frac{f(w)}{w-z} dw$$

The following corollary is of great importance in the study of elliptic curves.

Corollary 2.5. *Let Ω be an open, connected set in $\mathbb{C}$, and let γ be a curve in Ω with interior G. Then* $\dfrac{1}{2\pi i}\displaystyle\int_\gamma \varphi(z)\frac{f'(z)}{f(z)}dz = \sum_{a\in G} v_a(f)\varphi(a)$

Proof. f has finitely many zeros and poles in G, and so for each pole or zero a there is a circle γ_a that contains that pole or zero but no others. Thus

$$\frac{1}{2\pi i}\int_\gamma \varphi(z)\frac{f'(z)}{f(z)}dz = \sum_{a\in G}\left(\frac{1}{2\pi i}\int_{\gamma_a}\varphi(z)\frac{f'(z)}{f(z)}dz\right)$$

f is meromorphic and so by Proposition 2.2, $f(z) = (z-a)^m g(z)$ where $m = v_a(f)$, g is holomorphic, and $g(a) \neq 0$.Then by calculation,

$$\varphi(z)\frac{f'(z)}{f(z)} = \varphi(z)\left(\frac{m}{z-a} + \frac{g'(z)}{g(z)}\right)$$

Thus,

$$\frac{1}{2\pi i}\int_\gamma \varphi(z)\frac{f'(z)}{f(z)}dz = \sum_{a\in G}\left(\frac{1}{2\pi i}\int_{\gamma_a}\varphi(z)\left[\frac{m}{z-a} + \frac{g'(z)}{g(z)}\right]dz\right) \tag{2.6}$$

We can choose γ_a arbitrarily small, and $g(a) \neq 0$, therefore $g'(z)/g(z)$ is holomorphic in γ_a. Thus the integral of $\varphi(z)\dfrac{g'(z)}{g(z)}$ is zero around γ_a. By the Cauchy Integral Theorem,

$$\int_{\gamma_a}\varphi(z)\frac{m}{z-a}dz = m\varphi(a) = v_a(f)\varphi(a)$$

Substituting this into Equation 2.6 finishes the proof. □

Definition 2.7. Let w_1 and w_2 be a basis for $\mathbb{C}$ over $\mathbb{R}$. A *lattice* over $\mathbb{C}$ is a set $\Lambda = w_1\mathbb{Z} \oplus w_2\mathbb{Z} = \{a_1w_1 + a_2w_2 | a_1, a_2 \in \mathbb{Z}\}$.

The symbol Λ will now denote such a lattice with basis w_1, w_2. The ultimate object of study in this section is the group $\mathbb{C}/\Lambda$, which associates every point in $\mathbb{C}$ with a point in the parallelogram spanned by w_1 and w_2. Because opposite sides of this parallelogram are associated, $\mathbb{C}/\Lambda$ is topologically a torus. The following defines a function that draws a correspondence between $\mathbb{C}/\Lambda$ and elliptic curves.

Definition 2.8. Define the *Weierstrass $\wp$-function* on $\mathbb{C}\backslash\Lambda$ by

$$\wp(z) = \frac{1}{z^2} + \sum_{w\in\Lambda\backslash\{0\}} \frac{1}{(w-z)^2} - \frac{1}{w^2}$$

Note that

$$\wp'(z) = \sum_{w\in\Lambda} \frac{1}{(z-w)^3}$$

Definition 2.9. A function f is Λ-*periodic* if for all w in Λ and all z in $\mathbb{C}\backslash\Lambda$, $f(z) = f(z+w)$.

Lemma 2.10. *If $f(z) = f(z+w_1) = f(z+w_2)$ then f is Λ-periodic.*

Proof. Let $w = nw_1 + mw_2$. Clearly the statement holds if $n = 1$ and $m = 0$ or vice versa. Now induct over m, n. i.e.,

$$\begin{aligned} f(z+nw_1+mw_2) &= f([z+(n-1)w_1+mw_2]+w_1) \\ &= f([z+(n-1)w_1+(m-1)w_2]+w_2) \\ &= f(z+(n-1)w_1+(m-1)w_2) = f(z) \end{aligned}$$

□

Lemma 2.11. *Let f be an even function and f' be Λ-periodic. Then f is Λ-periodic.*

Proof. Let w be in Λ and define $g(z) = f(z+w) - f(z)$. Then

$$g'(z) = f'(z+w) - f'(z) = 0$$

g is therefore constant. But f is even, so $g(-w/2) = f(w/2) - f(-w/2) = f(w/2) - f(w/2) = 0$. g is therefore identically zero. □

Theorem 2.12. *$\wp$ is Λ-periodic*

Proof. Let $\Lambda' = w_1 - \Lambda = \{w_1 - w | w \in \Lambda\}$. Clearly $\Lambda' = \Lambda$. But

$$\begin{aligned} \wp'(z+w_1) &= -2\sum_{w\in\Lambda} \frac{1}{(z+w_1-w)^3} \\ &= -2\sum_{w\in\Lambda'} \frac{1}{(z-w)^3} = \wp'(z) \end{aligned}$$

The same is true for w_2 by the same argument. Thus $\wp'(z)$ is Λ-periodic. $\wp(z)$ is even, and is therefore also Λ-periodic. □

Proposition 2.13. *Define the* Eisenstein series *by*

$$G_k(\Lambda) = \sum_{w\in\Lambda\backslash\{0\}} \frac{1}{w^k}$$

For all $k > 2$, the sum in G_k converges.

Lemma 2.14. *$G_k(\Lambda) = 0$ if k is odd.*

Proof. It is clear that $\Lambda = -\Lambda = \{-w | w \in \Lambda\}$, therefore $G_k(\Lambda) = G_k(-\Lambda)$. But if k is odd, $G_k(-\Lambda) = -G_k(\Lambda)$. Thus $G_k(\Lambda)$ must be zero. □

Lemma 2.15. *For all w in $\mathbb{C}$ and for all z with $0 < |z| < |w|$,*

$$\wp(z) = \frac{1}{z^2} + \sum_{n=1}^{\infty}(n+1)G_{n+2}(\Lambda)z^n$$

Note that by Lemma 2.14, the odd terms of the sum are zero.

Proof. By definition, $\wp(z) = \dfrac{1}{z^2} + \displaystyle\sum_{w\in\Lambda} \frac{1}{(z-w)^2} - \frac{1}{w^2}$. Expanding the sum,

$$\sum_{w\in\Lambda} \frac{1}{(z-w)^2} - \frac{1}{w^2} = \sum_{w\in\Lambda} \frac{1}{w^2}\left(\frac{1}{(1-\frac{z}{w})^2} - 1\right)$$

But $\dfrac{1}{(1-\frac{z}{w})^2} - 1 = \left(\displaystyle\sum_{n=0}^{\infty}\left(\frac{z}{w}\right)^n\right)^2 - 1 = \sum_{n=1}^{\infty}(n+1)\left(\frac{z}{w}\right)^n$

Thus,

$$\begin{aligned}\sum_{w\in\Lambda} \frac{1}{(z-w)^2} - \frac{1}{w^2} &= \sum_{w\in\Lambda} \frac{1}{w^2}\sum_{n=1}^{\infty}(n+1)\left(\frac{z}{w}\right)^n \\ &= \sum_{n=1}^{\infty}(n+1)z^n \sum_{w\in\Lambda}\frac{1}{w^{n+2}} \\ &= \sum_{n=1}^{\infty}(n+1)G_{n+2}(\Lambda)z^n\end{aligned}$$

The result follows. □

Theorem 2.16. *Let $b = -60G_4(\Lambda)$ and $c = -140G_6(\Lambda)$. Then*

$$(\wp'(z))^2 = 4(\wp(z))^3 + b\wp(z) + c$$

Proof. From the expansion in Lemma 2.15 we obtain the following equations.

$$\begin{aligned}\wp(z) &= z^{-2} + 3G_4(\Lambda)z^2 + 5G_6(\Lambda)z^4 + \cdots \\ (\wp(z))^3 &= z^{-6} + 9G_4(\Lambda)z^{-2} + 15G_6(\Lambda) + \cdots\end{aligned}$$

Differentiating $\wp(z)$ yields

$$\begin{aligned}\wp'(z) &= -2z^{-3} + 6G_4(\Lambda)z + 20G_6(\Lambda)z^3 + \cdots \\ (\wp'(z))^2 &= 4z^{-6} - 24G_4(\Lambda)z^{-2} - 80G_6(\Lambda) + O(z^2)\end{aligned}$$

We then see that

$$4(\wp(z))^3 + b\wp(z) + c = 4z^{-6} - 24z^{-2} - 80G_6(\Lambda) + O(z^2)$$

Let $\Delta = 4(\wp(z))^3 + b\wp(z) + c - (\wp'(z))^2$. Since Δ has order z^2, is holomorphic and Λ-periodic, it is bounded. By Liouville's theorem it is constant. Because $\Delta = O(z^2)$, we have $\Delta \to 0$ as $z \to 0$. Thus Δ must be identically zero. □

Proposition 2.17. *Let f be a Λ-periodic function and let P be the parallelogram spanned by w_1 and w_2. Then*

$$\frac{1}{2\pi i}\int_{\partial P}\frac{f'(z)}{f(z)} = 0$$

Proof. If f is Λ-periodic, then f' is too. Define the Λ-periodic function $g(z) = f'(z)/f(z)$.We can split the integral into four integrals along the sides of the parallelogram,

$$\int_{\partial P} g(z) = \int_0^{w_1} g(z)dz + \int_{w_1}^{w_1+w_2} g(z)dz + \int_{w_1+w_2}^{w_2} g(z)dz + \int_{w_2}^{0} g(z)dz$$

By the periodicity of g,

$$\int_{w_1}^{w_1+w_2} g(z)dz = \int_0^{w_2} g(z)dz = -\int_{w_2}^{0} g(z)dz,$$

and

$$\int_{w_1+w_2}^{w_2} g(z)dz = \int_{w_1}^{0} g(z)dz = -\int_0^{w_1} g(z).$$

Thus the four terms sum to zero. □

Corollary 2.18. *If f is a Λ-periodic function, then f has the same number of zeros as poles in P (counting multiplicity).*

Proof. This follows from Corollary 2.5 with $\varphi(z) = 1$. □

Proposition 2.19. *With b and c as defined in Theorem 2.16, let $p(x) = 4x^3+bx+c$. Define $w_3 = w_1 + w_2$ and $x_i = \wp(w_i/2)$. Then*

(i) $p(x_j) = 0$.
(ii) $x_j \neq x_i$ *unless* $i = j$.

Proof. (i) Since $\wp'(z)$ is odd, we have $\wp'(w_i/2) = \wp'(-w_i/2) = -\wp'(w_i/2)$, which implies $(\wp'(w_i/2))^2 = 0$. Then

$$4(\wp(w_i/2))^3 + b\wp(w_i/2) + c = 0 = p(\wp(w_i/2))$$

(ii) From Corollary 2.18, $\wp$ has as many poles as zeros, counting multiplicity. Define $f_w(z) = \wp(z) - w$ for w in $\mathbb{C}$. Then f_w has the same poles as $\wp$. In particular, f_w and $\wp$ both have poles of multiplicity 2 at 0. This implies f_w has two zeros, or one zero of multiplicity 2. Therefore $\wp(z)$ maps P surjectively onto $\mathbb{C}$, taking each value twice.

From (i) we have $\wp'(w_1/2) = 0$. Thus $f_{x_i}(w_i/2) = 0 = f'_{x_i}(w_i/2)$. Or, $f_{x_i}(z)$ has a zero of multiplicity at least 2 at $w_i/2$. Therefore $\wp(z)$ cannot take on the value x_i anywhere except at $w_i/2$, or else $f_{x_i}(z)$ would have too many zeros. □

Proposition 2.20. *Let f be a Λ-periodic function and let γ be the parallelogram spanned by w_1 and w_2. Then*

$$\frac{1}{2\pi i}\int_{\gamma} z\frac{f'(z)}{f(z)} \in \Lambda$$

Proof. This follows from an argument identical to the proof of Proposition 2.17. □

Corollary 2.21. $\sum_{a\in P} v_a(f)a \in \Lambda$

Proof. This follows from Corollary 2.5 with $\varphi(z) = z$. □

We are now in a position to reconsider the group law on elliptic curves. First, Theorem 2.16 ensures that there is a map $\mathcal{F} : \mathbb{C}/\Lambda \to \{(x, y)|y^2 = 4x^3 + bx + c\} \cup \{\infty\}$, where $z + \Lambda \mapsto (\wp(z), \wp'(z))$ and $0 + \Lambda \mapsto \infty$. It Is not difficult to show that $\mathcal{F}$ is bijective.

Suppose now two points $(\wp(z_1), \wp'(z_1))$, $(\wp(z_2), \wp'(z_2))$ are given. Then they lie on the curve $y^2 = 4x^3+bx+c$, and there exists a line between them $Ax+By+C = 0$. Let $F(z) = A\wp(z) + B\wp'(z) + C$. F is clearly periodic. We know that $\wp'(z)$ has a pole of multiplicity 3, so if $B \neq 0$, F has three zeros by Corollary 2.18. By construction, z_1 and z_2 are two of them.

To find the third zero of F we employ Corollary 2.21 interpreted in $\mathbb{C}/\Lambda$,

$$\sum_{a \in P} v_a(f)(a + \Lambda) = 0 + \Lambda \tag{2.22}$$

F is Λ-periodic and has a pole of multiplicity 3 at 0, so letting $f = F$ in Equation 2.22,

$$(z_1 + \Lambda) + (z_2 + \Lambda) + (z_3 + \Lambda) = 0 + \Lambda \tag{2.23}$$

where z_3 is the third zero of F. This is an equation in the group $\mathbb{C}/\Lambda$, and can therefore be rewritten as $(z_1 + z_2) + \Lambda = -z_3 + \Lambda$. In the case $B = 0$, we have $(z_1 + z_2) + \Lambda = 0 + \Lambda$, so we let $z_3 = 0$ and define $\mathcal{F}(0) = \infty$. In any case, if P_1, P_2 are two points on an elliptic curve, we can define their sum,

$$P_1 + P_2 = \mathcal{F}(-z_3 + \Lambda) \tag{2.24}$$

Note that if $z_3 \neq 0$, then $\mathcal{F}(-z_3 + \Lambda) = (\wp(z_3), -\wp(z_3))$, which is precisely the group law described in the first section. Equations 2.23 and 2.24 show that the function $\mathcal{F}$ is a group isomorphism.

This proves the group law for complex elliptic curves. However, Theorem 1.9 shows that the group law can be expressed in purely algebraic terms, independent of the field K. Suppose now that there existed a field K' for which the group law did not hold. Then the group law would not hold for any field. Thus by proving the group law for $\mathbb{C}$ we have proven it for every field.

3. Finite Fields

The aim of this section is to construct and prove the existence of finite fields of order $q = p^n$ where p is prime and n is a positive integer. The use of finite fields for elliptic curves is less elegant than the use of $\mathbb{C}$, but has many applications, some of which will be discussed in Section 4.

Definition 3.1. Let F be a field. The derivative map $D : F[x] \to F[x]$ is the map defined by $D(x) = 1$, $D(fg) = gD(f) + fD(g)$, $D(f + g) = D(f) + D(g)$, and $D(ax) = a$, for a in F and f, g in $F[x]$.

Proposition 3.2. *If f, g are polynomials in $F[x]$ and a, b are in F, then*

(i) $D(1) = D(a) = 0$
(ii) $D(x^n) = nx^{n-1}$
(iii) $D(af + bg) = aD(f) + bD(g)$

Definition 3.3. Let F be a field and L be a field containing F. An element a in L is *algebraic over* F if there exists an f in $F[x]$ such that $f(a) = 0$. L is an *algebraic extension* of F if every a in L is algebraic. F is *algebraically closed* if for every f in $F[x]$ there is an c in F such that $f(c) = 0$.

Theorem 3.4. *Every field F has an algebraic extension $\overline{F}$ such that $\overline{F}$ is algebraically closed. Furthermore, if $\overline{F}'$ is an algebraically closed algebraic extension of F, then $\overline{F}' = \overline{F}$. We can pick $\overline{F}$ so that every algebraic extension of F is contained in $\overline{F}$.*

Proposition 3.5. *Let K be a finite extension of a field F. Then K is algebraic over F.*

Proof. If K is a finite extension of F then K is an $n$dimensional vector space over F. Let a be an element of K. Then the set $\{1, a, a^2, \ldots, a^n\}$ has $n+1$ elements and therefore is a linearly dependent set over F. That is, there exist $\beta_0, \ldots, \beta_n$ in F not all zero such that

$$\beta_0 + \beta_1 a + \cdots + \beta_n a^n = 0$$

Then a is a zero of the polynomial $f(x) = \beta_0 + \beta_1 x + \cdots + \beta_n x^n$. □

Definition 3.6. If F is a field and $\overline{F}$ its algebraic closure, then a polynomial f in $F[x]$ of degree n is *separable* if it has n distinct roots over $\overline{F}$.

Proposition 3.7. *A polynomial f is separable if and only if* $\gcd(f, D(f)) = 1$.

Proof. Suppose f is separable. Then $f(x) = (x - \alpha_1) \cdots (x - \alpha_n)$ for n distinct α_i in $\overline{F}$. This imples

$$\begin{aligned} D(f(x)) &= (x - \alpha_2) \cdots (x - \alpha_n) + \cdots + (x - \alpha_1)(x - \alpha_3) \cdots (x - \alpha_n) + \cdots \\ &= \sum_{i=1}^{n} \prod_{j \neq i} (x - \alpha_j) \end{aligned}$$

Since the $x - \alpha_i$ are irreducible and F[x] is a unique factorization domain, we see that the gcd is 1

Suppose now that f is not separable. Then $f(x) = (x-\alpha_1)^{m_1}(x-\alpha_2)^{m_2} \cdots (x-\alpha_n)^{m_n}$ where the m_i are positive integers, not all 1. Without loss of generality, suppose $m_1 > 1$. Then

$$D(f(x)) = m_1(x - \alpha_1)^{m_1 - 1}(x - \alpha_2)^{m_2} \cdots (x - \alpha_n)^{m_n} + \cdots$$

Thus $(x - \alpha_1)$ divides both f and $D(f)$ so $\gcd(f, D(f)) \neq 1$. □

Definition 3.8. The symbol $\mathbb{F}_p$ denotes $\mathbb{Z}/p\mathbb{Z}$ for prime p.

Proposition 3.9. *$\mathbb{F}_p$ is a field.*

Proof. $\mathbb{F}_p$ is a finite commutative ring with no zero divisors if p is prime. □

Proposition 3.10. *Let $F = \mathbb{F}_p$ for some prime p and $q = p^n$ for some positive integer n. Then the polynomial f over F defined by $f(x) = x^q - x$ is separable.*

Proof. If $f(x) = x^q - x$ then $D(f) = qx^{q-1} - 1 = -1$. Proposition 3.7 implies the result. □

Proposition 3.11. *Let F be a field of characteristic p. Let $q = p^n$ for some positive integer n. Then for any a, b in F, $(a + b)^q = a^q + b^q$.*

Proof. By the binomial theorem,

$$(a+b)^q = a^q + \binom{q}{1}a^{q-1}b + \cdots \binom{q}{q-1}ab^{q-1} + b^q,$$

but p divides $\binom{q}{k}$ for $k = 1, \ldots, q-1$. □

The following theorem constructs a single field of order q.

Theorem 3.12. *There exists a field of order $q = p^n$ for every prime p and positive integer n.*

Proof. Let $F = \mathbb{F}_p$ and $\overline{F}$ be the algebraic closure of F. Define $\mathbb{F}_q$ to be the set $\{\alpha \in \overline{F} | \alpha^q = \alpha\}$. By Proposition 3.10, $\mathbb{F}_q$ has q elements.

(i) $\mathbb{F}_q$ is closed under multiplication since if $a, b \in \mathbb{F}_q$ then $(ab)^q = a^q b^q = ab$.
(ii) It is closed under addition since $(a+b)^q = a^q + b^q = a + b$ using Proposition 3.11
(iii) It contains multiplicative inverses since $(a^{-1})^q = (a^q)^{-1} = a^{-1}$.

The remaining field properties can be verified directly. □

Lemma 3.13. *Let F be a field with $q = p^n$ elements. Then for every a in F, $a^q = a$.*

Proof. Clearly $0^q = 0$. The multiplicative group $F^\times$ has $q-1$ elements. By Lagrange's theorem, for any a in $F^\times$, the order of the subgroup generated by a is k, where k divides $q-1$. That is, $a^k = 1$, and for some m, $a^{q-1} = a^{km} = (a^k)^m = 1$. Then $a^q = a$. □

Theorem 3.14. *Suppose F and F' are two fields with $q = p^n$ elements. Then $F = F'$.*

Proof. Let F and F' be two such fields. Both contain the field $\mathbb{F}_p$ and so both are contained within its algebraic closure $\overline{\mathbb{F}}_p$. Then the set $E = \{a \in \overline{\mathbb{F}}_p \mid a^q = a\}$ contains both F and F' by Lemma 3.13. But each set has q elements, so $F = E = F'$. □

4. Elliptic Curve Cryptography

Having established the existence of $\mathbb{F}_q$, we can now discuss elliptic curves over finite fields in the context of cryptography. A *cryptosystem* is a method of secretly exchanging information between two parties in such a way that the two parties can communicate easily and freely, while anyone eavesdropping would have a difficult or impossible time understanding anything.

A common type of cryptography is called *public key cryptography*. In public key cryptography, if Alice wants to send a message to Bob, she encrypts a message with his *public key*, which is a piece of information Bob makes available to everyone. Bob can then decrypt the message using his *private key*, which he keeps secret. A secure cryptosystem uses private keys that are very difficult to compute given the public key. One such system is *ElGamal Public Key Encryption* described below.

Construction 4.1. *(ElGamal Public Key Encryption)* Before anything can be encrypted, Bob must first choose a public key. To do so, he first chooses an elliptic

curve E on a finite field $\mathbb{F}_q$. He then selects a point P on E and an integer n. Let $A = nP$. Then

$$\{E, \mathbb{F}_q, P, A\} \tag{4.2}$$

is Bob's public key. His private key is n. It may appear that computing A would be a tedious task for Bob, but there exist algorithms for computing nP very efficiently. The relative ease with which Bob can find A is central to usefulness of this method of encryption.

Now, for Alice to send a message to Bob, she must do several things.

(i) She must somehow encode her message M as a point on E.
(ii) Choose an integer m and compute $M_1 = mP$.
(iii) Compute $M_2 = M + mA$

There are many ways to do (i). For example, if q is sufficiently large (and in practical applications it usually is), Alice can associate an integer with each letter of her message. e.g., "hello" $\mapsto (8050, 120120150)$ in which "h" maps to 80, "e" to 50, and so on. The comma placement is arbitrary. Alice sends $\{M_1, M_2\}$ to Bob. Bob can now decrypt her message with the following calculation.

$$M = M_2 - nM_1$$

Proposition 4.3. *The described method of decryption works.*

Proof. By definition, $M_2 - nM_1 = (M + mA) - n(mP)$. But $mA = m(nP)$, so $M_2 - nM_1 = M + m(nP) - n(mP) = M$. □

Suppose now an eavesdropper, Eve, tries to read Alice's message to Bob. Eve knows $E, \mathbb{F}_q, P, A, M_1, M_2$. To decrypt M, she needs to calculate n. This is roughly analogous to Eve solving $a^n \equiv b \pmod{p}$ for n. This is known as the *discrete logarithm problem*, and it is computationally very difficult. In practical situations, the integer n is several hundred digits long. Eve's calculation, even with a powerful computer, could take years.

There are numerous other cryptographic algorithms based on the group law of elliptic curves, and some are even used in modern security software. But elliptic curves have applications in other areas, too. The various understandings of elliptic curves—as geometric or algebraic objects, as complex tori, etc.—allow for applications in many areas of math. For example, Andrew Wiles' proof of Fermat's Last Theorem involves reducing the claim, ultimately, to one about elliptic curves.

Acknowledgments. It is a pleasure to thank my mentor, Preston Wake, for his guidance, support, and limitless patience.

References

[1] Walter Rudin. Real And Complex Analysis. McGraw-Hill, Inc. 1974.
[2] Lawrence C. Washington. Elliptic Cuves: Number Theory And Cryptography, Second Edition. Chapman & Hall/CRC. 2008.