Elliptic Curve Cryptography

Gabrielle Gaudeau
Supervisor: Harry Braden

Year 4 Project
School of Mathematics
University of Edinburgh
2021/2022

Abstract

Non-singular elliptic curve may be given the structure of an abelian group. If we work with a finite field we have a finite group and the discrete logarithm problem given two points P and $n \cdot P$ (that is $P + P + \cdots$ n times) is to determine n. When determining n is hard this may be the basis for its use in public-key cryptography. The project aims to unpack the previous sentences by introducing elliptic curves and their properties.

Declaration

I declare that this report was composed by myself and that the work contained therein is my own, except where explicitly stated otherwise in the text.

(*Gabrielle Gaudeau*)

Contents

Chapter 1

Introduction

> It is possible to write endlessly on elliptic curves. (This is not a threat.)
>
> Serge Lang — *Elliptic curves: Diophantine Analysis* (1978)

In this chapter, we provide motivation for the study of elliptic curves.

The study of elliptic curves can be traced back to the ancient Greeks with Diophantus and the topic has remained close to Diophantine geometry throughout the centuries. However, the name "elliptic curve" itself was only coined in the middle of the 18^{th} century in the work of Giulio C. Fagnano who showed that computing the arc length of an ellipse leads to the integral:

$$I(x) = \int \frac{dx}{\sqrt{x^3 + a \cdot x^2 + b \cdot x + c}}.$$

As we will see in the next Chapter, $E : y^2 = x^3 + a \cdot x^2 + b \cdot x + c$ defines an elliptic curve [1, Section 3]. The name is somewhat unfortunate seeing as an ellipse is not an elliptic curve [2, Section 12.1], but this is where it comes from. Elliptic curves have come a long way from Diophantine equations (polynomial equations for which we try to find rational or integer solutions) and is now playing an increasingly important role in number theory [3, Preface]. In just the last few decades, its applications to cryptography have become widespread [4].

The outline of this report is as follows. Chapter 2 starts by presenting the more common definitions of elliptic curves. Chapter 3 slowly builds on these definitions to define the abelian group structure of elliptic curves over arbitrary fields which we then discuss in more detail in Chapter 4 as we focus on finite fields. These chapters give a thorough background in the theory of elliptic curves as grounds to explore some of their applications to cryptography in Chapter 5.

This report is addressed to my peers with interests in mathematics and its applications to cryptography, and should be understandable for an advanced undergraduate that has taken

courses in Number and Group theory, Algebra and Computer Science. Throughout, you will find a mixture of original and cited proofs. I have tried to fill any gaps I encountered along the way, adding more steps and comments to convince myself and the reader of the truth of the results.

Note that a Python Jupyter notebook is attached to this report with my implementation of elliptic curve point addition over the real numbers and over finite fields. I also included examples for the Classic and Elliptic Curve Diffie-Hellman and fully implemented Lenstra's factorisation algorithm using the equations presented in literature and [5]. Unless otherwise stated, the examples and figures are my own. The code for all the plots is also included in the notebook.

Chapter 2

Definitions

> It is true that a mathematician who is not somewhat of a poet, will never be a perfect mathematician.
>
> Karl Weierstrass — *Letter to Sofia Kovalevskaya* (August 27, 1883)

You can find as many different definitions of elliptic curves as different elliptic curves themselves. These definitions, while in appearance quite similar, vary greatly in scope, formality and precision. In this chapter, we give the more common definitions and highlight some important properties of elliptic curves.

2.1 Weierstrass Equations

Elliptic curves are very concrete objects: they are essentially a set of points satisfying a cubic equation. You will find, however, that there are some subtleties in the way we define them. Perhaps surprising at first, these subtleties give rise to a rich structure we will discuss in later chapters.

Definition 2.1.1 (Elliptic curve)**.** An **elliptic curve** E over a field K is defined as the set of points $(x, y) \in K^2$ satisfying the equation

$$E : y^2 + a_1 \cdot x \cdot y + a_3 \cdot y = x^3 + a_2 \cdot x^2 + a_4 \cdot x + a_6, \tag{2.1}$$

where a_1, a_2, a_3, a_4, a_6 are constants, along with a **point at infinity** denoted $\mathcal{O}$. This equation is called the **generalised Weierstrass equation** for an elliptic curve.

Remark 2.1.2. For now, we will treat $\mathcal{O}$ as a formal symbol with some useful properties. We will define it formally in Section 2.3 and investigate its properties in Chapter 3.

For geometrical intuition, it is often useful to think of elliptic curves in terms of graphs over the real numbers. Figure 2.1 shows the cubic $y^2 = x^3 - x$ over $\mathbb{R}$ with three distinct real roots. However, it is important not to forget that elliptic curves can be defined over any field K. Later chapters will focus on elliptic curves over the **finite fields**.

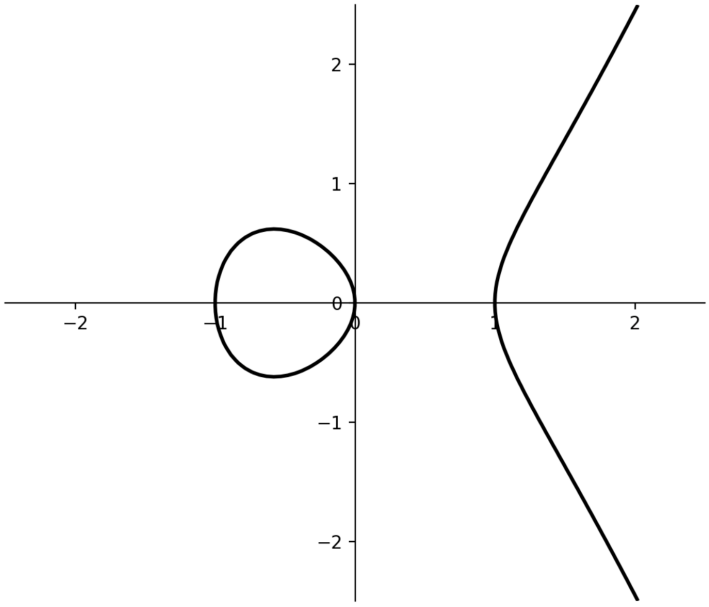


Figure 2.1: Plot of the elliptic curve $E : y^2 = x^3 - x$ over $\mathbb{R}$

The generalized Weierstrass equation is particularly useful when working with fields of **characteristic** 2 and 3. Recall that the **characteristic** of a field is the smallest number of times you need to add the multiplicative identity to get the additive identity.

If the characteristic of the field is any other, we can transform the equation of any elliptic curve into a simpler equation using the following method [3, p.10]:

- If the characteristic of the field is not 2, divide (2.1) by 2 and complete the square to obtain

$$\left(y + \frac{a_1 \cdot x}{2} + \frac{a_3}{2}\right)^2 = x^3 + \left(a_2 + \frac{a_1^2}{4}\right) \cdot x^2 + \left(a_4 + \frac{a_1 \cdot a_3}{2}\right) \cdot x + \left(\frac{a_3^2}{4} + a_6\right).$$

- Let $Y = y + \frac{a_1 \cdot x}{2} + \frac{a_3}{2}$, and replace the constant factors by a_2', a_4', a_6', to rewrite as

$$Y = x^3 + a_2' \cdot x^2 + a_4' \cdot x + a_6'.$$

- Seeing as the characteristic of the field is not 3, we can let $X = x + \frac{a_2'}{3}$ and obtain an equation of the form (2) for some constants a, b

$$Y^2 = X^3 + a \cdot X + b.$$

Definition 2.1.3 (Weierstrass equation)**.** Let E be an elliptic curve over a field K with an equation of the form

$$E : y^2 = x^3 + a \cdot x + b, \tag{2.2}$$

where a and b are constants. We call (2.2) the **(short) Weierstrass equation** for E.

Henceforth, we will assume that any considered elliptic curve is given in (short) Weierstrass form unless stated otherwise for simplicity purposes.

2.2 Singularity

In this section, we review another aspect of the definition of elliptic curves. There is some divide over the **singularity** of elliptic curves. Some split elliptic curves into two groups, singular or non-singular, while others solely define elliptic curves as "smooth" (i.e. non-singular) as we have done in Definition 2.1.1. The choice is somewhat an arbitrary one but there is a definite need to look into what singularity for elliptic curves means. Before we can go any further, we need to introduce several objects and their properties.

Definition 2.2.1 (Sylvester matrix)**.** Given two univariate polynomials over a field K

$$\begin{aligned} p(x) &= a_n \cdot x^n + a_{n-1} \cdot x^{n-1} + \cdots + a_1 \cdot x + a_0, \\ q(x) &= b_m \cdot x^m + b_{m-1} \cdot x^{m-1} + \cdots + b_1 \cdot x + b_0 \end{aligned}$$

of degrees n and m, respectively, the associated **Sylvester matrix** S is a square matrix of size $(n + m)$ given by:

$$\mathrm{S}(p, q) = \begin{pmatrix} a_n & a_{n-1} & \cdots & a_0 & 0 & \cdots & 0 \\ 0 & a_n & a_{n-1} & \cdots & a_0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_n & a_{n-1} & \cdots & a_0 \\ b_m & b_{m-1} & \cdots & b_0 & 0 & \cdots & 0 \\ 0 & b_m & b_{m-1} & \cdots & b_0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & b_m & b_{m-1} & \cdots & b_0 \end{pmatrix}.$$

Definition 2.2.2 (Resultant)**.** Given two univariate polynomials p and q over a field K, the **resultant** $\rho(p, q)$ is commonly defined as the determinant of their Sylvester matrix.

Proposition 2.2.3. The polynomials p and q have a common a common root in K if, and only if, the resultant $\rho(p, q)$ vanishes.

Proof. Let p and q be two univariate polynomials over a field K of degrees n and m respectively. Assume that p and q have a common factor h, say $p(x) = s(x) \cdot h(x)$ and $q(x) = -r(x) \cdot h(x)$ for non-zero polynomials

$$r(x) = c_{m-1} \cdot x^{m-1} + c_{m-2} \cdot x^{m-2} + \cdots + c_1 \cdot x + c_0$$

and

$$s(x) = d_{n-1} \cdot x^{n-1} + d_{n-2} \cdot x^{n-2} + \cdots + d_1 \cdot x + d_0$$

of degrees $m-1$ and $n-1$, respectively. Then r and s form a linear combination of p and q,

$$r(x) \cdot p(x) + s(x) \cdot q(x) = 0.$$

Multiplying out and collecting the terms yields

$$\begin{aligned} 0 = (c_{m-1} \cdot a_n + d_{n-1} \cdot b_m) \cdot x^{n+m-1} + (c_{m-1} \cdot a_{n-1} + c_{m-2} \cdot a_n + d_{n-1} \cdot b_{m-1} + d_{n-2} \cdot b_m) \cdot x^{n+m-2} + \\ \cdots + \left(\sum_{j=1}^{i} c_{m-j} \cdot a_{n+j-i} + \sum_{k=1}^{i} d_{n-k} \cdot b_{m+k-i} \right) \cdot x^{n+m-i} + \cdots \\ + (c_1 \cdot a_0 + c_0 \cdot a_1 + d_1 \cdot b_0 + d_0 \cdot b_1) \cdot x + (c_0 \cdot a_0 + d_0 \cdot b_0). \end{aligned}$$

In order for this to be equal to the zero polynomial, all the individual coefficients must be zero. We can write this as a homogeneous system of linear equations:

$$(0, \cdots, 0) = (c_{m-1}, \cdots, c_0, d_{n-1}, \cdots, d_0) \cdot \begin{pmatrix} a_n & a_{n-1} & \cdots & a_0 & 0 & \cdots & 0 \\ 0 & a_n & a_{n-1} & \cdots & a_0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_n & a_{n-1} & \cdots & a_0 \\ b_m & b_{m-1} & \cdots & b_0 & 0 & \cdots & 0 \\ 0 & b_m & b_{m-1} & \cdots & b_0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & b_m & b_{m-1} & \cdots & b_0 \end{pmatrix}.$$

We recognise the matrix on the right as the Sylvester matrix of p and q.

Now, in order for the system to have a non-trivial solution, it is necessary and sufficient that the determinant of the Sylvester matrix (i.e. the resultant) be zero. This concludes our proof that the polynomials p and q have a common a common root in K if, and only if, the resultant $\rho(p, q)$ vanishes.

□

Now we are ready to look at what **singularity** means for elliptic curves.

Definition 2.2.4 (Singularity)**.** Let E be a cubic plane curve in x, y over the field K. E is said to be **singular** (conversely, **non-singular** or **smooth**) if the partial derivatives $\frac{\partial E}{\partial x}$ and $\frac{\partial E}{\partial y}$ vanish simultaneously at one or more points on the curve, meaning that they share one or more common roots.

Geometrically, non-singular elliptic curves have no **cusps** or **self-intersections** over $\mathbb{R}$ (see Figure 3.4 for graphs of singular elliptic curves), and algebraically, their equations have three distinct roots. This is equivalent to the following condition.

Proposition 2.2.5. An elliptic curve E is said to be non-singular if, and only if,

$$\Delta = 4 \cdot a^3 + 27 \cdot b^2 \neq 0. \tag{2.3}$$

We call this number the **discriminant** of E.

Proof. Let E be an elliptic curve over the field K given by

$$E(x, y) = y^2 - (x^3 + a \cdot x + b) = 0.$$

We compute

$$\frac{\partial E}{\partial x} = -(3 \cdot x^2 + a) \quad \text{and} \quad \frac{\partial E}{\partial y} = 2 \cdot y.$$

Suppose the partial derivatives vanish at the point $(x_0, y_0) \in E$, we must have $2 \cdot y_0 = 0$ which implies $-p'(x_0) = 0$. Substituting $y_0 = 0$ into the original equation, we also get $p(x_0) = 0$.

We thus want to know the condition on a and b under which the polynomials $p(x)$ and $p'(x)$ do not have common roots in K. Consider their resultant:

$$\rho(p, p') = \det \begin{pmatrix} 1 & 0 & a & b & 0 \\ 0 & 1 & 0 & a & b \\ 3 & 0 & a & 0 & 0 \\ 0 & 3 & 0 & a & 0 \\ 0 & 0 & 3 & 0 & a \end{pmatrix} = 4 \cdot a^3 + 27 \cdot b^2.$$

Let $\Delta = 4 \cdot a^3 + 27 \cdot b^2$. According to Proposition 2.2.3, $p(x)$ and $p'(x)$ have a common root in K if, and only if, Δ is zero. Using Definition 2.2.4, this proves our claim.

□

Note that there exists a more general form of the discriminant for elliptic curves whose equations cannot be reduced to the Weierstrass short form [6, Chapter 3].

Remark 2.2.6. Why do we focus on non-singular curves? It turns out that in Cryptography, the use of singular curves is insecure as we will discuss in Section 5.1.

2.3 Projective Space

In this section we describe the "space" in which elliptic curves arise. Back when we were in school, we were taught that parallel lines never met at infinity (in Euclidean geometry). However, imagine you are standing between the rails of a train track, looking out towards the horizon. From your standpoint, the rails are parallel lines and appear to intersect as they recede into the distance. **Projective space**, as an extension of Euclidean space, allows us to make sense of this.

Definition 2.3.1 (Projective Space)**.** Let K be a field. The **two-dimensional projective space** is the set $\mathbb{P}^2(K)$ of all nonzero triples (x, y, z), with $x, y, z \in K$, modulo the equivalence relation

$$(x, y, z) \sim (\lambda \cdot x, \lambda \cdot y, \lambda \cdot z).$$

The **projective point** $(x : y : z)$ is the equivalence class of (x, y, z).

Take $(x : y : z) \in \mathbb{P}^2(K)$. If $z \neq 0$, then $(x : y : z) \sim (x/z : y/z : 1)$. These are called **affine points** and they form the two-dimensional **affine (Euclidean) plane**

$$\mathbb{A}^2(K) = \{(x, y) \in K \times K\} \hookrightarrow \mathbb{P}^2(K).$$

On the other hand, if $z = 0$, then dividing by z can be thought as having ∞ in the x-coordinate or the y-coordinate. We call points of the form $(x : y : 0)$ **points at infinity**.

Remark 2.3.2. Notice that the point $(0 : 0 : 0)$ is not allowed in projective space.

An elliptic curve in the affine plane $\mathbb{A}^2(K)$, can be projected as a cubic curve in the projective plane to become a **projective plane curve**.

Definition 2.3.3 (Projective Plane Curve)**.** A **plane projective curve** $C_f(K)$ is a homogeneous polynomial $f(x, y, z)$ with coefficients in K, where the set

$$C_f(K) = \{(x : y : z) \in \mathbb{P}^2(K) \mid f(x, y, z) = 0\}.$$

More concretely, let $E(K)$ be an elliptic curve given by (2.2). Set $x = X/Z$ and $y = Y/Z$. The equation becomes

$$\left(\frac{Y}{Z}\right)^2 = \left(\frac{X}{Z}\right)^3 + a \cdot \left(\frac{X}{Z}\right) + b.$$

Multiplying by Z^3 to clear the denominators yields

$$Y^2 \cdot Z = X^3 + a \cdot X \cdot Z^2 + b \cdot Z^3, \tag{2.4}$$

a homogeneous equation in three variables (all monomials have degree 3) called the **projectivisation** of E.

Remark 2.3.4. Notice that if (x, y, z) is a solution to the equation, then so is $(\lambda \cdot x, \lambda \cdot y, \lambda \cdot z)$ for any non-zero λ. Indeed, in projective space, these are one and the same solution $(x : y : z)$.

As noted previously, there are two types of such solutions. If $z \neq 0$, then the projective point $(x/z : y/z : 1)$ is a point on the projective curve. However, substituting $z = 0$ into (2.4) yields $x = 0$, and y can take any non-zero values. So there is precisely one projective point $(0 : 1 : 0)$ on E since $(0, y, 0) \sim (0, 1, 0)$. Geometrically, it can be thought of as a point infinitely high and infinitely low on the y-axis. The point at which all vertical parallel lines meet on E.

Proposition 2.3.5. The only point at infinity on an elliptic curve is $(0 : 1 : 0)$, denoted by $\mathcal{O}$.

It is no coincidence that we denote the point at infinity on an elliptic curve in the same way as the distinguished point $\mathcal{O}$ in Definition 2.1.1. They are one and the same. That is, we consider $E(K)$ to be the set

$$E(K) = \{\mathcal{O}\} \cup \{(x, y) \in K \times K \mid y^2 = x^3 + a \cdot x + b\}.$$

2.4 Genus

In this last section, we mention an alternative way of defining elliptic curves. Addressing it sheds further light on the relationship between elliptic curves and projective space.

Definition 2.4.1. An elliptic curve E over a field K is a smooth projective algebraic plane curve of **genus** 1 on which we distinguish a K**-rational point** $\mathcal{O}$.

Remark 2.4.2. By K**-rational point** we mean that it is a point whose coordinates (x, y) belong to a given field. Notice again that we distinguish a point $\mathcal{O}$ on E. It so happens that not every smooth projective curve of genus 1 is an elliptic curve. The presence of a rational point is a necessary condition. It may very well seem unnatural but we will later see that including it has very useful consequences.

The **genus** of an algebraic curve can be intuitively understood as the measure of its geometric complexity [7, Remark 3, p.10]. One might reasonably seek to classify algebraic curves according to the **degree** of the polynomial that define their curve equations. Unfortunately, this classification method has its limits.

Example 2.4.3. Consider two non-singular algebraic curves $C_i \in \mathbb{P}^2(K)$ for $i = 1, 2$,

$$C_1 : x = 0 \qquad \text{and} \qquad C_2 : x \cdot z - y^2 = 0,$$

and the mappings $u : C_1 \to C_2$ and $v : C_2 \to C_1$,

$$u : (x : y : z) \mapsto (y^2 : y \cdot z : z^2)$$
$$v : (x : y : z) \mapsto (0 : y : z).$$

We can show that under these mappings, C_1 and C_2 are isomorphic.

For all $(x : y : z) \in C_1$ we have $x = 0$. Then

$$(v \circ u)(x : y : z) = v(y^2 : y \cdot z : z^2) = (0 : y \cdot z : z^2).$$

We must have $z \neq 0$ (recall that $(0 : 0 : 0)$ is not allowed in projective space), then

$$(v \circ u)(x : y : z) = (0 : y \cdot z : z^2) = (0 : y : z) = (x : y : z).$$

And for all $(x : y : z) \in C_2$,

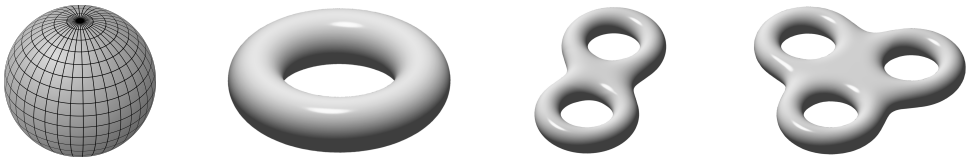
$$(u \circ v)(x : y : z) = u(0 : y : z) = (y^2 : y \cdot z : z^2).$$

On C_2, if $z = 0$ then $y = 0$. However, we cannot have $(0 : 0 : 0)$ in projective space, so it must be that $z \neq 0$. Using the equation for C_2 we have $x = y^2/z$. Then

$$(u \circ v)(x : y : z) = (y^2 : y \cdot z : z^2) = (y^2/z : y : z) = (x : y : z).$$

We would want a measure of geometric complexity to be invariant under isomorphism. However, the above example presents two isomorphic curves of different degrees: $\deg(C_1) = 1$ while $\deg(C_2) = 2$. In this instance, the degree is not a good enough measure. The genus on the other hand is a topologically invariant property of algebraic curves, notably invariant under isomorphism [8].

Over $\mathbb{C}$, the genus of an (orientable) surface is defined as the number of holes (or handles) it has (Figure 2.2). The genus can however be defined algebraically over any field (the formal definition of the genus is actually quite complex and outside of the scope of this report).



(a) Genus 0 (b) Genus 1 (c) Genus 2 (d) Genus 3

Figure 2.2: Plots of orientable surfaces of different genera [9]

Definition 2.4.1 states that elliptic curves are of genus 1 and "it is also possible to show that elliptic curves always have genus 1" [3, p.370]. In particular, genus 1 curves are birationally isomorphic to smooth cubic curves in two-dimensional projective space $\mathbb{P}^2(K)$ [8], that is, elliptic curves can always be projected as cubic curves in the the projective plane.

Chapter 3

Group Law

> *You said something [...] about addition of points on an elliptic curves being defined in a peculiar way for a particular reason. What's the reason?*
> Now *that's* a story I'd love to tell - but the margin of this paper is not large enough to contain it. As my grandmother used to say, "Tell you tomorrow!"
>
> Erza Brown — *Three Fermat Trails to Elliptic Curves* (2000)

The group law on elliptic curves is what makes the theory of elliptic curves so special, and sheds light on some of the particularities encountered in the previous chapter. Unlike Brown, I will tell you all about it today.

3.1 Groups

In this section we recall from Fundamentals of Pure Mathematics some basic definitions in **group** theory.

Definition 3.1.1 (Group)**.** A **group** $G = (S, \cdot)$ is a set S together with a binary operation

$$\cdot : S \times S \to S$$

such that:

- the operation $\cdot$ is **associative**, that is, for all $P, Q, R \in S$ we have

$$(P \cdot Q) \cdot R = P \cdot (Q \cdot R);$$

- there exists an **identity** $\mathcal{O} \in S$ such that for all $P \in S$ we have $P \cdot \mathcal{O} = \mathcal{O} \cdot P = P$;

- for all $P \in S$, there exists an **inverse** $P^{-1} \in S$ such that $P \cdot P^{-1} = P^{-1} \cdot P = \mathcal{O}$.

Groups are algebraic structures that naturally arise in the study of geometry (groups of symmetries), and the analysis of polynomials and their roots (Galois group). They have many applications, and are widely used in cryptosystems, including elliptic curve cryptography.

Remark 3.1.2. Looking at the definition, the condition on associativity implies that parentheses can be dropped altogether. We can write the product of n elements $P_1, P_2, \cdots, P_n \in G$

$$P_1 \cdot P_2 \cdots P_n,$$

without worrying about the order in which we evaluate each individual term. However, the order of individual elements matters, and there is no guarantee that commutativity will hold, that is,

$$P \cdot Q = Q \cdot P,$$

is not necessarily true for all $P, Q \in G$. This brings us to the definition of an **abelian group**.

Definition 3.1.3 (Abelian Group)**.** A group whose operation $\cdot$ is also **commutative**, that is, for all $P, Q \in S$ we have

$$P \cdot Q = Q \cdot P,$$

is called an **abelian group** (or **commutative group**).

Recall that two groups, $(G_1, *)$ and $(G_2, \cdot)$, are said to be **isomorphic** if there exists a bijection $\psi : G_1 \to G_2$ such that $\psi(g * h) = \psi(g) \cdot \psi(h)$ for all $g, h \in G_1$. The following theorem is very important result of group theory taken from [3, Theorem B.3].

Theorem 3.1.4 (Classification of finite abelian groups)**.** A finite abelian group G is **isomorphic** to a group of the form

$$C_{n_1} \oplus C_{n_2} \oplus \cdots \oplus C_{n_s}$$

with $n_i \mid n_{i+1}$ for $i = 1, 2, \cdots, s-1$. The integers n_i are uniquely determined by G.

Here, C_n denotes the **finite cyclic group of order** n, the group order n that can be finitely generated from a single element. We will need to distinguish the **group order** (the number of elements in a group) from the **order** of an element of a group.

Definition 3.1.5 (Order)**.** An element P of a group is said to have **order** m if

$$P^m = \underbrace{P \cdot P \cdots P}_{m \text{ times}} = \mathcal{O}$$

and for all $m' \in \mathbb{N}$, with $1 \leq m' < m$,

$$P^{m'} \neq \mathcal{O}.$$

If such m exists, P is said to have **finite order**, otherwise it has **infinite order**.

A final result we will need is the following.

Proposition 3.1.6. If G is an abelian group and let $H \subset G$ be the subset of elements with finite order, then H is a **subgroup** of G. We write $H \leq G$.

Proof. Note first that H is nonempty since the identity element $\mathcal{O} \in G$ has order 1 which implies that $\mathcal{O} \in H$. Now, in order for H to be a **subgroup** of G, we need to show that it is closed under multiplication and inverses.

- Take any $P, Q \in H$ and let m, n be their respective orders, such that:

  $$P^m = \mathcal{O} \qquad \text{and} \qquad Q^n = \mathcal{O}.$$

  Using the fact that G is abelian, we compute

  $$\begin{aligned}(P \cdot Q)^{m \cdot n} &= P^{m \cdot n} \cdot Q^{m \cdot n} \\ &= (P^m)^n \cdot (Q^n)^m \\ &= \mathcal{O}^n \cdot \mathcal{O}^m \\ &= \mathcal{O}.\end{aligned}$$

  By definition, the order of $P \cdot Q$ is then at most $m \cdot n$, but since m, n are both finite integers, so is $m \cdot n$. Hence $P \cdot Q \in H$.

- Take $P \in H$ with finite order m. Consider now

  $$(P^{-1})^m = (P^m)^{-1} = \mathcal{O}^{-1} = \mathcal{O}.$$

  This implies that P^{-1} also has finite order and hence inverse of P is also in H.

We have shown that H is closed under multiplication and inverses, hence $H \leq G$.

□

3.2 Addition Laws

In this section, we show how the set of points of an elliptic curve given in the (short) Weierstrass form can be endowed with a binary operation to form an abelian group.

First we introduce the following notation. Take two distinct points $P_1, P_2 \in E(K)$ and draw the line $\mathcal{L}$ through them. We denote $P_1 * P_2$ the third point of intersection of the line $\mathcal{L}$ with E (we will return to the existence of this point later in this section, for now assume that it exists). If P_1 and P_2 are coincident (that is $P_1 = P_2$), then we denote $P_1 * P_1$ the intersection of the tangent line to E at P_1.

We define the binary operation of the group law as a point addition denoted $(+)$. To give the reader a better intuition of the algebraic **addition laws**, Figure 3.1 shows what the addition of points on elliptic curves looks like graphically.

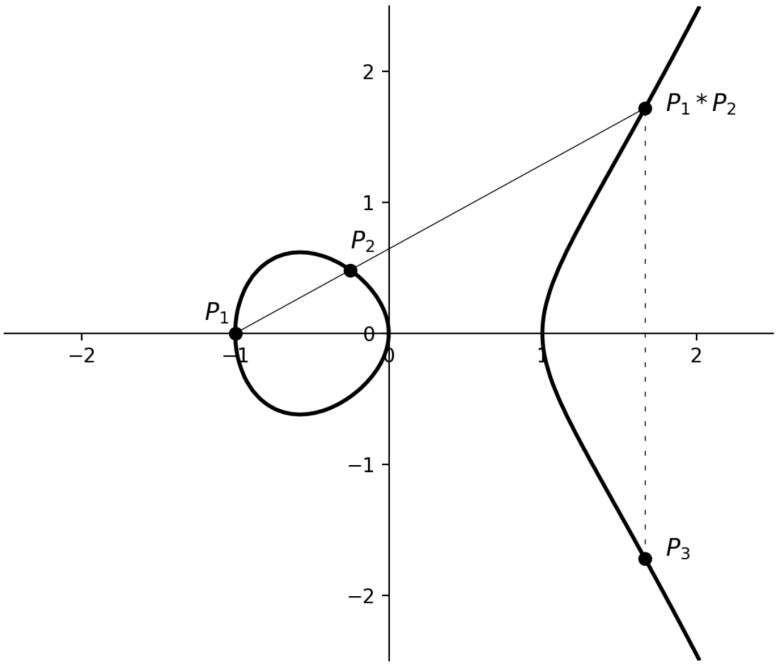


Figure 3.1: Adding $P_1 = (-1, 0)$ and $P_2 = (-\frac{1}{4}, \frac{\sqrt{15}}{8})$ on $E : y^2 = x^3 - x$ over $\mathbb{R}$

Consider a non-singular elliptic curve E over a field K given in the form (2.2):

$$E : y^2 = x^3 + a \cdot x + b$$

for a, b constants. Start with two points on the curve P_1, P_2 and draw the line $\mathcal{L}$ through them. We will see below that $\mathcal{L}$ intersects with E in one and only one other point which we denote $P_1 * P_2$. Reflect it across the x-axis to find P_3 which we define as:

$$P_3 = P_1 + P_2.$$

Put simply, the addition of two points on an elliptic curve is the negation of the point resulting from the intersection of the curve E and the line through those two points. Algebraically, this yields the **addition laws**.

Proposition 3.2.1 (Addition Laws)**.** Let $P_1 = (x_1, y_1), P_2 = (x_2, y_2)$ be two points on $E(K)$. We define the **addition** $(+)$ of P_1 and P_2 on E as $P_3 = (x_3, y_3) = P_1 + P_2$ where:

$$\begin{cases} x_3 = s^2 - x_1 - x_2, \\ y_3 = s \cdot (x_1 - x_3) - y_1. \end{cases}$$

When P_1 and P_2 are distinct, with $x_1 \neq x_2$, P_3 is the negation of the point resulting from the intersection of the curve E and the line $\mathcal{L}$ through P_1 and P_2. The slope s of $\mathcal{L}$ is:

$$s = \frac{y_2 - y_1}{x_2 - x_1}.$$

When P_1 and P_2 are coincident, with $y_1 \neq 0$, P_3 is the negation of the point resulting from the intersection of the curve E and the tangent to E at P_1. Implicit differentiation gives us:

$$s = \frac{3 \cdot x_1^2 + a}{2 \cdot y_1}.$$

These addition laws hold provided neither point is the point at infinity $\mathcal{O}$ encountered in Section 2.4.

Proof. We follow [3, Section 2.2] and consider two separate cases.

When P_1 and P_2 are distinct, with $x_1 \neq x_2$ ($\mathcal{L}$ is not vertical), the slope s of the line $\mathcal{L}$ is given by

$$s = \frac{y_2 - y_1}{x_2 - x_1}.$$

When P_1 and P_2 are coincident, we take $\mathcal{L}$ to be the tangent line to E through P_1. Using implicit differentiation on (2.2), we find

$$2 \cdot y \cdot \frac{\partial y}{\partial x} = 3 \cdot x^2 + a \implies s = \frac{3 \cdot x_1^2 + a}{2 \cdot y_1},$$

where s is the slope of $\mathcal{L}$, provided $y_1 \neq 0$.

Either way, the equation of $\mathcal{L}$ is given by:

$$y = s \cdot (x - x_1) + y_1. \tag{3.1}$$

Substituting (3.1) into (2.2), we find the three intersections of $\mathcal{L}$ and E by finding the three roots of the cubic

$$(s \cdot (x - x_1) + y_1)^2 = x^3 + a \cdot x + b. \tag{3.2}$$

Which can be rearranged in the form of a cubic in x:

$$0 = x^3 - s^2 \cdot x^2 + \cdots$$

The fact that the intersection of $\mathcal{L}$ and E consists of exactly three points (with multiplicities) is a special case of **Bézout's theorem** [10, Theorem I.7.8]. Since P_1 and P_2 both belong to E and $\mathcal{L}$, we already know two roots x_1 and x_2. We find the third root x_3' using Viète's formulas,

$$x_1 + x_2 + x_3' = -(-s^2) \implies x_3' = s^2 - x_1 - x_2.$$

Substituting into (3.1), we obtain $y_3' = s \cdot (x_3' - x_1) + y_1$. Reflect $P_1 * P_2 = (x_3', y_3')$ across the x-axis to find P_3:

$$\begin{cases} x_3 = s^2 - x_1 - x_2, \\ y_3 = s \cdot (x_1 - x_3) - y_1, \end{cases}$$

as required.

□

Notice that we have conveniently avoided some cases (for example when $\mathcal{L}$ is vertical). Recall from Section 2.4 that the set of points on an elliptic curve contains a unique point at infinity $\mathcal{O}$ which can be thought as a point both infinitely high and infinitely low on the y-axis, the point at which all vertical parallel lines meet on E. We complete our addition laws with the following proposition.

Proposition 3.2.2. Let $P_1 = (x_1, y_1), P_2 = (x_2, y_2)$ be two points on $E(K)$.

- If $x_1 = x_2$ and $y_1 \neq y_2$, then we have $P_1 + P_2 = \mathcal{O}$. In particular, $P + (-P) = \mathcal{O}$ for all $P \in E(K)$ where $-P$ is the reflection of P across the x-axis.
- If $P_1 = P_2$ and $y_1 = 0$, then we have $P_1 + P_2 = \mathcal{O}$.
- For all $P \in E(K)$, we have that $P + \mathcal{O} = P$.

Proof. We follow [3, p.13–14] and prove by cases.

- Suppose first that $x_1 = x_2$ and $y_1 \neq y_2$. Then the line $\mathcal{L}$ through P_1 and P_2 is vertical and intersects E at $\mathcal{O}$. This is the only other intersection since a line and a cubic can intersect in at most three distinct points. Reflecting $\mathcal{O}$ across the x-axis yields the same point $\mathcal{O}$ since we think of that point as the top and the bottom of the y-axis. We thus find that $P_1 + P_2 = \mathcal{O}$.
- Now suppose that $P_1 = P_2$ and $y_1 = 0$. The the line through $\mathcal{L}$ through P_1 and P_2 is the vertical tangent to E at P_1 which intersects E at $\mathcal{O}$. As before, we reflect $\mathcal{O}$ across the x-axis and find the same point $\mathcal{O}$. We thus find that $P_1 + P_2 = \mathcal{O}$.
- Finally, choose any $P \in E(K)$. The line $\mathcal{L}$ through P and $\mathcal{O}$ is a vertical line that intersects E in the point $-P = P * \mathcal{O}$, the reflection of P across the x-axis. Reflecting $-P$ across the x-axis to get $P + \mathcal{O}$, we are back at P. Therefore, $P = P + \mathcal{O}$.

□

Remark 3.2.3. The above proof lacks somewhat in rigour. Indeed, if we want to properly deal with $\mathcal{O}$ as any other point we need to use projective coordinates. However, projective coordinates tend to make formulas and proofs significantly more complicated and lengthy. Should you wish to convince yourself that we have not strayed away from the definition of the point at infinity given in Section 2.3, see [3, p.67–68] for the group law given in projective coordinates.

Using the addition laws and Proposition 3.2.2, we can now formalise the **group law** on elliptic curves.

Theorem 3.2.4 (Group Law)**.** The addition $(+)$ of points on an elliptic curve has the following properties:

- the operation is **associative**, that is, for all $P_1, P_2, P_3 \in E(K)$,

$$(P_1 + P_2) + P_3 = P_1 + (P_2 + P_3);$$

- there exists an **identity** $\mathcal{O} \in E(K)$ such that for all $P \in E(K)$ we have

$$P + \mathcal{O} = \mathcal{O} + P = P;$$

- for all $P \in E(K)$, there exists an **inverse** $-P \in E(K)$ such that

$$P + (-P) = (-P) + P = \mathcal{O};$$

- the operation is **commutative**, that is, for all $P_1, P_2 \in E(K)$,

$$P_1 + P_2 = P_2 + P_1.$$

Proof. We devote the next section to the proof of associativity. The identity and inverse properties hold by Proposition 3.2.2.

Commutativity is straightforward to prove. Graphically, since the line $\mathcal{L}$ through P_1 and P_2 is the same as the one through P_2 and P_1, it follows that $P_1 + P_2 = P_2 + P_1$. Note that you can also check this algebraically using the addition laws (but we omit this here).

□

Coming back to the quote that introduced this Chapter: we have seen the "particular way" in which the addition of points on an elliptic curve is defined. As to the "particular reason" for it: Theorem 3.2.4 implies that the set $E(K)$ together with the point addition $(+)$ binary operation form an abelian group.

3.3 Proof of Associativity

Associativity is by far the more subtle of the four properties of point addition. In fact, it is quite surprising that an operation defined in such an apparently arbitrary way could prove to be associative. In this section we will go through the proof as promised, but before we can go any further we need to mention the following result.

Theorem 3.3.1 (Cubic Cayley-Bacharach Theorem)**.** Let C_1 and C_2 be cubic curves in $\mathbb{P}^2$ without common components of respective degrees, and suppose that C_1 and C_2 intersect in nine distinct points. Let E be a cubic curve also in $\mathbb{P}^2$. If E passes through all but one of the points of $C_1 \cap C_2$, then it must also pass through the remaining point.

For a proof of this fundamental result see [11, A.3]. In this report, we give one of the classical proofs of associativity whose credit of which lies entirely with [11, Section 1.2]. In this section, we only seek to illustrate the proof further, and invite the reader to follow along and convince themselves that everything checks out geometrically.

Proposition 3.3.2 (Associativity)**.** Point addition $(+)$ is associative. That is,

$$(P_1 + P_2) + P_3 = P_1 + (P_2 + P_3),$$

for all points P_1, P_2, and P_3 on $E(K)$.

Proof. We follow [11, p.14–15]. Take P_1, P_2 and P_3 to be three points on $E(K)$. First note that $(P_1 + P_2) + P_3 = ((P_1 + P_2) * P_3) * \mathcal{O}$ and $P_1 + (P_2 + P_3) = (P_1 * (P_2 + P_3)) * \mathcal{O}$. To prove associativity then amounts to showing that

$$(P_1 + P_2) * P_3 = P_1 * (P_2 + P_3). \tag{3.3}$$

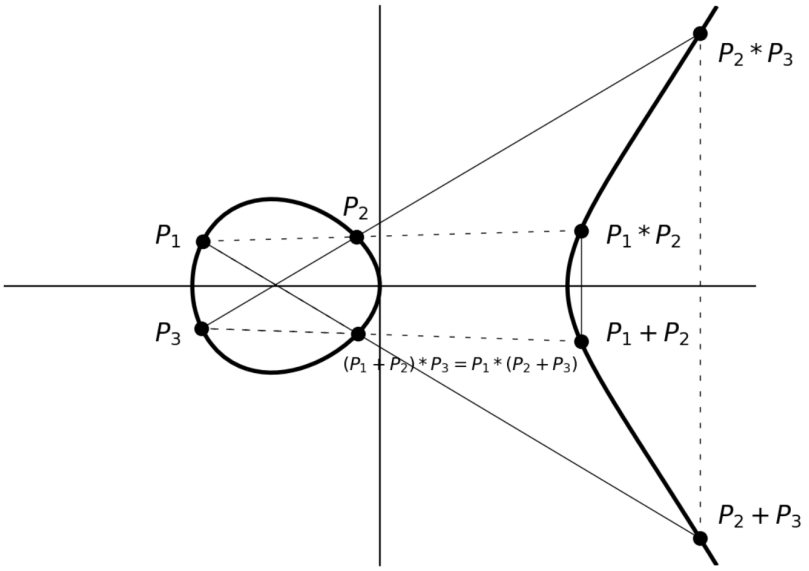


Figure 3.2: Associativity on $E : y^2 = x^3 - x$ over $\mathbb{R}$

To form $(P_1 + P_2) * P_3$, we first need $P_1 + P_2$. Start by drawing the line through P_1 and P_2 to find $P_1 * P_2$, join that to $\mathcal{O}$, and take the third intersection of that line with E, which is $P_1 + P_2$. Then we must join this point to P_3 to find the third intersection of the line with E, which is $(P_1 + P_2) * P_3$. We can form $P_1 * (P_2 + P_3)$ in a similar fashion.

Figure 3.2 shows a visualisation of the lines we have drawn so far for the curve $E : y^2 = x^3 - x$ over $\mathbb{R}$. We can see that each of the points

$$\mathcal{O}, P_1, P_2, P_3, P_1 * P_2, P_1 + P_2, P_2 * P_3, P_2 + P_3, \tag{3.4}$$

lie on one of the dashed lines and on one of the solid lines. In particular, $\mathcal{O}$ is on a dashed line and on a solid line since all (parallel) vertical lines pass through $\mathcal{O}$ in projective space. Note that each line is defined by a linear equation.

Consider the dashed line through $P_1 + P_2$ and P_3, and the solid line through P_1 and $P_2 + P_3$. If their intersection lies on E, that is, they intersect with E at the same point, then we have proven (3.3). Denote this intersection by Q.

Let C_1 to be the curve obtained by multiplying the linear equations of the three dashed lines and C_2 to be the curve obtained by multiplying the linear equations of the three solid lines. We obtain two cubic curves without common components and nine distinct intersection points: the eight points in (3.4) and Q. Since E passed through all eight points of $C_1 \cap C_2$, it follows from the Cayley-Bacharach theorem that E must therefore also pass through Q, proving (3.3) and our initial claim.

□

Remark 3.3.3. Throughout this proof we have made no mention of the Weierstrass equation for the elliptic curve, and in fact, the proof holds for any non-singular cubic curve.

3.4 Group Identity

In this section, we revisit our choice of the identity element for the group law. Indeed, recall that we *chose* to define the group law with $\mathcal{O} = (0 : 1 : 0)$ as our identity. Why this point in particular? Several answers exist to this question. The point at infinity $\mathcal{O}$ is invariably present on all elliptic curves (the K-rational point in Definition 2.4.1). It is also the convention. But more importantly perhaps, it is an **inflection point**.

Before we can define what an **inflection point** is, we need to introduce a new object.

Definition 3.4.1 (Hessian matrix)**.** Given a homogeneous polynomial equation in three variables $F(X, Y, Z)$, the associated **Hessian matrix** is a 3-by-3 square matrix given by:

$$\mathrm{H}_F(X, Y, Z) = \begin{pmatrix} \dfrac{\partial F^2}{\partial X^2} & \dfrac{\partial F^2}{\partial X \partial Y} & \dfrac{\partial F^2}{\partial X \partial Z} \\ \dfrac{\partial F^2}{\partial Y \partial X} & \dfrac{\partial F^2}{\partial Y^2} & \dfrac{\partial F^2}{\partial Y \partial Z} \\ \dfrac{\partial F^2}{\partial Z \partial X} & \dfrac{\partial F^2}{\partial Z \partial Y} & \dfrac{\partial F^2}{\partial Z^2} \end{pmatrix}.$$

Definition 3.4.2 (Inflection Point)**.** Let F be a homogeneous polynomial equation in three variables. The **inflection points** of the projective plane curve given by the implicit equation $F(X, Y, Z) = 0$ are exactly the non-singular points at which the determinant of its Hessian

matrix vanishes.

We have thus far claimed that $\mathcal{O} = (0 : 1 : 0)$ is an inflection point of any elliptic curve E. Let us prove this formally.

Proposition 3.4.3. Let E be an elliptic curve over K. The point $\mathcal{O} = (0 : 1 : 0)$ is an inflection point of $E(K)$.

Proof. Using the projectivisation of E given by (2.4), we let

$$F(X, Y, Z) = X^3 + a \cdot X \cdot Z^2 + b \cdot Z^3 - Y^2 \cdot Z,$$

so that E is the vanishing locus in $\mathbb{P}^2$ of the polynomial F, that is, E is the set of points where F vanishes.

Consider the Hessian matrix of F:

$$\mathrm{H}_F(X, Y, Z) = \begin{pmatrix} 6 \cdot X & 0 & 2 \cdot a \cdot Z \\ 0 & -2 \cdot Z & -2 \cdot Y \\ 2 \cdot a \cdot X & -2 \cdot Y & 2 \cdot a \cdot X + 6 \cdot b \cdot Z \end{pmatrix}.$$

Now evaluating the matrix at $\mathcal{O}$ yields

$$\mathrm{H}_F(\mathcal{O}) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -2 \\ 0 & -2 & 0 \end{pmatrix},$$

which has determinant equal to zero. We find that $\mathcal{O}$ must be an inflection point.

□

Remark 3.4.4. Notice that we define elliptic curves in a way that ensures that they always have an K-rational inflection point over any field K, that is, the point $\mathcal{O}$. For your interest, we can show that there are exactly 9 inflection points over $\mathbb{C}$ [12].

Why is this important? Let $I \in E$ be an inflection point of E. The reason we define the identity element as an inflection point is the following: we want any three points on E to sum to the identity element I if they lie on the same line. Then, and only then, can the negation of a point P can be defined as the third intersection of the line through I and P with E. Associativity of the group law with I as identity element will follow just as it did with

$\mathcal{O}$ [3, Remark 2.12]. The discussion of the group law is thus simpler when the identity element is an inflection point, and indeed, it is simplest when we choose $\mathcal{O}$ as the identity. As a result, most mathematicians put themselves in that situation when discussing the group law.

However, there are some exceptions, and it is indeed possible to define the group law choosing any point on E as the identity. In fact, all groups defined with different identities on the same curve E, while seemingly different, are actually isomorphic to each other [11, Section 1.2, p.15]. We state this formally.

Proposition 3.4.5. Let $\dot{\mathcal{O}}$ be any point on E. The map

$$\phi : P \mapsto P + \dot{\mathcal{O}}$$

is an isomorphism from the group $(E, \mathcal{O}, +)$ to the group $(E, \dot{\mathcal{O}}, \dot{+})$, where the new addition law is defined by

$$P_3 = P_1 \dot{+} P_2 = P_1 + P_2 - \dot{\mathcal{O}}.$$

We give a geometrical example of point addition with an arbitrary identity point $\dot{\mathcal{O}}$ in the figure below.

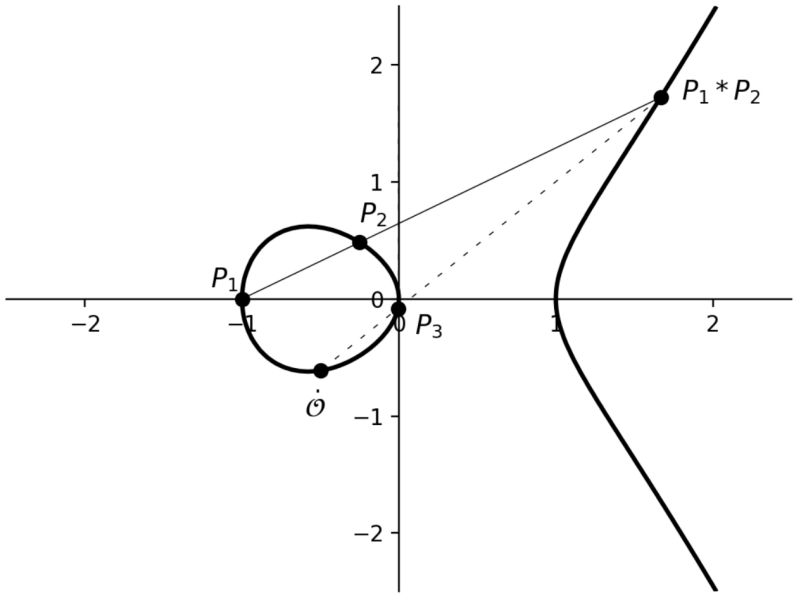
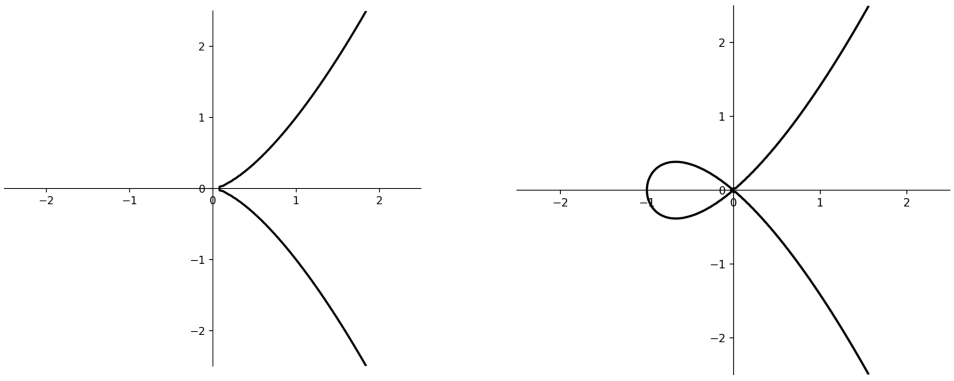


Figure 3.3: Adding $P_1 + P_2$ on $E : y^2 = x^3 - x$ over $\mathbb{R}$ with identity point $\dot{\mathcal{O}} = (-\frac{1}{2}, -\frac{\sqrt{6}}{4})$

3.5 Singular Curves

In this section we look at what happens to the Group Law when we consider elliptic curves with multiple roots (singular). Recall that an elliptic curve is singular when $\Delta = 4 \cdot a^3 + 27 \cdot b^2 = 0$ as given in Proposition 2.2.5. We will need to distinguish two cases (see Figure 3.4):

- when $a = 0$, the equation for E becomes $y^2 = x^3$ with a triple root, and geometrically, a **cusp**, at $(0, 0)$ as in (a);
- when $a \neq 0$, the equation for E has a double root, and geometrically, a **node**, at that point as in (b) for example.



(a) $E : y^2 = x^3$ (b) $E : y^2 = x^3 + x^2$

Figure 3.4: Example plots of singular elliptic curves over $\mathbb{R}$: (a) cusp and (b) node

Cusp

In the first case, notice that any line through the only singular point $(0, 0)$ intersects with E in at most one other point, and point addition as we have defined it cannot be performed. We therefore choose to exclude it and denote $E_{ns}(K)$ the remaining K-smooth points on E which include the point at infinity $\mathcal{O}$. Suppose that we then define the addition on $E_{ns}(K)$ much like we did in Section 3.2. Does it define a group? We follow [3, Theorem 2.30] as we prove that point addition on $E_{ns}(K)$ reduces to an addition of elements in K.

Let us define the parameter $t \in K$ such that for all points $(x, y) \in E_{ns}(K)$,

$$\begin{cases} t = 0, & \text{if } (x, y) = \mathcal{O} \\ t = x/y, & \text{otherwise.} \end{cases}$$

Theorem 3.5.1. Let E be the curve $y^2 = x^3$ and let $E_{ns}(K)$ be the K-smooth points on this curve. The map:

$$\begin{aligned} \psi : E_{ns}(K) &\to K \\ (x, y) &\mapsto t \end{aligned}$$

is a group isomorphism between $E_{ns}(K)$ and K, regarded as an additive group.

Proof. Using the equation for E, we can rewrite $x = (y/x)^2 = t^{-2}$ and $y = x/t = t^{-3}$ to express all points $(x, y) \in E_{ns}(K)$ in terms of the parameter t. It follows that ψ is a bijection.

It remains for us to prove that $P_1 + P_2 = P_3$ if, and only if, $t_1 + t_2 = t_3$ where $P_i = (x_i, y_i) \in E_{ns}(K)$ and $t_i = x_i/y_i \in K$ for $i = 1, 2, 3$. Using the addition laws and substituting in what we found above for x and y in terms of t:

- if P_1 and P_2 are distinct,

$$\begin{cases} t_3^{-2} = \left(\dfrac{t_2^{-3} - t_1^{-3}}{t_2^{-2} - t_1^{-2}}\right)^2 - t_1^{-2} - t_2^{-2} = (t_1 + t_2)^{-2} \\ \\ t_3^{-3} = \left(\dfrac{t_2^{-3} - t_1^{-3}}{t_2^{-2} - t_1^{-2}}\right) \cdot (t_1^{-2} - t_3^{-2}) - t_1^{-3} = (t_1 + t_2)^{-3} \end{cases} \implies t_3 = t_1 + t_2;$$

- if $P_1 = P_2$,

$$\begin{cases} t_3^{-2} = \left(\dfrac{3 \cdot (t_1^{-2})^2}{2 \cdot t_1^{-3}}\right)^2 - 2 \cdot t_1^{-2} = \dfrac{t_1^{-2}}{4} \\ \\ t_3^{-3} = \left(\dfrac{3 \cdot (t_1^{-2})^2}{2 \cdot t_1^{-3}}\right) \cdot (t_1^{-2} - t_3^{-2}) - t_1^{-3} = \dfrac{t_1^{-3}}{8} \end{cases} \implies t_3 = 2 \cdot t_1;$$

- if either point is $\mathcal{O}$, it is trivial to show that $t_3 = t_2$ if say $P_1 = \mathcal{O}$ (so $P_3 = P_2$), or reciprocally, that $t_3 = t_1$ if say $P_2 = \mathcal{O}$ (so $P_3 = P_1$);
- if both points are $\mathcal{O}$, it is trivial to show that $t_3 = 0$ and equally $P_3 = \mathcal{O}$.

This completes our proof that ψ is a group isomorphism between $E_{ns}(K)$ and K.

□

Remark 3.5.2. To further convince yourself that it is a group, we must ensure that no sum of two points in $E_{ns}(K)$ yields $(0, 0)$. We previously mentioned that a line through $(0, 0)$ has at most one other intersection point with the curve, and in fact, a line through two K-smooth points cannot pass through the origin. The sum of two K-smooth points on E is never $(0, 0)$.

Node

In the second case, assume without loss of generality that the double root is at $(0,0)$ and the curve E has the equation $y^2 = x^2 \cdot (x+a)$ for $a \neq 0$ (we can always translate the singular point left and right). Again, the point $(0,0)$ is the only singularity on E and we choose to exclude it for the same reasons as in the previous case. We follow [3, Theorem 2.31] as we show that the group law becomes a multiplication of elements in $K^\times$ or in an extension of K.

Let $\alpha^2 = a$ so that α might lie in an extension of K, and let us define the parameter $t \in K$ such that for all points $(x,y) \in E_{ns}(K)$,

$$\begin{cases} t = 1, & \text{if } (x,y) = \mathcal{O} \\ t = \dfrac{y + \alpha \cdot x}{y - \alpha \cdot x}, & \text{otherwise.} \end{cases}$$

Theorem 3.5.3. Let E be the curve $y^2 = x^2 \cdot (x+a)$ and let $E_{ns}(K)$ be the K-smooth points on this curve. The map:

$$\begin{aligned} \psi : E_{ns}(K) &\to K^\times \\ (x,y) &\mapsto t \end{aligned}$$

is a group isomorphism between $E_{ns}(K)$ and $K^\times$, regarded as an multiplicative group if $\alpha \in K$. If $\alpha \notin K$, then ψ gives an isomorphism

$$E_{ns}(K) \cong \{u + \alpha \cdot v \mid u, v \in K, u^2 - a \cdot v^2 = 1\},$$

where the right-hand side is a group under multiplication.

Proof. Using the equation for E, we have that $x + a = (y/x)^2$. Further, we can obtain y/x from the parameter t:

$$\frac{y}{x} = \alpha \cdot \frac{t+1}{t-1}.$$

From this we can rewrite $x = \dfrac{4 \cdot \alpha^2 \cdot t}{(t-1)^2}$ and $y = \dfrac{4 \cdot \alpha^3 \cdot t \cdot (t+1)}{(t-1)^3}$ to express all points $(x,y) \in E_{ns}(K)$ in terms of the parameter t provided that $\alpha \in K$. It follows that ψ is a bijection.

If $\alpha \notin K$, multiply the numerator and denominator of t by $y + \alpha \cdot x$ to obtain

$$t = \frac{y + \alpha \cdot x}{y - \alpha \cdot x} = u + \alpha \cdot v$$

for $u, v \in K$. Similarly, if we multiply the numerator and denominator of t^{-1} by $y - \alpha \cdot x$,

we obtain $t^{-1} = u - \alpha \cdot v$. Then:

$$u^2 - a \cdot v^2 = (u - \alpha \cdot v) \cdot (u + \alpha \cdot v) = t \cdot t^{-1} = 1 \in K.$$

Suppose now that $u^2 - a \cdot v^2 = 1 \in K$ and let

$$x = \left(\frac{u+1}{v}\right)^2 - a, \qquad y = \left(\frac{u+1}{v}\right) \cdot x.$$

Then, $(x, y) \in E_{ns}(K)$, and we can check that $\psi(x, y) = u + \alpha \cdot v$. It follows that ψ is surjective and also a bijection.

It remains for us to prove that $P_1 + P_2 = P_3$ if, and only if, $t_1 \cdot t_2 = t_3$ where $P_i = (x_i, y_i) \in E_{ns}(K)$ and $t_i = \dfrac{y_i + \alpha \cdot x_i}{y_i - \alpha \cdot x_i} \in K$ for $i = 1, 2, 3$. Using the addition laws, substituting in what we found above for x and y in terms of t and simplifying:

- if P_1 and P_2 are distinct,

$$\begin{cases} \dfrac{4 \cdot \alpha^2 \cdot t_3}{(t_3 - 1)^2} = \dfrac{4 \cdot \alpha^2 \cdot t_1 \cdot t_2}{(t_1 \cdot t_2 - 1)^2} \\ \dfrac{4 \cdot \alpha^3 \cdot t_3 \cdot (t_3 + 1)}{(t_3 - 1)^3} = \dfrac{4 \cdot \alpha^3 \cdot t_1 \cdot t_2 \cdot (t_1 \cdot t_2 + 1)}{(t_1 \cdot t_2 - 1)^3} \end{cases} \implies t_3 = t_1 \cdot t_2;$$

- if $P_1 = P_2$,

$$\begin{cases} \dfrac{4 \cdot \alpha^2 \cdot t_3}{(t_3 - 1)^2} = \dfrac{4 \cdot \alpha^2 \cdot t_1^2}{(t_1^2 - 1)^2} \\ \dfrac{4 \cdot \alpha^3 \cdot t_3 \cdot (t_3 + 1)}{(t_3 - 1)^3} = \dfrac{4 \cdot \alpha^3 \cdot t_1^2 \cdot (t_1^2 + 1)}{(t_1^2 - 1)^3} \end{cases} \implies t_3 = t_1^2;$$

- if either point is $\mathcal{O}$, it is trivial to show that $t_3 = t_2$ if say $P_1 = \mathcal{O}$ since $P_3 = P_2$, or reciprocally, that $t_3 = t_1$ if say $P_2 = \mathcal{O}$ since $P_3 = P_1$;
- if both points are $\mathcal{O}$, then $P_3 = \mathcal{O}$ and equally $t_3 = 1$.

This completes our proof.

□

We will make use of these group laws when we look into what makes cryptography on singular elliptic curves insecure (Remark 2.2.6).

Chapter 4

Finite Fields

> Our goal throughout has been to illuminate the coherence and the beauty of the arithmetic theory of elliptic curves; we happily leave the task of being encyclopedic to the authors of more advanced monographs.
>
> J. H. Silverman, J. Tate — Preface of *Rational Points on Elliptic Curves* (1992)

So far, the real numbers have proved very useful in visualising elliptic curves and their group law. In practice, however, point addition over $\mathbb{R}$ is slow and inaccurate due to computer-imposed limitations like rounding errors. In this chapter, we focus on elliptic curves over **finite fields** as a starting point for the cryptographic applications we discuss in Chapter 5.

4.1 Revision

In this section, we take some time to recollect all that we know about **finite fields**.

Definition 4.1.1 (Finite field)**.** A **finite field** is a field that contains a finite number of elements. For a prime number p and a positive integer k, we denote the finite field of **characteristic** $q = p^k$ as $\mathbb{F}_q$.

Recall from Honours Algebra that the **algebraic closure** of a field K is an algebraic extension of that field for which every non-constant polynomial in $K[x]$ (the univariate polynomial ring with coefficients in K) has a root in K. We give here the algebraic closure of $\mathbb{F}_q$, leaving the proof to [13, Secrion 2.2]. It will come in handy in the coming sections.

Proposition 4.1.2. The **algebraic closure** of $\mathbb{F}_q$, denoted $\overline{\mathbb{F}_q}$, is the union

$$\bigcup_{k=1}^{\infty} \mathbb{F}_{q^k}.$$

The simplest examples of finite fields are those of prime characteristic. For a prime number p, the finite field $\mathbb{F}_p$ can be constructed as the ring of **integers modulo** p, denoted $\mathbb{Z}/p\mathbb{Z}$. These will be our main focus when it comes to the elliptic curve group law.

Definition 4.1.3 (Integers modulo n)**.** The ring $\mathbb{Z}/n\mathbb{Z}$ of **integers modulo** n is the set of equivalence classes of integers modulo n. It is endowed with its natural ring structure:

$$(a \pmod n) + (b \pmod n) = (a+b) \pmod n$$
$$(a \pmod n) \cdot (b \pmod n) = (a \cdot b) \pmod n.$$

You might recall that every additive group $\mathbb{Z}/n\mathbb{Z}$ is isomorphic to the finite cyclic group of order n [14]. The next theorem follows simply from Theorem 3.1.4 and is formally proven by [3, Theorem 4.1].

Theorem 4.1.4. Let E be an elliptic curve over the finite field $\mathbb{F}_q$. Then $E(\mathbb{F}_q) \cong \mathbb{Z}/n\mathbb{Z}$ or $\mathbb{Z}/n_1\mathbb{Z} \oplus \mathbb{Z}/n_2\mathbb{Z}$ for some integer $n \geq 1$, or for some integers $n_1, n_2 \geq 1$ with $n1 \mid n_2$.

Finally, we define **invertibility** in $\mathbb{Z}/n\mathbb{Z}$.

Proposition 4.1.5. For all $a \in \mathbb{Z}/n\mathbb{Z}$ there exists an element $b \in \mathbb{Z}/n\mathbb{Z}$ such that

$$a \cdot b \equiv b \cdot a \equiv 1 \pmod n$$

if, and only if, a and n are coprime.

If b exists, we say that a is **invertible** and call b the **modular inverse** of a modulo n and denote it as a^{-1}.

Proof. ($\implies$) Suppose that $a \cdot b \equiv b \cdot a \equiv 1 \pmod n$ for some $a, b \in \mathbb{Z}/n\mathbb{Z}$ and integer n. We can rewrite this as the following equation: $a \cdot b + q \cdot k = 1$ for some integer k. Using Bézout's Identity, it follows that $\gcd(a, n) | 1$ and since the greatest common divisor of two numbers is strictly positive, we find that $\gcd(a, n) = 1$, i.e. a and n are coprime.

($\impliedby$) Suppose now that a and n are coprime. Then using Bézout's Identity, it follows that there exists integers b, k such that $a \cdot b + n \cdot k = 1$. Rewriting this as a congruence equation and using the fact that multiplication is commutative yields $a \cdot b \equiv b \cdot a \equiv 1 \pmod n$.

We have shown both directions of the double implication which concludes our proof.

□

When n is not prime however, $\mathbb{Z}/n\mathbb{Z}$ fails to be a field since its elements are not all invertible. Indeed, any non-trivial divisors of n are not invertible. We will see an example of this in the next section as we look at the group law over finite fields.

4.2 Addition over $\mathbb{F}_p$

We define an abelian group on elliptic curves over finite fields of prime characteristic much as we did in the previous chapter. Note however that we denote x/y to mean $x \cdot y^{-1}$ in $\mathbb{F}_p$ where y^{-1} is the modular inverse of y. Let us look at some examples.

Example 4.2.1. Consider the elliptic curve $E : y^2 = x^3 - x$ over $\mathbb{F}_5$. depicted in Figure 4.1.

To determine the order of $E(\mathbb{F}_5)$, we to list all possible values of $x \in \mathbb{F}_5$ and compute $x^3 - x \pmod 5$, then we find the square roots of $x^3 - x$ in $\mathbb{F}_5$ which gives points in $E(\mathbb{F}_5)$ as shown in Table 4.1. Therefore, $\#E(\mathbb{F}_q) = 8$.

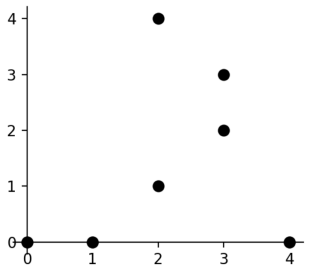


Figure 4.1: Plot of $E : y^2 = x^3 - x$ over $\mathbb{Z}/5\mathbb{Z}$

x	$x^3 - x$	y	(x, y)
0	0	0	$(0, 0)$
1	0	0	$(1, 0)$
2	1	± 1	$(2, 1), (2, 4)$
3	4	± 2	$(3, 2), (3, 3)$
4	0	0	$(4, 0)$
$\mathcal{O}$	$\mathcal{O}$	$\mathcal{O}$	$\mathcal{O}$

Table 4.1: Points on $E(\mathbb{F}_5)$

Let us demonstrate the addition of points on over $\mathbb{F}_q$ using the addition laws defined in Section 3.2. Taking $P_1 = (4, 0)$ and $P_2 = (2, 1)$, we want to add $P_1 + P_2 = P_3$ on $E(\mathbb{F}_5)$. We start by computing the slope of the line through the two points:

$$s = \frac{1-0}{2-4} = \frac{1}{-2} \equiv \frac{1}{3} \equiv 2 \cdot 3^{-1} \equiv 1 \cdot 2 \equiv 2 \pmod 5,$$

where the modular inverse of 3 is 2 modulo 5 since $3 \cdot 2 \equiv 1 \pmod 5$.

Hence the line through the two points is given by

$$y = s \cdot (x - 4) \equiv 2 \cdot x + 2 \pmod 5.$$

We substitute into the equation for E to obtain

$$(2 \cdot x + 2)^2 = x^3 - x \implies 0 = x^3 - 4 \cdot x^2 + \cdot x - 4.$$

We know the roots 2 and 4, therefore the remaining root is $x \equiv 3$, and since $y \equiv 2 \cdot x + 2 \pmod 5$, we have $y \equiv 3$. Finally, we reflect across the x-axis to find $P_3 = (3, 2) \in E(\mathbb{F}_q)$. Note that we could have used the algebraic formulas directly.

A little calculation shows that $(0, 0), (1, 0), (4, 0)$ have order 2 and the remaining four points have order 4 (note that $\mathcal{O}$ always has order 1), it follows that the group is isomorphic to $\mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/4\mathbb{Z}$.

Example 4.2.2. Suppose we now consider the elliptic curve $E : y^2 = x^3 - x$ over $\mathbb{Z}/25\mathbb{Z}$, where $25 = 5^2$, depicted in Figure 4.2, and use the same addition laws as above.

Suppose we want to add $P_1 = (2, 9)$ and $P_2 = (7, 6)$ on $E(\mathbb{F}_{25})$. As before, we start by computing the slope of the line through the two points

$$s = \frac{6 - 9}{7 - 2} = \frac{-3}{5}.$$

Notice that the denominator 5 is neither zero nor invertible in $\mathbb{F}_{25}$ since it divides the characteristic 25: the slope is neither infinite nor finite. But we already have $P_1 + (-P_1) = (2, 9) + (2, 16) \equiv \mathcal{O} \pmod{25}$, so we cannot also have $P_1 + P_2 \equiv \mathcal{O} \pmod{25}$. It must therefore be undefined.

As mentioned at the end of the previous section, this is because $\mathbb{Z}/n\mathbb{Z}$ is not a field when n is composite and since we defined our addition and group law over a field, we find ourselves with some undefined values. This is at the heart of the **Lenstra's Prime Factorization** we will study in Chapter 5.

4.3 Rational Points

Rational points can mean a great many things. We previously defined them in Section 2.4 as points whose coordinates (x, y) belong to a given field. In the case of finite fields, there are finitely many such pairs. As a result, an elliptic curve over a finite field $E(\mathbb{F}_q)$ has finitely many points within that finite field [3, p.95], we these the **rational points**.

The **group order** is the number of rational points on $E(\mathbb{F}_q)$ denoted $\#E(\mathbb{F}_q)$. It is considered to be an arithmetic quantity of great significance [15, Chapter V, Introduction]. In

elliptic curve cryptography, notably, the order helps us judge the difficulty of solving the **logarithm problem** in $E(\mathbb{F}_q)$. This brings us to the following theorem to which we dedicate the next section.

Theorem 4.3.1 (Hasse Bound)**.** Let E be an elliptic curve over $\mathbb{F}_q$, then

$$q + 1 - 2 \cdot \sqrt{q} \leq \#E(\mathbb{F}_q) \leq q + 1 + 2 \cdot \sqrt{q}. \tag{4.1}$$

Despite knowing that elliptic curves have a finite number of rational points, how do we go about finding the group order. By providing a bound, the Hasse Bound is a great leap towards finding out $\#E(\mathbb{F}_q)$.

Incidentally, the most commonly used algorithm to compute the order of a finite group is known as the **Baby Step, Giant Step** method [3, Section 4.3.4]. However, it is not very efficient when applied to elliptic curves over large finite fields. In 1985, Schoof introduced the first polynomial-time algorithm [15, Algorithm 3.1], which proved a lot more efficient in comparison to previously existing algorithms. It comes as no surprise that Hasse's theorem plays an instrumental part in both these algorithms. We sketch here the basic idea of **Schoof's algorithm**. For a more detailed explanation see [16].

Algorithm 4.3.2 (Schoof's algorithm)**.** Let $E : y^2 = x^3 + a \cdot x + b$ over $\mathbb{F}_q$ and set $a = q + 1 - \#E(\mathbb{F}_q)$. Hasse's theorem tells us that $\mid a \mid \leq 2 \cdot \sqrt{q}$. Let $S = \{2, 3, 5, 7, \cdots, L\}$ be the smallest set of primes such that

$$N = \prod_{l \in S} l > 4 \cdot \sqrt{q}.$$

We compute $a \pmod{l}$ for each prime $l \in S$ using the **Frobenius endomorphism** (more on this later) in a ring R defined in terms of division polynomials. For more details on this, see [3, Section 4.5].

Finally, we use the **Chinese Remainder Theorem** [17, Theorem 3.10] and compute

$$a \pmod{N}$$

to uniquely determine a using Hasse's Bound.

Still today, "extensions of Schoof's algorithm remain the point-counting method of choice when the characteristic of $\mathbb{F}_q$ is large (for example, when q is a cryptographic size prime)" [18].

4.4 Hasse Bound

As promised, we devote this section to the proof of the Hasse Bound. The Hasse Bound was originally conjectured in 1924 by Emil Artin in his thesis [19] and eventually proven by Hasse in 1933. His proof was published shortly after in a series of papers [20]. It is more commonly known as the **Hasse-Weil Bound** which is in fact a generalization of Hasse's theorem to algebraic curves of higher genera [21]. It is a deep result which finds a wide range of applications in mathematics but also in elliptic-curve cryptography as we will see in Chapter 5.

We start by introducing the **Frobenius map** as we build our proof from the ground up.

Definition 4.4.1 (Frobenius map)**.** We define the **Frobenius map** for $\mathbb{F}_q$ as

$$\begin{aligned} \phi_q : \overline{\mathbb{F}_q} &\to \overline{\mathbb{F}_q} \\ x &\mapsto x^q, \end{aligned}$$

where $\overline{\mathbb{F}_q}$ is the **algebraic closure** of $\mathbb{F}_q$.

Let E be an elliptic curve over $\mathbb{F}_q$, then the map ϕ_q acts on the coordinates of points in $E(\overline{\mathbb{F}_q})$ in the following way:

$$\phi_q(x, y) = (x^q, y^q), \qquad \phi_q(\mathcal{O}) = \mathcal{O}.$$

The Frobenius map is commonly called the **Frobenius endomorphism** of E, and for good reason. We give the definition of **endomorphism** from [3, Section 2.9, p.50].

Definition 4.4.2 (Endomorphism)**.** By an **endomorphism** of E, we mean a homomorphism

$$\alpha : E(\overline{K}) \to E(\overline{K})$$

that is given by rational functions. In other words, for all $P_1, P_2 \in E(\overline{K})$,

$$\alpha(P_1 + P_2) = \alpha(P_1) + \alpha(P_2),$$

and there are rational functions (quotients of polynomials), $R_1(x, y)$ and $R_2(x, y)$ with coefficients in K such that

$$\alpha(x, y) = (R_1(x, y), R_2(x, y))$$

for all $(x, y) \in E(\overline{K})$.

In particular, we have that $y^2 = x^3 + a \cdot x + b$ for all $(x, y) \in E(\overline{\mathbb{F}_q})$ and so we can rewrite any arbitrary endomorphism α defined as above as

$$\alpha(x, y) = (r_1(x), r_2(x) \cdot y),$$

with rational functions $r_1(x), r_2(x)$.

Let $r_1 = p(x)/q(x)$. We define the **degree** of α as:

$$\deg(\alpha) = \begin{cases} 0, & \text{if } \alpha = 0; \\ \max\big\{ \deg(p(x)), \deg(q(x)) \big\}, & \text{otherwise.} \end{cases}$$

We call α a **separable endomorphism** (conversely **inseparable**) if the derivative r_1' is not identically zero.

So far, we have claimed that the Frobenius map is an endomorphism of E, let us now prove it.

Lemma 4.4.3. Let E be an elliptic curve defined over $\mathbb{F}_q$, then ϕ_q is an inseparable endomorphism of E of degree q.

Proof. We want to show that $\phi_q : E(\overline{\mathbb{F}_q}) \to E(\overline{\mathbb{F}_q})$, equipped with our standard point addition, is a homomorphism. We let $P_1 = (x_1, y_1), P_2 = (x_2, y_2) \in E(\overline{\mathbb{F}_q})$ and prove by cases.

- Suppose first that P_1 and P_2 are distinct points.

  If $x_1 \neq x_2$. The sum $P_3 = (x_3, y_3) = P_1 + P_2$ is given by the addition laws:

  $$\begin{cases} x_3 = s^2 - x_1 - x_2, \\ y_3 = s \cdot (x_1 - x_3) - y_1. \end{cases}$$

  Applying the Frobenius map $\phi : (x, y) \mapsto (x^q, y^q)$ in $\overline{\mathbb{F}_q}$, our point addition becomes:

  $$\begin{cases} x_3^q = s_q^2 - x_1^q - x_2^q, \\ y_3^q = s_q \cdot (x_1^q - x_3^q) - y_1^q, \end{cases}$$

  where $s_q = \dfrac{y_2^q - y_1^q}{x_2^q - x_1^q}$. It is easy to see that $\phi(x_3, y_3) = \phi(x_1, y_1) + \phi(x_2, y_2)$.

If $x_1 = x_2$ and $y_1 \neq y_2$ (P_1 and P_2 are still distinct). By the addition laws, we have that $P_3 = \mathcal{O} = P_1 + P_2$. Now applying the Frobenius map, we get that $\phi(x_3, y_3) = \phi(x_1, y_1) + \phi(x_2, y_2) = \mathcal{O}$ much in the same way since $x_1^q = x_2^q$.

Finally, let us look at the case when either point is $\mathcal{O}$. Without loss of generality, take $P_1 = \mathcal{O}$. Then by the addition laws we get that $P_3 = \mathcal{O} + P_2 = P_2$. Applying the Frobenius map we find

$$\phi_q(P_3) = \phi_q(P_2) = \mathcal{O} + \phi_q(P_2) = \phi_q(\mathcal{O}) + \phi_q(P_2) = \phi_q(P_1) + \phi_q(P_2).$$

- Suppose now that P_1 and P_2 are coincident. The addition laws state that the sum $P_3 = (x_3, y_3) = 2 \cdot P_1$ is as above, with slope $s = \dfrac{3 \cdot x_1^2 + a}{2 \cdot y_1}$. Raising s to the q^{th} power yields:

$$s_q = \frac{3^q \cdot (x_1^q)^2 + a^q}{2^q \cdot y_1^q}.$$

Since $2, 3, a \in \mathbb{F}_q$, it follows from Fermat's Little Theorem [22] that

$$2^q = 2, \qquad 3^q = 3, \qquad A^q = A, \qquad \text{and} \qquad s_q = \frac{3 \cdot (x_1^q)^2 + a}{2 \cdot y_1^q}.$$

It follows that $\phi(x_3, y_3) = 2 \cdot \phi(x_1, y_1)$.

Thus, ϕ_q is a homomorphism of rational functions, from which we can conclude that it is an endomorphism of E.

Taking $r_1(x) = x^q$, we can easily see that ϕ_q has degree q. Finally, the derivative $r_1'(x) = q \cdot x^{q-1} = 0$ since $q = 0$ in characteristic q. Thus, we find that ϕ_q is inseparable which concludes our proof.

□

To keep our proof as short as possible, we only claim the following result which is proved in [3, Proposition 2.29].

Proposition 4.4.4. Let E be an elliptic curve defined over $\mathbb{F}_q$ and let r, s be integers not both zero. The endomorphism $r \cdot \phi_q + s$ is separable if, and only if, $p \nmid s$.

Given that $(E(\mathbb{F}_q), +)$ is an abelian group, the sum of two endomorphisms is an endomorphism, and multiplication by the integer -1 is an endomorphism [23]. In particular, $\phi_q - 1$ is an endomorphism given by

$$(\phi_q - 1)(x, y) = \phi_q(x, y) + (-1)(x, y).$$

It follows from Proposition 4.2.5 that $\phi_q - 1$ is a separable endomorphism.

The following result is a nice property of finite fields which will be needed in the next proof.

Lemma 4.4.5. For $a, b \in \mathbb{F}_q$ and any positive integer k, $(a+b)^{q^k} = a^{q^k} + b^{q^k}$.

Proof. Let $a, b \in \mathbb{F}_q$ and let k be a positive integer. Using the binomial expansion, we have that

$$(a+b)^{q^k} = \sum_{n=0}^{q^k} \binom{q^k}{n} \cdot a^n \cdot b^{q^k - n}.$$

Now, recall that for all n, the binomial coefficient is given by:

$$\binom{q^k}{n} = \frac{q^k!}{n! \cdot (q^k - n)!}.$$

Note that for $n = 1, \cdots, q^k - 1$, the binomial coefficient are divisible by q^k. As a result, they are thus zero in $\mathbb{F}_q$ and we are then left with, as claimed,

$$(a+b)^{q^k} = \binom{q^k}{0} \cdot b^{q^k} + \binom{q^k}{q^k} \cdot a^{q^k} = a^{q^k} + b^{q^k}.$$

□

Building on, the following lemma ties in the notion of cardinality and lies at the heart of the proof of Hasse's theorem.

Lemma 4.4.6. Let E be an elliptic curve defined over $\mathbb{F}_q$, then $\#E(\mathbb{F}_q) = \deg(\phi_q - 1)$.

Proof. Note that $(x, y) \in E(\mathbb{F}_q)$ if, and only if, $\phi_q(x, y) = (x, y)$. This is straightforward enough using the equation for E and Lemma 4.5.5. Finally $a^q = a$ when $a \in \mathbb{F}_q$.

Now,

$$\begin{aligned}
(x, y) \in E(\mathbb{F}_q) &\iff \phi_q(x, y) = (x, y) \\
&\iff \phi_q(x, y) - (x, y) = 0 \\
&\iff (\phi_q - 1)(x, y) = 0 \\
&\iff (x, y) \in \mathrm{Ker}(\phi_q - 1).
\end{aligned}$$

Thus, $\#E(\mathbb{F}_q) = \#\mathrm{Ker}(\phi_q - 1)$. And since $\phi_q - 1$ is separable, it follows from [3, Proposition 2.21] that $\deg(\phi_q - 1) = \#\mathrm{Ker}(\phi_q - 1)$ which concludes our proof.

□

Finally, we will need the following version of the **Cauchy-Schwarz inequality** [15, Chapter V, Lemma 1.2].

Lemma 4.4.7 (Cauchy-Schwarz inequality)**.** Let A be an abelian group, and let $d : A \to \mathbb{Z}$ be a positive definite quadratic form. Then

$$| \, d(a-b) - d(a) - d(b) \, | \leq 2 \cdot \sqrt{d(a) \cdot d(b)}$$

for all $a, b \in A$.

The **degree map** $\deg : \mathrm{End(E)} \to \mathbb{Z}$ first described in Definition 4.2.3 is an example of positive definite quadratic form. This is quite straightforward so we leave the proof to [15, Chapter III, Corollary 6.3]. We are now ready to prove the Hasse Bound.

Proof. Setting $a = \phi_q$, $b = 1$ and d to the degree map in the Cauchy-Schwarz inequality given above, we find

$$| \, \deg(\phi_q - 1) - \deg(\phi_q) - \deg(1) \, | \leq 2 \cdot \sqrt{\deg(\phi_q) \cdot \deg(1)}$$

Using the fact that $\deg(\phi_q) = q$ and $\deg(1) = 1$, as well as Lemma 4.2.5, we write

$$| \, \#E(\mathrm{F}_q) - q - 1 \, | \leq 2 \cdot \sqrt{q},$$

which is just another way of writing the Hasse Bound given in (4.1).

□

It all seems beautifully short and simple, but let us not forget all the work that has gone into this proof. We finish this section with an illustration of the theorem in an example.

Example 4.4.8. Picking a larger prime, consider the elliptic curve $E : y^2 = x^3 - x$ over $\mathbb{F}_{37}$. The Hasse Bound yields

$$25 < 38 - 2 \cdot \sqrt{37} \leq \#E(\mathbb{F}_{37}) \leq 38 + 2 \cdot \sqrt{37} < 51.$$

Looking at Figure 4.2, or counting the points in Python, we find that in fact $\#E(\mathbb{F}_{37}) = 40$.

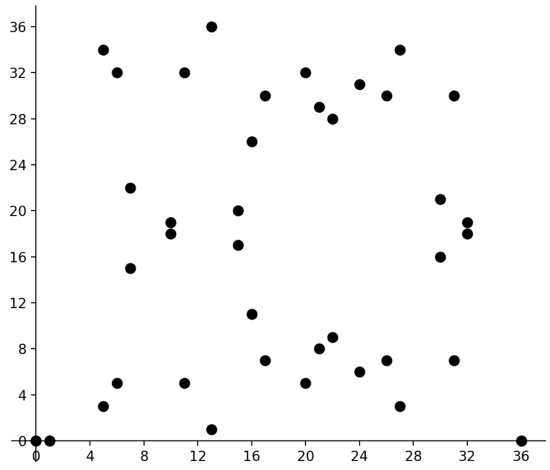


Figure 4.2: Plot of $E : y^2 = x^3 - x$ over $\mathbb{F}_{37}$

4.5 Reduction Modulo p

Given an elliptic curve $E : y^2 = x^3 + a \cdot x + b$ over the rational numbers $\mathbb{Q}$ with $a, b \in \mathbb{Z}$, we can reduce its coefficients modulo a fixed prime p to obtain a (possibly singular) curve over $\mathbb{F}_p$. We write:

$$\overline{E} : y^2 = x^3 + \overline{a} \cdot x + \overline{b},$$

where $\overline{a}, \overline{b} \in \mathbb{F}_p$ and $\overline{E}$ is the **reduction of** E **modulo** p or **reduced curve**.

We now ask: can we establish a mapping between the points on $E(\mathbb{Q})$ and the points on $\overline{E}(\mathbb{F}_p)$ by reducing the points modulo p? In this section, we introduce and formalise an interesting application of the theory we have developed so far: the **modulo-p reduction map** for any fixed prime p.

Remark 4.5.1. To keep things short, notice that we have introduced the following notation: for all $x \in \mathbb{Z}$ we write $\overline{x}$ to mean $x \pmod{p}$ (when the characteristic we are working in is obvious).

Points of Finite Order

Our first task is to define the reduction map. Having reduced E modulo p, it is natural to try taking points in $E(\mathbb{Q})$ and reducing them modulo p to get points on $\overline{E}$. We can do this provided that the coordinates of the point have their denominators coprime with p. Otherwise, they are not invertible (similar to Example 4.2.2).

Trivially, if a point $P = (x, y) \in E(\mathbb{Q})$ has integer coordinates x, y, then P can be reduced modulo p to a point $\overline{P} = (\overline{x}, \overline{y})$ that lies on $\overline{E}$. Thus, for all $(x, y) \in E(\mathbb{Q})$ with $x, y \in \mathbb{Z}$, we have:

$$(x, y) \mapsto (\overline{x}, \overline{y}) \in \overline{E}(\mathbb{F}_q).$$

With this in mind, consider the following theorem. Note that while the proof of **Nagell-Lutz theorem** is not in the scope of this report, it can be found in [11, Theorem 2.5].

Theorem 4.5.2 (Nagell-Lutz theorem)**.** Let $E : y^2 = x^3 + a \cdot x + b$ be a non-singular elliptic curve with integer coefficients a, b and discriminant Δ. Let $P = (x, y)$ be a rational point of finite order, then x and y are integers, and

- if $y = 0$, then P has order two;
- if $y \neq 0$, then $y \mid \Delta$.

The Nagell-Lutz theorem tells us that all points of finite order on $E(\mathbb{Q})$ have integer coordinates (except for $\mathcal{O}$), giving us a correspondence between the points of finite order on $E(\mathbb{Q})$, and a subgroup of $E(\mathbb{F}_p)$ generated by these points. We denote $E_{tor}(\mathbb{Q})$ the collection of points of finite order on $E(\mathbb{Q})$, also called the **torsion subgroup of** $E(\mathbb{Q})$,

$$E_{tor}(\mathbb{Q}) = \big\{P = (x, y) \in E(\mathbb{Q} \mid P \text{ has finite order}\big\} \cup \big\{\mathcal{O}\big\}.$$

We are now ready to define our **reduction map** following [11, p.134]:

Definition 4.5.3 (Reduction Map)**.** Let $E : y^2 = x^3 + a \cdot x + b$ be an elliptic curve over $\mathbb{Q}$ with $a, b \in \mathbb{Z}$. For every point of **finite order** $P = (x, y) \in E(\mathbb{Q})$, we define the **reduction map modulo** p as

$$\begin{aligned} r_p : E_{tor}(\mathbb{Q}) &\to \overline{E}(\mathbb{F}_p) \\ P \mapsto \overline{P} &= \begin{cases} \overline{\mathcal{O}}, & \text{if } P = \mathcal{O} \\ (\overline{x}, \overline{y}), & \text{otherwise} \end{cases} \end{aligned}$$

where $\overline{E}(\mathbb{F}_p)$ is the **modulo** p **reduced curve** of E.

Singularity

When is $\overline{E}$ non-singular? We give two conditions. First, recall that an elliptic curve is non-singular if it has three distinct roots. If $p < 3$, we do not even have three distinct elements

in $\mathbb{F}_p$, let alone three distinct roots of $\overline{E}$, so it must be that $p \geq 3$.

Further, we previously defined the discriminant $\Delta = 4 \cdot a^3 + 27 \cdot b^2$ and proved that $\Delta \neq 0$ was necessary and sufficient condition for E to be non-singular. Note that since a, b are chosen to be integers, we have that $\Delta \in \mathbb{Z}$, which means that we can easily reduce it modulo p. It follows that:

Proposition 4.5.4. For a fixed prime $p \geq 3$, the reduced curve $\overline{E}$ is said to be non-singular if, and only if,

$$p \nmid \Delta,$$

When $\overline{E}$ is non-singular, we call it a **good reduction** (conversely a **bad reduction** when it is singular).

Homomorphism

It follows from Proposition 3.1.6 that $E_{tor}(\mathbb{Q})$ is a subgroup of $E(\mathbb{Q})$. This implies that it is a group itself. Additionally, provided that p does not divide $2 \cdot \Delta$, we know that $\overline{E}(\mathbb{F}_p)$ is also a group [11, p.135]. So r_p is a map between two groups, but is the group structure preserved in some way? That is, does it define a group homomorphism? In fact, yes.

Theorem 4.5.5 (Reduction Modulo p theorem)**.** If r_p is a good reduction, then it defines a group homomorphism from $E(\mathbb{Q})$ to $\overline{E}(\mathbb{F}_p)$.

As a preliminary step, we show that negativity is preserved under the mapping.

Lemma 4.5.6. Let $P = (x, y) \in E(\mathbb{Q})$, then the following holds:

$$r_p(-P) = -r_p(P).$$

Proof. For $P = \mathcal{O}$, this is trivial. Suppose then that $P \neq \mathcal{O}$. We prove

$$r_p(-P) = \overline{-P} = (\overline{x}, \overline{-y}) = (\overline{x}, -\overline{y}) = -\overline{P} = -r_p(P)$$

as claimed.

□

Now let us prove Theorem 4.5.5 following [11, Theorem 4.4].

Proof. We want to demonstrate that the map is a homomorphism. Let $P_1 = (x_1, y_1), P_2 = (x_2, y_2)$ and $P_3 = (x_3, y_3)$ be any three points in $E_{tor}(\mathbb{Q})$, it suffices for us to show that

$$P_1 + P_2 + P_3 = \mathcal{O} \implies r_p(P_1) + r_p(P_2) + r_p(P_3) = \overline{\mathcal{O}}.$$

Suppose that $P_1 + P_2 + P_3 = \mathcal{O}$.

- If at least one of P_1, P_2, or P_3 is $\mathcal{O}$: without loss of generality, say $P_3 = \mathcal{O}$. Then $r_p(P_3) = \overline{\mathcal{O}}$ and $P_1 = -P_2$. Using Lemma 4.5.6, we find that $r_p(P_1) = -r_p(P_2)$ and so:

$$r_p(P_1) + r_p(P_2) + r_p(P_3) = -r_p(P_2) + r_p(P_2) + \overline{\mathcal{O}} = \overline{\mathcal{O}}.$$

- If P_1, P_2, and P_3 are all not equal to $\mathcal{O}$, then from the definition of our group law, they must be sitting on a single line $\mathcal{L}$. Let the equation of that line be given by

$$\mathcal{L} : y = \lambda \cdot x + \nu.$$

  The same group laws give us that the coordinates of P_3 are

$$x_3 = \lambda^2 - x_1 - x_2 \qquad \text{and} \qquad y_3 = \lambda \cdot x_3 + \nu.$$

  Note that because the coordinates x_i, y_i are integers for $i = 1, 2, 3$, we have $\lambda, \nu \in \mathbb{Z}$. This important fact means that we can reduce the coefficient of $\mathcal{L}$ modulo p.

  Substituting the equation of $\mathcal{L}$ into the equation of E yields

$$0 = x^3 + a \cdot x + b - (\lambda \cdot x + \nu)^2 = (x - x_1) \cdot (x - x_2) \cdot (x - x_3),$$

  since x_1, x_2, and x_3 are the three roots (P_1, P_2 and P_3 belong to both E and $\mathcal{L}$). Because the coefficients are all integers, we can reduce them modulo p to find

$$x^3 + \overline{a} \cdot x + \overline{b} - (\overline{\lambda} \cdot x + \overline{\nu})^2 = (x - \overline{x_1}) \cdot (x - \overline{x_2}) \cdot (x - \overline{x_3}),$$

  and of course reducing $y_i = \lambda \cdot x + \nu$, we find $\overline{y_i} = \overline{\lambda} \cdot \overline{x_i} + \overline{\nu}$, for $i = 1, 2, 3$.

  Thus, the line $y = \overline{\lambda} \cdot x + \overline{\nu}$ intersects $E(\mathbb{F}_p)$ at the points $r_p(P_1), r_p(P_2)$, and $r_p(P_3)$ which implies $r_p(P_1) + r_p(P_2) + r_p(P_3) = \overline{\mathcal{O}}$ as required.

This concludes our proof that r_p is a homomorphism.

□

We conclude this chapter with an example taken from [24] of the above theorem which ties in the notion of group order previously studied.

Example 4.5.7. Let $E : y^2 = x^3 + 3$ be an elliptic curve defined over the rationals $\mathbb{Q}$. We have $\Delta = -243 = -3^5$. Applying Theorem 4.5.5, we know that there is an injective homomorphism correspondence between $E_{tor}(\mathbb{Q})$ and the points in the reductions $\overline{E}(\mathbb{F}_5)$ and $\overline{E}(\mathbb{F}_7)$. **Lagrange's Theorem** [22, Theorem 26.1] requires that:

$$\#E_{tor}(\mathbb{Q}) \mid \#E(\mathbb{F}_5) \qquad \text{and} \qquad \#E_{tor}(\mathbb{Q}) \mid \#E(\mathbb{F}_7).$$

Now $\#E(\mathbb{F}_5) = 6$ and $\#E(\mathbb{F}_7) = 13$ (which we compute using our Python). Since they are coprime, we must have that $\#E_{tor}(\mathbb{Q}) = 1$, that is, $\mathcal{O}$ is the only finite order point in is $E(\mathbb{Q})$.

Chapter 5

Cryptography

> The problem of secure communication is probably almost as old as civilization itself. In sending a message to a distant correspondent, with adversaries somewhere along the way, it has always been necessary to ensure that the messages cannot be understood by the adversaries in the middle.
>
> Duncan A. Buell — *Fundamentals of Cryptography - Introducing Mathematical and Algorithmic Foundations* (2021)

In this chapter, we look at some of the applications of elliptic curves to Cryptography.

5.1 Background

Cryptography is the art and science of encrypting data so that no one can understand it save for the person it is intended to. We distinguish two main types of encryption: **symmetric** and **asymmetric-key encryption**.

In symmetric-key encryption, the **keys** used for encryption and decryption are similar. These systems are very fast and use relatively small keys. However, sharing the keys safely can prove difficult. On the other hand, asymmetric-key encryption, also known as **public-key encryption**, uses a pair of keys: a **public key** and **private key**. Everyone can encrypt data using the public key (it is made public), but only someone with the private key can decrypt it. These are more easily shareable.

Remark 5.1.1. "For our scenarios we suppose that A and B (also known as Alice and Bob) are two users of a public-key cryptosystem" [25]. This was the first mention of Alice (the sender) and Bob (the receiver) in secure transmission examples. With the paper's rising fame, the fictional characters, originally invented to make research in cryptography easier to understand, became without doubt the most famous cryptographic couple.

In this chapter, we will often talk about key size. This is because key size is often closely related to the **security level** of a cryptoscheme, but the bigger the key, the more computationally expensive it becomes. We define the security level as a number n in bits: n-bit security means that an attacker needs to perform 2^n operations to break a system. Introducing new methods which achieve the same level of security for smaller keys is the focus of much research in cryptography. This is where elliptic curve cryptography comes in.

5.2 Discrete Logarithm Problem

In this section we introduce the **discrete logarithm problem** as an entry-point for the applications of elliptic curve in cryptography.

Definition 5.2.1 (Discrete Logarithm)**.** Let $(G, \cdot)$ be a cyclic group with a known generator g. If the group is cyclic, then any element $a \in G$ can be written as some power g^n in G. Given such an element $a \in G$, an integer n that solves the equation $g^n = a$ is termed the **discrete logarithm** of a in G.

In some very few special cases, discrete logarithms can be computed quickly. However, in general, no method is known to compute them efficiently: it is considered to be a computationally intractable problem [26, Section 2.3]. This is what we call the **discrete logarithm problem** ("discrete" distinguishes the finite group situation from the classic continuous situation we omit here). In fact, several important asymmetric-key algorithms base their security on the assumption that the discrete logarithm problem has no efficient solution over carefully chosen groups. This is why the discrete logarithm problem is considered to be the "engine" of public-key cryptography.

What does it look like for elliptic curves? Let E be a non-singular elliptic curve over $\mathbb{F}_p$ and P be a point on that curve. Given a multiple Q of P, the discrete logarithm problem for elliptic curves consists of finding $n \in \mathbb{Z}$ such that

$$n \cdot P = \underbrace{P + P + \cdots + P}_{n \text{ times}} = Q.$$

Notice that we require E to be non-singular. This is because, as mentioned at the end of Section 2, singular curves are considered insecure for the discrete logarithm problem: the discrete logarithms are too easy to compute. We demonstrate it here.

Suppose E is singular, then E falls in one of the two cases we distinguished in Section 3.5 which are both insecure for cryptography:

- if $E : y^2 = x^3$ over $\mathbb{F}_p$, then ψ is an efficient map to the additive group of $\mathbb{F}_p$ where the discrete logarithm is trivial;
- if $E : y^2 = x^2 \cdot (x + a)$, then ψ is an efficient map to a multiplicative group of $F_p^\times$ or an extension of $\mathbb{F}_q$ where the discrete logarithm is a lot easier to compute (but not as easy as the previous case). We look at an example from [27].

Example 5.2.2. Let $p = 23981$. Consider the singular curve $E : y^2 = x^2 \cdot (x - 555)$ over $\mathbb{F}_p$ (check that the discriminant is zero). It has a singularity at the point $(0, 0)$. To determine $\alpha^2 = -555$, we compute the square root of -555 in characteristic p and we find $\alpha = 7020$. We can now define the map ψ as we did in Theorem 3.5.3.

Take $P = (1451, 1362)$ and $Q = (3141, 12767)$, two points on $E(\mathbb{F}_p)$. Now, with the help of Python, we perform the following operations modulo p:

$$\psi(P) = \frac{1362 + \alpha \cdot 1451}{1362 - \alpha \cdot 1451} = 19402 \cdot 4182 = 11441 \in \mathbb{F}_p,$$
$$\psi(Q) = \frac{3141 + \alpha \cdot 12767}{3141 - \alpha \cdot 12767} = 10484 \cdot 8669 = 21787 \in \mathbb{F}_p.$$

We want to know the discrete logarithm n such that $n \cdot P = Q$ which is hard to determine. Now instead, we can compute the discrete logarithm n such that $\psi(P)^n = \psi(Q)$ in $\mathbb{F}_p$ which is easier to do (see [28, IV, 3, p.102–103] for some algorithms to find discrete logs in finite fields). We find $n = 8279$.

5.3 Diffie-Hellman

A **key-exchange protocol**, also called **key-agreement protocol**, is a method of establishing a shared secret key by communicating solely over an insecure channel, without any previous private communication. The classic **Diffie-Hellman Key Exchange**, originally imagined by Ralph Merkle and named after Whitfield Diffie and Martin Hellman [29], was one of the first key-exchange protocols in 1976. It illustrates perfectly the discrete logarithm problem we described in the previous section.

Classic Diffie-Hellman

The original implementation of the protocol used the multiplicative group of integers modulo p, where p is prime and went like this:

Algorithm 5.3.1 (Classic Diffie-Hellman Key Exchange)**.** Let G be a group over a finite field $\mathbb{F}_p$ such that the discrete logarithm problem is hard (more on this later).

(1) Alice and Bob agree on a point $g \in G$ such that the subgroup generated by G has large order (usually, the curve and point are chosen so that the order is a large prime).

(2) Alice chooses a secret integer a, computes $A = g^a \pmod{p}$, and sends A to Bob. At the same time, Bob chooses a secret integer b, computes $B = g^b \pmod{p}$, and sends B to Alice.

(3) Alice computes the secret key $K = A^b = g^{a \cdot b} \pmod{p}$, and similarly, Bob computes $K = B^a = g^{b \cdot a} \pmod{p}$.

The security of the Diffie-Hellman protocol relies on the assumption that it is computationally infeasible for an attacker to compute the key $K = g^{a \cdot b}$ knowing only the eavesdropped values g^a and g^b. Indeed, recovering either a from g^a or b from g^b is equivalent to solving the discrete logarithm problem which we believe to be computationally hard as we discussed in the previous section.

Elliptic Curve Diffie-Hellman

You have to wait a decade before elliptic curves make their first appearance in public-key cryptography. In 1985, it was Victor Miller and Neal Koblitz who both suggested their use independently of each other in [30] and [31], and the idea of an **Elliptic Curve Diffie-Hellman** protocol emerged. While the algorithm is more or less the same the classic version, it makes novel use of the group structure on elliptic curve we defined in Section 3.2.

Algorithm 5.3.2 (Elliptic Curve Diffie-Hellman Key Exchange)**.** In first a instance, both parties need to agree on a non-singular elliptic curve E and a finite field $\mathbb{F}_p$ such that the discrete logarithm problem is hard (more on this later).

(1) Alice and Bob agree on a point $G \in E(\mathbb{F}_p)$ such that the subgroup generated by G has large order (usually, the curve and point are chosen so that the order is a large prime).

(2) Alice chooses a secret integer a, computes $G \cdot a = a \cdot G$, and sends $G \cdot a$ to Bob. On the other hand, Bob chooses a secret integer b, computes $G \cdot b = b \cdot G$, and sends $G \cdot b$ to Alice.

(3) Alice computes the secret key $K = a \cdot G \cdot b = a \cdot b \cdot G$, and similarly, Bob computes $K = b \cdot G \cdot a = b \cdot a \cdot G$.

In comparison to the classic Diffie-Hellman, Elliptic Curve Diffie Hellman can use smaller keys for the same level of security [3, Chapter 6]. This is because there is wide belief that the discrete logarithm problem is much harder to solve for elliptic curves. It easily follows that using elliptic curves is also much faster: in fact, its key and message sizes are 5 to 10 times smaller than those for other system [15, Chapter XI, Section 4, p.377].

5.4 Prime Factorisation

In this section we discuss a final application of elliptic curves to Cryptography: prime factorisation. Prime factorisation dates back to the ancient Greeks with Euclid's **Fundamental Theorem of Arithmetic** [32, Chapter VII, prop. 30–32] which states that every positive integer can be uniquely factorised (up to ordering) as a product of prime numbers. See for a proof [22, Theorem 11.1]. The implications of this theorem are as relevant today as they were then: indeed, many cryptographic protocols including the public-key **Rivest–Shamir–Adlema (RSA)** cryptosystem rely on the difficulty of factoring large composite integers for secure data transmission.

At the time of writing this, no **polynomial-time** algorithm has been found to factor all integers (of any size). By **polynomial-time** we mean that its running time T is upper bounded by a polynomial expression in the size n of the input. The existence (or non-existence) of such an algorithm has yet to be proven but it is generally accepted that such a proof does not exist, putting the problem in the **non-deterministic polynomial-time (NP) class**. [33, p. 203]

What does all this have to do with elliptic curves? In the late 1980s, the Dutch mathematician Hendrik Willem Lenstra published a paper outlining an efficient method for finding non-trivial factors of integers [34]. This was the first application of elliptic curves to cryptography, obtained as a generalisation of the classical **Pollard's** $(p-1)$ **method** that worked on multiplicative groups $\mathbb{Z}_n^\times$ [35].

Algorithm 5.4.1 (Lenstra's elliptic-curve factorisation)**.** Given a composite odd integer n, we wish to find a non-trivial divisor of $d \mid n$ such that $1 < d < n$.

(1) Choose a non-singular elliptic curve $E : y^2 = x^3 + a \cdot x + b$ over $\mathbb{Q}$ with $a, b \in \mathbb{Z}$. Pick a point $P = (x, y) \in E(\mathbb{Q})$.

(2) Let $d = \gcd(\Delta, n)$.

- If $1 < d < n$, then d is a non-trivial divisor of n and we are done.
- If $d = n$, then go back to step (1).
- Otherwise, continue to step (3).

(3) Take some bounds $B, C \in \mathbb{N}$, and let k the product of small powers of primes not exceeding B which are less than some bound C. That is, set

$$k = \prod_{l \leq B} l^{\alpha_l}$$

where l is prime and α_l is the largest exponent such that $l^{\alpha_l} \leq C$. B is picked to be small enough so that B-wise point addition can be performed in reasonable time.

(4) Attempt to compute $k \cdot P = \underbrace{P \cdot P \cdots P}_{k \text{ times}}$ in $\mathbb{Z}/n\mathbb{Z}$.

- If we finish the calculation without encountering non-invertible elements modulo n, go back to step (1) and choose a new (E, P) pair.
- If the calculation fails, we have a denominator d' which is not invertible modulo n. Setting $d = \gcd(d', n)$. If $d = n$ go back to step (1) and choose a new (E, P) pair. Otherwise, d is non-trivial divisor of n and we are done.

The intuition behind this algorithm is the following. Suppose we run into a difficulty when attempting to compute $k' \cdot P$ modulo n where $k' \cdot P$ is a partial sum encountered along the way of our computation of $k \cdot P$ in step (4). By the Addition Laws, this means that the denominator d' of the coordinates of $k' \cdot P$ is not invertible modulo n, that is, d' is a multiple of n. Setting $d = \gcd(n, d')|n$. Either d is a proper divisor of n, or it is n itself. Here is the formal statement of this.

Proposition 5.4.2. Let $E : y^2 = x^3 + a \cdot x + b$ be an elliptic curve over $\mathbb{Q}$ with $a, b \in \mathbb{Z}, n \in \mathbb{N}$ and $\gcd(\Delta, n) = 1$. Let $P_1, P_2 \in E(\mathbb{Q})$ with coordinates which have denominator prime to n, $P_1 \neq -P_2$. Then the point $P_1 + P_2$ has coordinates with denominator not prime to n if, and only if, there exists a prime $p \mid n$ such that $r_p(P_1 + P_2) = \overline{\mathcal{O}}$.

Note that this $p \neq 2$ because $p \mid n$, $\gcd(\Delta, n) = 1$ and $2 \mid \Delta$ by definition of the discriminant with $a, b \in \mathbb{Z}$.

Proof. We prove the claim following [36, Chapter I, 1.4].

($\Longleftarrow$) Let $P_1 = (x_1, y_1), P_2 = (x_2, y_2)$ and $P_1 + P_2 = (x_3, y_3)$ have coordinates whose denominators are prime to n. Suppose that there exists a prime $p \mid n$ with $r_p(P_1 + P_2) = \overline{\mathcal{O}}$. We will argue by contradiction.

Suppose first that $P_1 = P_2$. We then have that $P_1 + P_2 = 2 \cdot P_1$. By hypothesis, we find that $2 \cdot \overline{P_1} = \overline{\mathcal{O}}$. This can only be true if the tangent to E at P_1 is vertical, i.e. the slope of the tangent at $\overline{P_1}$ is undefined. The slope is given by implicit differentiation:

$$s = \frac{dy_1}{dx_1} = \frac{3 \cdot x_1^2 + a}{2 \cdot y_1}. \tag{5.1}$$

For s to be undefined mod p, we must have that $y_1 \equiv 0 \pmod{p}$. Using the equation of E and rearranging (5.1), it follows that $x_1^3 + a \cdot x + b \equiv 0 \pmod{p}$.

Now, using the addition laws defined above, we have that

$$\begin{aligned} x_3 &= s^2 - 2 \cdot x_1 \\ &= \frac{(3 \cdot x_1^2 + a)^2}{4 \cdot y_1^2} - 2 \cdot x_1 \\ &= \frac{(3 \cdot x_1^2 + a)^2 - 8 \cdot x_1 \cdot y_1^2}{4 \cdot y_1^2}. \end{aligned}$$

Note that by assumption, no divisor of n divides the denominator of x_3, in particular p does not divide the denominator of x_3. However, we see that the denominator of x_3 is a factor of y_1 and we know from above that $p \mid y_1$. It must then be that p also divides the numerator of x_3, i.e.

$$(3 \cdot x_1^2 + a)^2 - 8 \cdot x_1 \cdot y^2 \equiv (3 \cdot x_1^2 + a)^2 \equiv 0 \pmod{p}.$$

Using the fact that p is prime, we have that the derivative $3 \cdot x_1^2 + a \equiv 0 \pmod{p}$ also by Euclid's Lemma. Therefore, $\overline{P_1}$ is a singular point. Thus, E is singular mod p, which implies that $p \mid \Delta$. This leads to a contradiction since p now divides both Δ and n, which would imply that $\gcd(\Delta, n) \geq p$, yet we have $\gcd(\Delta, n) = 1$.

Suppose now that $P_1 \neq P_2$.

- Suppose first that $x_1 \not\equiv x_2 \pmod{p}$. Note that the denominators of the coordinates of $P_1 + P_2$ are both $x_2 - x_1$. By assumption, $x_2 - x_1 \not\equiv 0 \pmod{p}$. Thus, the coordinates of $r_p(P_1 + P_2)$ are defined and we get that $r_p(P_1 + P_2) \neq \overline{\mathcal{O}}$ which contradicts our assumptions. Hence if $r_p(P_1 + P_2) = \overline{\mathcal{O}} \pmod{p}$, we must have $\overline{x_1} = \overline{x_2}$.

- Suppose now that $\overline{x_1} = \overline{x_2}$. Note that by assumption, no divisor of n divides the denominator of x_3, in particular p does not divide the denominator of x_3. Using the addition laws, we have that:

$$x_3 = s^2 - x_1 - x_2 = \left(\frac{y_2 - y_1}{x_2 - x_1}\right)^2 - x_1 - x_2.$$

  By assumption, we have $p \mid x_2 - x_1$ which implies $p^2 \mid (x_2 - x_1)^2$. Multiply by x_3 to

find:

$$p^2 \mid (y_2 - y_1)^2 - (x_2 - x_1)^2 \cdot (x_2 + x_1) \implies p^2 \mid (y_2 - y_1)^2.$$

By Euclid's Lemma, we find $p \mid y_2 - y_1$, i.e. $\overline{y_1} = \overline{y_2}$. Hence, $\overline{P_1} = \overline{P_2}$.

As in the first part of our proof, we now have $2 \cdot \overline{P_1} = \overline{\mathcal{O}}$. This can only be true if the tangent to E at P_1 is vertical, i.e. the slope of the tangent at $\overline{P_1}$ is undefined. This can only be true for $\overline{y_1} = \overline{y_2}$. Substituting in the equation of E for y_1, we can conclude that $x_1^3 + a \cdot x + b \equiv 0 \pmod{p}$.

Now write $x_2 = x_1 + p^k \cdot x$ with $k \in \mathbb{N}$ such that neither the denominator nor the numerator of x is divisible by p. It is easy to see that then $y_2 = y_1 + p^k \cdot y$ is of the same form where y is chosen to have neither the denominator nor the numerator divisible by p also. Consider now the equation:

$$\begin{aligned} y_2^2 &= (x_1 + p^k \cdot x)^3 + a \cdot (x_1 + p^k \cdot x) + b \\ &\equiv x_1^3 + a \cdot x_1 + b + p^k \cdot (3 \cdot x \cdot x_1^2 + a \cdot x) \pmod{p} \\ &\equiv y_1^2 + p^k \cdot x \cdot (3 \cdot x_1^2 + a) \pmod{p}. \end{aligned}$$

This yields

$$y_2^2 - y_1^2 \equiv p^k \cdot x \cdot (3 \cdot x_1^2 + a) \pmod{p}.$$

As $y_2^2 - y_1^2 = (y_2 - y_1) \cdot (y_2 + y_1)$ is divisible by p^{k+1} and x is chosen to be not divisible by p, it follows that the derivative $(3 \cdot x_1^2 + a)$ must be divisible by p. Therefore $\overline{P_1}$ is a singular point and we can thus conclude that p divides the discriminant Δ which contradicts our hypothesis.

($\implies$) Let p be a prime divisor of n and suppose now that $r_p(P_1 + P_2) \neq \overline{\mathcal{O}}$. We want to prove the contrapositive, i.e. that in this case p does not divide the denominator of the coordinates of $P_1 + P_2$.

Suppose first that $\overline{x_1} \neq \overline{x_2}$. Then x_1, x_2 cannot both be divisible by p, if at all, so p does not divide the denominator of the coordinates of $P_1 + P_2$. This is trivial.

Suppose now that $\overline{x_1} = \overline{x_2}$. Looking at the equation of E, we can thus conclude that $\overline{y_1} = \pm\overline{y_2}$. Using our assumption that $P_1 \neq -P_2$, we cannot have $y_1 = -y_2 = -\overline{y_2}$, which implies that $\overline{y_1} = \overline{y_2}$. Now, using the fact that $r_p(P_1 + P_2) = 2 \cdot \overline{P_1} \neq \overline{\mathcal{O}}$, it follows that $\overline{y_2} = \overline{y_1} \neq 0$ or else the slope at $\overline{P_1}$ would be undefined.

- If $P_1 = P_2$, it then follows directly that the denominator of the coefficients of $P_1 + P_2 = 2 \cdot P_1$ is not divisible by p.

- If $P_1 \neq P_2$, we write as before $x_2 = x_1 + p^k \cdot x$ with $k \in \mathbb{N}$ such that neither the denominator nor the numerator of x is divisible by p and then $y_2 = y_1 + p^k \cdot y$ where y is chosen to have neither the denominator nor the numerator divisible by p also.

  Consider now the equation:

$$\begin{aligned} y_2^2 &\equiv y_1^2 + p^k \cdot x \cdot (3 \cdot x_1^2 + a) \pmod{p} \\ y_2^2 - y_1^2 &\equiv p^k \cdot x \cdot (3 \cdot x_1^2 + a) \pmod{p} \\ \frac{y_2^2 - y_1^2}{p^k} &\equiv x \cdot (3 \cdot x_1^2 + a) \pmod{p} \end{aligned}$$

  to find

$$\frac{y_2^2 - y_1^2}{x_2 - x_1} \equiv 3 \cdot x_1^2 + a \pmod{p}.$$

  Multiply by $y_2 + y_1$. Since p does not divide $y_2 + y_1 = 2 \cdot y_1 \pmod{p}$, it follows that p does not divide the denominator of:

$$\frac{y_2 - y_1}{x_2 - x_1} = \frac{y_2^2 - y_1^2}{(y_2 + y_1) \cdot (x_2 - x_1)}$$

  and hence does not divide the denominator of the coordinates of $P_1 + P_2$ as required.

We have thus shown that the point $P_1 + P_2$ has coordinates with denominator not prime to n if, and only if, there exists a prime $p \mid n$ such that $r_p(P_1 + P_2) = \overline{\mathcal{O}}$.

□

The above algorithm allows you to find a non-trivial divisor d of n eventually. But how do you find the prime decomposition of n? Once the algorithm terminates, you find yourself with not one, but two non-trivial divisors of n, d and a number $k \in \mathbb{N}$ such that $n = d \cdot k$. Check if d and k are prime. If they both are then you are done, otherwise repeat the algorithm setting $n = d$ and/or $n = k$ to find their respective non-trivial divisors. Repeat this until you find yourself with a list of prime numbers only which form the prime decomposition of your initial n. This method is a very good example of **dynamic programming**: we are breaking down the problem into simpler sub-problems. Divide and conquer!

If we pick n to have two prime divisors p and q, then $y^2 = x^3 + a \cdot x + b \pmod{n}$ implies the same equation also holds modulo p and modulo q as a consequence of the Chinese Remainder Theorem [17, Theorem 3.10]. The set of solutions to these two equations over $\mathbb{F}_p$ and $\mathbb{F}_q$ respectively define groups with the usual Addition Laws.

The Hasse-Weil Bound tells us that the order of such a group is in the interval:

$$p + 1 - 2 \cdot \sqrt{p} \leq \#\overline{E}(\mathbb{F}_p) \leq p + 1 + 2 \cdot \sqrt{p}.$$

If our k is of reasonable size (perhaps 10^8), the density of k-**smooth** integers in the interval is high and the distribution of orders of random elliptic curves is sufficiently uniform [3, Section 7.1]. Therefore, if we repeat step (1) and choose several random elliptic curves, then it is highly likely that at least one will have k-smooth order and lead to a non-trivial divisor of n [37].

Remark 5.4.3. The pair (E, P) is generated in some random way. There are many ways to do this and in practice we will choose to use some deterministic method capable of generating many such pairs. See [28] for some algorithms.

Example 5.4.4. Coming back to Example 4.2.2, we have $E : y^2 = x^3 - x$ over $\mathbb{Q}$. Suppose we wish to find a non-trivial divisor of 25.

Suppose we want to add $P_1 = (2, 9)$ and $P_2 = (7, 6)$ on $E(\mathbb{Z}/25\mathbb{Z})$ as before. We start by computing the slope of the line through the two points

$$s = \frac{6 - 9}{7 - 2} = \frac{-3}{5}.$$

Notice that the denominator 5 is neither zero nor invertible in $\mathbb{Z}/25\mathbb{Z}$ since it divides the characteristic 25. We have thus found a non-trivial divisor of the characteristic 25, that is $5 \mid 25$.

We could have easy computed that ourselves, but we take this as an illustration of Lenstra's method which works for much larger characteristics in practice.

Asymptotically, Lenstra's prime factorisation is the third fastest integer factorisation algorithm, beaten only by the **Quadratic Sieve** [38] and the **General Number Field Sieve** [39]. However, it remains the fastest whose running time depends on the size of the smallest prime factor. In fact, although Lenstra's method is slightly slower on products of two similarly sized primes, it will be quicker when the number in question has a small prime factor. It is for this reason that Lenstra's algorithm is still widely used. Notably, it is used as part of these faster methods to look for medium sized prime factors of numbers that appear in intermediate steps [3, Section 7.1]. For modern implementations and performance techniques, refer to [40].

Chapter 6

Conclusion

Through the diversity of applications we have encountered, it becomes apparent that the abelian group structure of elliptic curves has far-reaching consequences. There are many more applications of elliptic curves out there that we did not have the time to cover like the famous **ElGammal encryption** method [3, Section 6.4]. Still today, new applications of elliptic curve Cryptography are being developed. Had there been more time, some topics of interest worth mentioning would have been hyper-elliptic curve cryptography which offers potentially higher bit-securities at the cost of higher computational costs (the group law is much more complicated). See [41].

Bibliography

[1] Ezra Brown. Three fermat trails to elliptic curves. *The College Mathematics Journal*, 31:162–172, 2000.

[2] John Stillwell. *Mathematics and Its History*. Springer, 3 edition, 2010.

[3] Lawrence C. Washington. *Elliptic Curves: Number Theory and Cryptography*. Chapman and Hall/CRC, 2 edition, 2008.

[4] Scott A. Vanstone & Alfred Menezes Darrel Hankerson. *Guide to Elliptic Curve Cryptography*. Springer Professional Computing, 2004.

[5] Wikipedia contributors. Lenstra elliptic-curve factorization — Wikipedia, the free encyclopedia, 2022. [Online; accessed 11-March-2022].

[6] Anthony W. Knapp. *Elliptic Curves*, volume 40. Princeton University Press, 1992.

[7] Evmorfia-Iro Bartzia. *A formalization of elliptic curves for cryptography*. Theses, Université Paris Saclay, February 2017.

[8] Genus of a curve — encyclopedia of mathematics. [Online: accessed 23-February-2022].

[9] Wikipedia contributors. Genus (mathematics) — Wikipedia, the free encyclopedia, 2022. [Online: accessed 23-February-2022].

[10] Robin Hartshorne. *Algebraic geometry*. Number 52 in Graduate Texts in Mathematics. New York: Springer-Verlag, 1977.

[11] Joseph H. Silverman & John T. Tate. *Rational Points on Elliptic Curves*. New York: Springer-Verlag, 2 edition, 2015.

[12] Richard E. Schwartz. Lecture notes in math 1540 (galois theory) - the weierstrass elliptic curve group law. [Online: accessed 25-February-2022].

[13] Joel V. Brawley & George E. Schnibben. *Infinite Algebraic Extensions of Finite Fields*, volume 95. American Mathematical Society, 1989.

[14] Wikipedia contributors. Cyclic group — Wikipedia, the free encyclopedia, 2021. [Online; accessed 9-March-2022].

[15] Joseph H. Silverman. *The arithmetic of elliptic curves*. New York: Springer-Verlag, 1986.

[16] René Schoof. Elliptic curves over finite fields and the computation of square roots mod p. *Mathematics of Computation*, 44:483–494, 1985.

[17] Gareth A. Jones & Josephine M. Jones. *Elementary Number Theory*. London: Springer London, Limited, 1998.

[18] Andrew V. Sutherland. Lecture notes for 18.783 - elliptic curves (lecture 9), May 2015. [Online; accessed 8-March-2022].

[19] Emil Artin. Quadratische körper im gebiete der höheren kongruenzen. ii. analytischer teil. *Mathematische Zeitschrift*, 19:207–246, 1924.

[20] Helmut Hasse. Zur theorie der abstrakten elliptischen funktionenkörper. i, ii & iii. *Crelle's Journal*, 1936, 1936.

[21] André Weil. Numbers of solutions of equations in finite fields. *Bulletin of the American Mathematical Society*, 55:497–508, 1949.

[22] Martin Liebeck. *A Concise Introduction to Pure Mathematics*. Chapman and Hall/CRC., 4 edition, 2016.

[23] Wikipedia contributors. Endomorphism ring — Wikipedia, the free encyclopedia, 2021. [Online; accessed 5-March-2022].

[24] Michael Galerpin. *Torsion points of elliptic curves*. PhD thesis, The University of Chicago, August 2013.

[25] Adi Shamir & Leonard M. Adleman Ronald L. Rivest. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, 21:120–126, 1978.

[26] Michael Goodrich & Roberto Tamassia. Introduction to computer security: Pearson new international edition, 2014.

[27] Seungyun Baek, September 2022. [Online; accessed 10-March-2022].

[28] Neal Koblitz. *A course in number theory and cryptography*. Graduate Texts in Mathematics, vol. 114. Springer-Verlag, New York, 2 edition, 1994.

[29] Whitfield Diffie & Martin E. Hellman. New directions in cryptography. *IEEE Transactions on Information Theory*, 22:644–654, 1976.

[30] Neal Koblitz. Elliptic curve cryptosystems. *Mathematics of Computation*, 48:203–209, January 1987.

[31] Victor S. Miller. Use of elliptic curves in cryptography. *Conference on the theory and application of cryptographic techniques*, pages 417–426, August 1985.

[32] Euclid & Thomas Little Heath. *The Thirteen Books of Euclid's Elements, Books 1 and 2*. Dover Publications, Inc., 1956.

[33] Steven G. Krantz. *The Proof is in the Pudding: The Changing Nature of Mathematical Proof*. New York: Springer, 2011.

[34] Hendrik Willem Lenstra Jr. Elliptic curves and number-theoretic algorithms. *A. M. Gleason (ed.), Proceedings of the International Congress of Mathematicians*, pages 99–120, 1986.

[35] John M. Pollard. Theorems on factorization and primality testing. *Mathematical Proceedings of the Cambridge Philosophical Society*, 76(3):521–528, 1974.

[36] Susanne Schmitt & Horst G. Zimmer. *Elliptic Curves: A Computational Approach*. De Gruyter, 2008.

[37] Hendrik Willem Lenstra Jr. Factoring integers with elliptic curves. *Annals of Mathematics*, 126:649—673, 1987.

[38] Carl Pomerance. Analysis and comparison of some integer factoring algorithms. *Computational Methods in Number Theory, Mathematical Centre Tracts*, 154:89–139, 1982.

[39] Matt Briggs. *An Introduction to the General Number Field Sieve*. PhD thesis, Faculty of the Virginia Polytechnic Institute and State University, 1998.

[40] Peter Birkner & Tanja Lange Daniel J. Bernstein and Christiane Peters. Ecm using edwards curves. *Mathematics of Computation*, 82(282):1139–1179, 2013.

[41] Fangguo Zhang. Bit security of the hyperelliptic curves diffie-hellman problem. *IACR Cryptol. ePrint Arch.*, 2015:614, 2017.