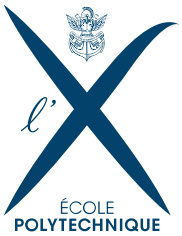

Elliptic Curves: Cryptography and Mordell–Weil theorem

Bachelor Thesis

Sirawit Pongnakintr

sirawit.pongnakintr@polytechnique.edu

March 17, 2024

Supervised by

Prof. Diego Izquierdo

Abstract

This bachelor thesis is a continuation of the lab research project "Introduction to Elliptic Curves". In this thesis, emphasis is given on understanding the applications of elliptic curves in cryptography and the theory behind elliptic curves over $\mathbb{Q}$, i.e. Mordell–Weil theorem. This thesis will therefore be divided into two corresponding parts: Cryptographic applications and Mordell–Weil theorem. The implementation of the algorithms are given in the appendix.

Contents

Acknowledgement

The author greatly thanks his supervisor, Professor Diego Izquierdo, for his help with guiding through further topics in elliptic curves. The author thanks his friend, Pitchayut Saengrungkongka, for minor help with algebraic number theory. The author also thanks his parents (Taweesuk Pongnakintr and Tipawan Pongnakintr), his friends (Dr. Natkamon Tovanich and Yoshimi-Théophile Etienne), and the clinical psychologist (Ms. Anne Mortureux) for emotional support through tough times.

Chapter 1

Cryptographic Applications

1.1 Introduction

The first part of the thesis, i.e., this chapter, concerns about the computational aspects of the theory of elliptic curves, mainly with a focus towards cryptography. The most notable one is the ElGamal Public Key Cryptosystem, which is now a standard cryptosystem in asymmetric cryptography. We start by an important result of Hasse and go on towards the two main computational problems in cryptography: Integer factorization and Discrete Logarithm Problem (DLP). After that we will consider other miscellaneous algorithms involving the use of elliptic curves.

Recall that we denote by K the base field, which, in this chapter, is often a finite field: $\mathbb{F}_{p^n}$ with p prime and $n \in \mathbb{N}^*$. The notation $\bar{K}$ denotes the algebraic closure of K, which is well-defined and uniquely defined. An elliptic curve can be described as a smooth projective curve of genus one with a specified base point. However, in this context and later on we will often use the fact [22, III.3.1] that such curve can be parametrized by a Weierstrass equation of the form

$$E\colon Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3,$$

with $a_1, a_2, a_3, a_4, a_6 \in \bar{K}$. If $a_1, a_2, a_3, a_4, a_6 \in K$, we say that E is defined over K and write E/K. We dehomogenize the equation and write $E\colon y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$, together with an extra point at infinity $O = [0, 1, 0]$. We will normally use O to be the base point.

Here we recall the basic notions involving varieties and maps between them, to be used later. The main reference is [22, I, II.1, II.2]. Since the basic notions are not the main focus of this thesis, this section will try to give only what we need to use, as quickly as possible, with an obvious lack of examples.

Definition (Affine space)**.** *We denote by $\mathbb{A}^n(K)$ the space K^n. Implicitly, $\mathbb{A}^n$ means $\mathbb{A}^n(\bar{K})$.*

Definition (Affine algebraic set)**.** *A set $V \subseteq \mathbb{A}^n$ is said to be an affine algebraic set if it is the set of zeros of an ideal, i.e., there exists an ideal $I \subseteq \bar{K}[X_1, \ldots, X_n]$ such that*

$$V = \{(x_1, \ldots, x_n) \in \mathbb{A}^n\colon \text{for all } f \in I, f(x_1, \ldots, x_n) = 0\}.$$

Definition (Ideal of an affine algebraic set)**.** *For a given affine algebraic set V, we denote by $I(V)$ the ideal*

$$\{f \in \bar{K}[X_1, \ldots, X_n]\colon f(P) = 0 \text{ for all } P \in V\}.$$

Recall that for a given ideal I, $I(V_I)$ isn't necessarily I, but for a given algebraic set V, $V_{I(V)}$ is V.

Definition (K-rational points of an affine algebraic set)**.** *An affine algebraic set $V \subseteq \mathbb{A}^n$ is said to be defined over K if $I(V)$ can be generated by polynomials in $K[X_1, \ldots, X_n]$. In such case, we denote by $V(K)$ the set of K-rational points $V \cap \mathbb{A}^n(K)$.*

Definition (Affine algebraic variety)**.** *An affine algebraic set over $\mathbb{A}^n$ is said to be a variety if its ideal is a prime ideal in $\bar{K}[X_1, \ldots, X_n]$.*

Definition (Coordinate ring)**.** *Given an algebraic set $V \subseteq \mathbb{A}^n$, we denote by $\bar{K}[V]$ the quotient*

$$\bar{K}[T]/I(V)$$

and call this structure the coordinate ring *of V. Similarly $K[V]$ is defined as $K[T]/(I(V) \cap K[T])$.*

Definition (Function field)**.** *The field of fractions of the coordinate ring $\bar{K}[V]$ is called the* function field*, and is written as $\bar{K}(V)$, i.e., $\bar{K}(V) := \mathrm{Frac}(\bar{K}[V])$. Similarly, $K(V) := \mathrm{Frac}(K[V])$.*

Definition (Projective space)**.** *We denote by $\mathbb{P}^n(K)$ the space $K^{n+1} \setminus \{0\}$ modulo the equivalence relation $\sim$ defined by $(x_0, \ldots, x_n) \sim (y_0, \ldots, y_n)$ if and only if there exists $\lambda \in \bar{K}^\times$ such that $x_i = \lambda y_i$ for all $i \in \{0, \ldots, n\}$. We denote such equivalence class of $(x_0, \ldots, x_n)$ by $[x_0, \ldots, x_n]$. We denote by $\mathbb{P}^n$ the space $\mathbb{P}^n(\bar{K})$.*

Definition (Homogeneous polynomial)**.** *A polynomial $f \in \bar{K}[X_0, \ldots, X_k]$ is said to be homogeneous of degree d if $f(\lambda X_0, \ldots, \lambda X_k) = \lambda^d f(X_0, \ldots, X_k)$ for all $\lambda \in \bar{K}$. An ideal $I \subseteq \bar{K}[X_0, \ldots, X_k]$ is said to be a homogeneous ideal if it is generated by homogeneous polynomials.*

Definition (Projective algebraic set)**.** *A set $V \subseteq \mathbb{P}^n$ is said to be a projective algebraic set if it is the set of zeros of a homogeneous ideal, i.e. there exists a homogeneous ideal $I \subseteq \bar{K}[X_0, \ldots, X_n]$ such that*

$$V = \{[x_0, \ldots, x_n] \in \mathbb{P}^n \colon \textit{for all homogeneous } f \in I, f(x_0, \ldots, x_n) = 0\}.$$

In general, given a homogeneous ideal I, we denote by V_I this set, and call it "the algebraic set generated by I".

Definition (Ideal of a projective algebraic set)**.** *Let $V \subseteq \mathbb{P}^n$ be a projective algebraic set, then we define its ideal $I(V)$ to be the ideal generated by the set*

$$\{f \in \bar{K}[X_0, \ldots, X_n] \colon f \textit{ is homogeneous and for all } P \in V, f(P) = 0\}.$$

Definition (K-rational points of a projective algebraic set)**.** *A projective algebraic set $V \subseteq \mathbb{P}^n$ is said to be defined over K if $I(V)$ can be generated by homogeneous polynomials in $K[X_0, \ldots, X_n]$. In such case, we denote by $V(K)$ the set of K-rational points $V \cap \mathbb{P}^n(K)$.*

Definition (Projective variety)**.** *A projective algebraic set V is said to be a projective variety if $I(V)$ is a prime ideal in $\bar{K}[X_0, \ldots, X_n]$.*

Now, for a projective variety, one can identify a subset of it with an affine variety by using the map

$$\phi_i \colon \begin{cases} \mathbb{A}^n & \to \mathbb{P}^n \\ (y_1, \ldots, y_n) & \mapsto [y_1, \ldots, y_{i-1}, 1, y_i, \ldots, y_n] \end{cases}$$

which is injective. The image of this map is the set

$$\{[x_0, \ldots, x_n] \in \mathbb{P}^n \colon x_i \neq 0\}.$$

We denote this set by U_i, so that $\phi_i^{-1} \colon U_i \to \mathbb{A}^n$ is a well-defined bijection. Now we introduce the notions of coordinate ring and function field for projective variety as follows. Note that the notion of ϕ_i and U_i will be used in the next two definitions.

Definition. *Let V be a projective variety and let $W := \phi_i^{-1}(V \cap U_i)$ for some $i \in \{0, \ldots, n\}$. Then W is an affine variety in $\mathbb{A}^n$. The coordinate ring $\bar{K}[V]$ for the projective variety V is defined as $\bar{K}[W]$, and the function field $\bar{K}(V)$ for the projective variety V is defined as $\bar{K}(W)$. It is admitted that this is well-defined, i.e., it doesn't depend on the choice of i.*

Definition (Smooth points)**.** *For an affine variety $V \subseteq \mathbb{A}^n$, let $f_1, \ldots, f_m \in \bar{K}[X_1, \ldots, X_n]$ be such that $\{f_1, \ldots, f_m\}$ is a generating set for the ideal $I(V)$. Let $P \in V$, then V is smooth at P if the $m \times n$ matrix*

$$\left(\frac{\partial f_i}{\partial X_j}(P)\right)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$$

has rank $n - \dim V$. If V is smooth at every point, then we say that V is smooth. Now, for a projective variety $V \subseteq \mathbb{P}^n$, a point $P \in V$ is said to be smooth if $\phi_i^{-1}(P)$ is smooth in the affine variety $\phi_i^{-1}(V \cap U_i)$ for some $i \in \{0, \ldots, n\}$ such that $P \in U_i$.

Definition (Rational map)**.** *Let $V_1, V_2 \subseteq \mathbb{P}^n(\bar{K})$ be projective varieties. A collection of functions $f_0, \ldots, f_n \in \bar{K}(V_1)$ is said to define a rational map ϕ if for every point $P \in V_1$ at which $f_0, \ldots, f_n$ are all defined, $\phi(P) = [f_0(P), f_1(P), \ldots, f_n(P)] \in V_2$.*

Furthermore, we say that ϕ is defined over K if there exists $\lambda \in \bar{K}^\times$ such that $\lambda f_0, \ldots, \lambda f_n \in K(V_1)$.

Remark. *Even if we write "$\phi\colon V_1 \to V_2$ is a rational map", this does not assume that ϕ is well-defined at every point of V_1! The notion of being defined at a point is given with the next definition.*

Definition (Regular points)**.** *A rational map $\phi = [f_0, \dots, f_n]\colon V_1 \to V_2$ is said to be regular at $P \in V_1$ if there is a function $g \in \bar{K}(V_1)$ such that for each $i \in \{0, \dots, n\}$, gf_i can be evaluated at P, and that there exists some i for which $(gf_i)(P) \neq 0$. If this is the case, we define the evaluation of ϕ at P as*

$$\phi(P) = [(gf_0)(P), \dots, (gf_n)(P)] \in V_2.$$

Note that it may be necessary to take different g's at different P's, and if a rational map is regular everywhere, we say that it is a morphism.

The following part about transcendence degree utilizes [23, Tag 030D] as the main reference.

Definition (Algebraic independence)**.** *Let K be a field. Let L/K be an extension. Let $\mathscr{F} = \{\alpha_i\}_{i\in I}$ be a family of elements in L. We say that $\mathscr{F}$ is* algebraically independent *over K if the evaluation map*

$$\begin{cases} K[\{X_i\}_{i\in I}] & \to L \\ P & \mapsto P((\alpha_i)_{i\in I}) \end{cases}$$

(evaluating the polynomial P at $(x_i)_{i\in I}$) is injective.

Definition (Transcendence basis)**.** *A transcendence basis of L/K is a set $\mathscr{F} = \{\alpha_i\}_{i\in I}$ of elements in L such that $\mathscr{F}$ is algebraically independent over K and $L/K((\alpha_i)_{i\in I})$ is an algebraic extension.*

Definition (Transcendence degree)**.** *Let L/K be a field extension. Then the transcendence degree of L over K, denoted by $\operatorname{trdeg}_K(L)$, is defined by the cardinality of a transcendence basis of L/K. Note that this requires proving that all transcendence bases have the same cardinality, which we refer to [23, Tag 030D, 9.26.3].*

Definition (Dimension of a variety)**.** *The dimension of a variety V, denoted by $\dim(V)$, is defined as $\operatorname{trdeg}_{\bar{K}}(\bar{K}(V))$. This is valid for both the affine case and the projective case.*

Definition (Curve)**.** *A curve is a projective variety of dimension one.*

After on, by convention, we let p denote any prime integer, and we let $q = p^n$ denote a (strictly-positive-integer-) power of prime.

1.2 Hasse's Bound

The main question about elliptic curve over finite fields is about bounding the number of points (in $\mathbb{F}_q$) on a curve $E/\mathbb{F}_q$. Consider a curve $E/\mathbb{F}_q$ defined by

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

with $a_1, a_2, a_3, a_4, a_6 \in \mathbb{F}_q$. For any fixed $x \in \mathbb{F}_q$, there exists at most two solutions $y \in \mathbb{F}_q$, so $\#E(\mathbb{F}_q) \leq 2q + 1$.

It is conjectured in the thesis of Emil Artin [3] that $q + 1 - 2\sqrt{q} \leq \#E(\mathbb{F}_q) \leq q + 1 + 2\sqrt{q}$. This turns out to be true as proven by Helmut Hasse. We will complete the proof later but here we note that there are a few claims to be proven first:

- The Galois group $\operatorname{Gal}(\bar{\mathbb{F}}_q/\mathbb{F}_q)$ is the topological closure of the subgroup generated by

  $$\operatorname{Frob}_q\colon x \mapsto x^q.$$

  (Note that $\operatorname{Frob}_q \in \operatorname{Gal}(\bar{\mathbb{F}}_q/\mathbb{F}_q)$, and the topology in $\operatorname{Gal}(\bar{\mathbb{F}}_q/\mathbb{F}_q)$ is the profinite group topology–we will discuss in detail later)

- The map $(1 - \phi)\colon E \to E$ is separable. ([22, III.5.5])
- If an isogeny $\phi\colon E_1 \to E_2$ is separable then $\#\ker\phi = \deg\phi$. ([22, II.4.10c])
- If A is an abelian group and $d\colon A \to \mathbb{Z}$ is a positive definite quadratic form, then

  $$|d(\psi - \phi) - d(\phi) - d(\psi)| \leq 2\sqrt{d(\phi)d(\psi)} \qquad \text{for all } \psi, \phi \in A.$$

  ([22, V.1.2], a form of Cauchy–Schwarz inequality, and we will define what it means to be a positive definite quadratic form in this context later)

To begin, it is better to recall and give a proper introduction to inverse limit and profinite group.

1.2.1 Inverse Limit and Profinite Group

Remark. *This subsection utilizes [20], the lecture notes of 18.785 at MIT, given by Pitchayut Saengrungkongka (Mark), a friend of mine. His notes is currently still unpublished publicly.*

Let us begin with a concrete example first.

Example 1 (p-adic integers)**.** *Let p be a prime number. We have the sequence of rings*

$$\cdots \to \mathbb{Z}/p^3\mathbb{Z} \to \mathbb{Z}/p^2\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}.$$

where $a_{i+1} \in \mathbb{Z}/p^{i+1}\mathbb{Z} \mapsto a_i \in \mathbb{Z}/p^i\mathbb{Z}$ is the modulo operation. If a sequence $(a_n)_{n\geq 1}$ such that $a_n \in \mathbb{Z}/p^n\mathbb{Z}$ satisfies the condition (for all $i \in \mathbb{N}^*$, $a_{i+1} \mapsto a_i$ with the modulo operation), then the sequence $(a_n)_{n\geq 1}$ is said to be a p-adic integer, and we denote by $\mathbb{Z}_p$ the set of all p-adic integers.

Example 2. *When $p = 5$, we have the sequence*

$$(2 \bmod 5, 17 \bmod 5^2, 67 \bmod 5^3, \dots)$$

as a p-adic integer. Note that it doesn't need to correspond to any integer.

We generalize this concept of p-adic integers to define the inverse limit.

Definition (Inverse limit)**.** *Given sets and maps*

$$\cdots \xrightarrow{f_2} X_2 \xrightarrow{f_1} X_1 \xrightarrow{f_0} X_0.$$

Define

$$\varprojlim X_i := \left\{ (x_i) \in \prod_{i\geq 0} X_i \colon f_i(x_{i+1}) = x_i \text{ for all } i \geq 0 \right\}.$$

Example 3 (Roots of unity)**.** *This example is given by [22, III.7.3.] Suppose $\bar{K}$ is an algebraically closed field. Let*

$$\boldsymbol{\mu}_{\ell^n} \subseteq \bar{K}^\times$$

be the group of ℓ^n-th roots of unity. Raising to the ℓ-th power gives the maps

$$\boldsymbol{\mu}_{\ell^{n+1}} \xrightarrow{z\mapsto z^\ell} \boldsymbol{\mu}_{\ell^n}.$$

Hence, this fits into the definition of the inverse limit, i.e. let X_n be $\boldsymbol{\mu}_{\ell^n}$ for all $n \in \mathbb{N}$ and let $f_n \colon X_{n+1} \to X_n$ be defined as $z \mapsto z^\ell$. This defines $\varprojlim_n X_n$, which is a Tate module of $\bar{K}^\times$.

Now let us consider a little generalization of the inverse limit.

Definition (Directed set)**.** *A directed set $(I, \preceq)$ is a set I equipped with a partial order $\preceq$ such that every finite subset has an upper bound.*

Definition (Inverse system)**.** *An inverse system indexed by a directed set $(I, \preceq)$ consists of a family of sets $(X_i)_{i\in I}$ and a family of maps $(f_{i,j} \colon X_j \to X_i)_{(i,j)\in\preceq}$[1], satisfying*

$$f_{i,i} = \mathrm{id}_{X_i} \text{ for all } i \in I \qquad \text{and} \qquad f_{i,j} \circ f_{j,k} = f_{i,k} \text{ for all } i, j, k \in I \text{ such that } i \preceq j \preceq k.$$

Definition (Inverse limit (generalization))**.** *For a given directed set $(I, \preceq)$ and an inverse system $(X_i)_{i\in I}, (f_{i,j})_{i\preceq j}$, its inverse limit is defined as*

$$\varprojlim X_i := \left\{ (x_i) \in \prod_{i\in I} X_i \colon f_{i,j}(x_j) = x_i \text{ for all } j \in I \text{ such that } i \preceq j \right\}.$$

Proposition 4. *(from [19, 2.2]) Let Ω/K be a Galois extension. Then*

$$\Phi \colon \begin{cases} \mathrm{Gal}(\Omega/K) & \to \varprojlim_{\Omega/L/K} \mathrm{Gal}(L/K) \\ \sigma & \mapsto (\sigma|_L)_{\Omega/L/K} \end{cases}$$

is a group isomorphism.

[1] The notation $(i,j) \in \preceq$ is same as $i \preceq j$. We write it this way to make it clear that we consider every possible pair $(i,j) \in I \times I$ satisfying $i \preceq j$.

Proof. Consider the family $\mathscr{F}$ of all field L such that Ω/L is Galois[2] (i.e. L is a sub-Galois-extension of Ω/K). This family forms a directed set because Ω is an upper bound for every L. Then, consider the family $(X_L)_{L\in\mathscr{F}}$ defined by $X_L := \mathrm{Gal}(L/K)$ with the corresponding family of maps $(f_{L,M})_{\bar{K}/M/L/K}$ defined by

$$f_{L,M}\colon \begin{cases} \mathrm{Gal}(M/K) & \to \mathrm{Gal}(L/K) \\ \sigma & \mapsto \sigma|_L \end{cases}$$

for all L, M such that $\Omega/M/L/K$. This forms an inverse system, so $\varprojlim X_L$ is well-defined.

It is easy to check that the image by Φ lands on $\varprojlim \mathrm{Gal}(L/K)$ and that Φ is a group homomorphism. Now, consider the kernel

$$\ker(\Phi) = \{\sigma \in \mathrm{Gal}(\Omega/K)\colon \sigma|_L = \mathrm{id}_L \text{ for all } \Omega/L/K\}.$$

Indeed, if σ is not id_Ω, then some element got permuted, and that element must be in some subextension L, so σ wouldn't be in the kernel. Hence, $\ker(\Phi) = \{\mathrm{id}_\Omega\}$. So Φ is injective. Now we're left with showing that it is surjective. Suppose $(\sigma|_L)_{\Omega/L/K}$ is in $\varprojlim \mathrm{Gal}(L/K)$ then we can define a preimage σ by covering it with each L since $\Omega = \bigcup_{\Omega/L/K} L$, i.e. $\sigma(x) := \sigma|_L(x)$ if $x \in L$. To check that it is well-defined, we must check that if $x \in L$ and $x \in L'$ then $\sigma|_L(x) = \sigma|_{L'}(x)$. This is true because $M := L \cap L'$ is lower in the poset such that L/M and L'/M are both Galois, and by definition of $\varprojlim \mathrm{Gal}(L/K)$, $\sigma|_M(x) = \sigma|_L(x)$ and $\sigma|_M(x) = \sigma|_{L'}(x)$ so they must coincide. This proves the surjectivity and hence completes the proof. □

Definition (Profinite Group)**.** *A group G is profinite if it is an inverse limit of finite groups $(G_i)_{i\in I}$. Usually, we endow each finite group G_i with the discrete topology, and then promote G to a topological group. The topology on G is given by the subset topology of the product topology on $\prod_{n\in I} G_i$.*

Proposition 5. *Every profinite group is compact and Hausdorff.*

Proof. Each G_i is finite with the discrete topology, hence compact and Hausdorff. Apply [5, 5.11.].[3] (Tychonoff's theorem together with another theorem which says the product of Hausdorff spaces is Hausdorff) □

Example 6. *Let K be a perfect field, and $\bar{K}$ be its algebraic closure. Then $\mathrm{Gal}(\bar{K}/K)$ is a compact group.*

Proof. Apply 4 to see that $\mathrm{Gal}(\bar{K}/K) \cong \varprojlim \mathrm{Gal}(L/K)$ which is profinite hence compact. □

Example 7 ($\mathrm{Gal}(\bar{\mathbb{F}}_p/\mathbb{F}_p) \cong \widehat{\mathbb{Z}}$)**.** *We define $\widehat{\mathbb{Z}}$, the set of profinite integers, as*

$$\varprojlim_n \mathbb{Z}/n\mathbb{Z}$$

using the divisibility poset as its directed set, and the natural modulo as maps in the inverse system. Now, apply 4 and the fact that $\mathrm{Gal}(\mathbb{F}_{p^m}/\mathbb{F}_{p^n}) \cong \mathbb{Z}/\frac{m}{n}\mathbb{Z}$ for any integers n dividing m [9, 5.4.5(i)] to see that

$$\mathrm{Gal}(\bar{\mathbb{F}}_p/\mathbb{F}_p) \cong \varprojlim_n \mathrm{Gal}(\mathbb{F}_{p^n}/\mathbb{F}_p) \cong \varprojlim_n \mathbb{Z}/n\mathbb{Z} \cong \widehat{\mathbb{Z}}.$$

We will prove the proposition using the following the following lemma. In this section, we utilize the following notation. For any prime power $q = p^n$ (p being prime, $n \geq 1$ being an integer), Frob_q denotes the map $x \mapsto x^q$ defined on $\bar{\mathbb{F}}_q$, so that $\mathrm{Frob}_q \in \mathrm{Gal}(\bar{\mathbb{F}}_q/\mathbb{F}_q)$, and Frob_q^k denotes $\underbrace{\mathrm{Frob}_q \circ \cdots \circ \mathrm{Frob}_q}_{k \text{ times}}$.

Lemma 8 (rearranged and summarized from [9])**.** *Let p be a prime number and $n \geq 2$ be an integer. The group $\mathrm{Gal}(\mathbb{F}_{p^{in}}/\mathbb{F}_{p^n})$ is generated by the map $\mathrm{Frob}_q|_{\mathbb{F}_{p^{in}}}$, so it is a cyclic group of order i.*

[2] L/K is automatically Galois.

[3] The terminology "compact" from [5] actually means "compact and Hausdorff" in the usual sense. And the "compact" in the usual sense is called "quasi-compact" in [5].

Proof. Let $\Phi := \mathrm{Frob}_q|_{\mathbb{F}_{p^{in}}} \in \mathrm{Gal}(\mathbb{F}_{p^{in}}/\mathbb{F}_{p^n})$. The map Φ^i then denotes the map $x \mapsto x^{q^i} = x^{p^{in}} = x$. So Φ^i is $1_{\mathrm{Gal}(\mathbb{F}_{p^{in}}/\mathbb{F}_{p^n})}$. Now let us check whether there exists another integer $2 \leq j < i$ such that $\Phi^j = 1_{\mathrm{Gal}(\mathbb{F}_{p^{in}}/\mathbb{F}_{p^n})}$. Well Φ^j is the map $x \mapsto x^{p^{jn}}$. This cannot be identical to the identity map because otherwise the polynomial $x^{p^{jn}} - x = 0$ would kill all $x \in \mathbb{F}_{p^{in}}$, i.e., the polynomial would contain p^{in} roots, but the degree of the polynomial is just $p^{jn} < p^{in}$, a contradiction. So $\Phi^j \neq 1_{\mathrm{Gal}(\mathbb{F}_{p^{in}}/\mathbb{F}_{p^n})}$ for all $2 \leq j < i$, hence the order of Φ is i. This proves that i divides the order of the Galois group. But $\#\mathrm{Gal}(\mathbb{F}_{p^{in}}/\mathbb{F}_{p^n}) = [\mathbb{F}_{p^{in}} : \mathbb{F}_{p^n}] = i$. The equality is true since the extension is normal and separable. Hence, $\{1, \Phi, \Phi^2, \ldots, \Phi^{i-1}\} = \mathrm{Gal}(\mathbb{F}_{p^{in}}/\mathbb{F}_{p^n})$ and so Φ is a generating element of the cyclic group $\mathrm{Gal}(\mathbb{F}_{p^{in}}/\mathbb{F}_{p^n})$ of order i. □

Now let us get to the main proposition.

Proposition 9. *Let G be $\mathrm{Gal}(\bar{\mathbb{F}}_q/\mathbb{F}_q)$ where $q = p^n$ with p prime. Then G is a profinite group as $G = \varprojlim_m \mathrm{Gal}(\mathbb{F}_{p^{mn}}/\mathbb{F}_{p^n})$. Let $S = \langle \mathrm{Frob}_q \rangle$ be the subgroup generated by $\mathrm{Frob}_q : x \mapsto x^q$. Then the topological closure of S is G.*

Proof. It is enough to show that every nonempty open subset of G intersects S. Because if this is the case, then every neighborhood of a point in G contains an open set, which, by this fact, intersects S, and so G would be a subset of the topological closure of S, by definition.

Now let T be an open subset of G, then it can be written as an intersection between the inverse limit and a union of sets in the basis of the product topology. Recall that the following collection is a basis of the product topology [5, 3.11(i)]:

$$\left\{ (x_i)_{i \in \mathbb{N}^*} \in \prod_{m \in \mathbb{N}^*} \mathrm{Gal}(\mathbb{F}_{p^{mn}}/\mathbb{F}_{p^n}) : \forall j \in J, x_j \in U_j \right\} = \bigcap_{j \in J} \mathrm{pr}_j^{-1}(U_j),$$

where $J \subseteq \mathbb{N}^*$ is finite and U_j is an open set (which is any set since the topology here is discrete) of $\mathrm{Gal}(\mathbb{F}_{p^{jn}}/\mathbb{F}_{p^n})$ for any $j \in J$. In our case, it is much simpler since all of $\mathrm{Gal}(\mathbb{F}_{p^{jn}}/\mathbb{F}_{p^n})$ are equipped with the discrete topology. We observe that $\bigcap_{j \in J} \mathrm{pr}_j^{-1}(U_j)$ is basically

$$\prod_{i \in \mathbb{N}^*} \begin{cases} \mathrm{Gal}(\mathbb{F}_{p^{in}}/\mathbb{F}_{p^n}) & ; \text{ if } i \notin J \\ U_j & ; \text{ if } i \in J. \end{cases} \qquad (\star)$$

In particular, since we picked T to be an open subset of G, and since T is an intersection between the inverse limit and a union of basis, say $T = G \cap (\bigcup_k B_k)$, there exists a set inside the basis (in the form $(\star)$), say B_k. Fix that set (let us rename it to B) and let $J \subseteq \mathbb{N}^*$ be the corresponding finite set and $(U_j)_{j \in J}$ be the corresponding family of finite sets where $U_j \subseteq \mathrm{Gal}(\mathbb{F}_{p^{jn}}/\mathbb{F}_{p^n})$. If J is not in the form $[1, k] \cap \mathbb{Z}$ then we can make it into that form by selecting k to be $\max(J)$ and for all $j' \in ([1, k] \cap \mathbb{Z}) \setminus J$, set $U_{j'} := \mathrm{Gal}(\mathbb{F}_{p^{j'n}}/\mathbb{F}_{p^n})$. This allows us to assume $J = [1, k] \cap \mathbb{Z}$ for some $k \in \mathbb{N}^*$.

Now, we have to show that $G \cap B \cap S = B \cap S \neq \emptyset$. Well, if we can find an element $\sigma \in \mathrm{Gal}(\mathbb{F}_{p^{k!n}}/\mathbb{F}_{p^n})$ such that the reduction of σ by the inverse system into $\mathrm{Gal}(\mathbb{F}_{p^{in}}/\mathbb{F}_n)$ gives a sequence which lies in B. Since the set J is finite, surely this is the case, i.e., there exists such $\sigma \in \mathrm{Gal}(\mathbb{F}_{p^{k!n}}/\mathbb{F}_{p^n})$. Furthermore, we can naturally lift this map into G. By the previous lemma, the group $\mathrm{Gal}(\mathbb{F}_{p^{k!n}}/\mathbb{F}_{p^n})$ is generated by $\mathrm{Frob}_q|_{\mathbb{F}_{p^{k!n}}}$, so $\sigma = \mathrm{Frob}_q|^{\ell}_{\mathbb{F}_{p^{k!n}}}$ for some $\ell \in \mathbb{N}$. Lifting σ to $\Phi = \mathrm{Frob}_q^{\ell}$, one sees that $\Phi \in B \cap S$, hence completes the proof.

□

Now we supply a few lemmas to prove another important tool.

Proposition 10. *Suppose $\bar{K}/K$ is a Galois extension and $\bar{K}/L/K$ is an algebraic extension. The map*

$$\rho_{\bar{K}/L} : \begin{cases} \mathrm{Gal}(\bar{K}/K) & \to \mathrm{Gal}(L/K) \\ \sigma & \mapsto \sigma|_L \end{cases}$$

is a surjective group homomorphism. In other words, any K-homomorphism from L to $\bar{K}$ can be extended to a homomorphism from $\bar{K}$ to $\bar{K}$.

Proof. (Edited from [7]) It is obvious that this map is a group homomorphism. Now let us show that for any $\sigma \in \mathrm{Gal}(L/K)$, one can find $\tilde{\sigma} \in \mathrm{Gal}(\bar{K}/K)$ such that $\tilde{\sigma}|_L = \sigma$. We do this by adjoining roots one by one from L into $\bar{K}$. Suppose $\alpha \in \bar{K} \setminus L$, then one can extend σ into

$\mathrm{Gal}(L(\alpha)/K)$ by observing that $L(\alpha) \cong L[T]/(\mu_\alpha)$ where μ_α is the minimal polynomial of α over $L[T]$. Now we define the function $\sigma' \colon L[T] \to \bar{K}$ as

$$\sum a_i T^i \mapsto \sigma(a_i)\alpha^i.$$

Since $\sigma'(\mu_\alpha) = \mu_\alpha^\sigma(\alpha) = \mu_\alpha(\alpha) = 0$, this induces $\tilde{\sigma} \colon L[T]/(\mu_\alpha) \to \bar{K}$ as $(\mu_\alpha) + P \mapsto \sigma'(P)$. Composing $\tilde{\sigma}$ with the isomorphism $L(\alpha) \cong L[T]/(\mu_\alpha)$ gives a map $\tilde{\sigma} \colon L(\alpha) \to \bar{K}$ which can easily be checked that it is a K-homomorphism from $L(\alpha)$ to $\bar{K}$. This gives an extension of σ from $\mathrm{Hom}_K(L, \bar{K})$ to $\mathrm{Hom}_K(L(\alpha), \bar{K})$. Next, we consider the partially ordered set $\mathscr{F}$ consisting of pairs of the form (M, ψ) where M is an intermediary field extension between L and $\bar{K}$, and $\psi \in \mathrm{Hom}_K(M, \bar{K})$ is such that $\psi|_L = \sigma$, ordered by

$$(M_1, \psi_1) \preceq (M_2, \psi_2) \Leftrightarrow M_1 \subseteq M_2 \text{ and } \psi_2|_{M_1} = \psi_1.$$

Each chain $C_I = \{(M_i, \psi_i)\}_{i \in I}$ has an upper bound. This is because we can define $M := \bigcup_{i \in I} M_i \subseteq \bar{K}$, and define $\psi \in \mathrm{Hom}_K(M, \bar{K})$ by $x \mapsto \psi_i(x)$ whenever $x \in M_i$. This is well-defined because by the chain property of C_I, $(M_i, \psi_i) \preceq (M_j, \psi_j)$ or $(M_j, \psi_j) \preceq (M_i, \psi_i)$ for all $i, j \in I$, so there cannot be two conflicting $(i, j) \in I \times I$ with $x \in M_i, x \in M_j$ but $\psi_i(x) \neq \psi_j(x)$. The pair (M, ψ) is a valid element of $\mathscr{F}$, and is an upper bound for the chain C_I. By Zorn's lemma, we deduce that there exists a maximal element of $\mathscr{F}$. Pick one such maximal element $(M_{\max}, \psi_{\max})$ of $\mathscr{F}$. If $M_{\max} \neq \bar{K}$ then there exists $\alpha \in \bar{K} \setminus M_{\max}$, and by the first part of this proof, one can extend $\psi_{\max}$ from $\mathrm{Hom}_K(M_{\max}, \bar{K})$ to $\tilde{\psi}_{\max} \in \mathrm{Hom}_K(M_{\max}(\alpha), \bar{K})$. Then $(M_{\max}(\alpha), \tilde{\psi}_{\max})$ is a greater element than $(M_{\max}, \psi_{\max})$, contradicting the fact that it is maximal. Therefore, $M_{\max} = \bar{K}$ and $\psi_{\max} \in \mathrm{Hom}_K(\bar{K}, \bar{K}) = \mathrm{Gal}(\bar{K}/K)$ is an extension of $\sigma \in \mathrm{Gal}(L/K)$, so that $\rho_{\bar{K}/L}(\psi_{\max}) = \sigma$. This proves that $\rho_{\bar{K}/L}$ is surjective. □

Corollary 11. *If $\bar{K}/K$ is a Galois extension, $\bar{K}/L$ is a Galois extension, and L/K is a finite extension, then $[L \colon K] = [\mathrm{Gal}(\bar{K}/K) \colon \mathrm{Gal}(\bar{K}/L)]$.*

Proof. By the previous proposition, we apply the first homomorphism theorem of groups to $\rho_{\bar{K}/L}$ to see that

$$\mathrm{Gal}(\bar{K}/K)/\mathrm{Gal}(\bar{K}/L) = \mathrm{Gal}(\bar{K}/K)/\ker(\rho_{\bar{K}/L}) \cong \mathrm{im}(\rho_{\bar{K}/L}) = \mathrm{Gal}(L/K).$$

Since $[L \colon K]$ is finite, we see that $[\mathrm{Gal}(\bar{K}/K) \colon \mathrm{Gal}(\bar{K}/L)]$ is finite too, and is equal to $[L \colon K]$ (see the cardinality of the equation above). □

Lemma 12. *A profinite group G is a topological group, i.e., for the usual profinite group topology on G, the maps $\varphi \colon G \times G \to G$ (multiplication on G) and $\iota \colon G \to G$ (inverse on G) are continuous.*

Proof. Recall that if $G = \varprojlim_{i \in I} G_i$ is a profinite group, then a basis of G is given by the collection of sets of the form

$$\bigcap_{j \in J} \mathrm{pr}_j^{-1}(U_j),$$

where $J \subseteq I$ is finite and $U_j \subseteq G_j$ is an open set (which is any set since G_i's are discrete). It is enough to show that the preimage of each of the member in this collection, through $\varphi \colon G \times G \to G$ gives an open set, and the preimage through $\iota \colon G \to G$ gives also an open set. Fix a finite subset $J \subseteq I$ with $U_j \subseteq G_j$ for each $j \in J$. Then let $V := \left(\bigcap_{j \in J} \mathrm{pr}_j^{-1}(U_j)\right) \cap G$. We have

$$\varphi^{-1}(V) = \varphi^{-1}\left(\bigcap_{j \in J} \mathrm{pr}_j|_G^{-1}(U_j)\right) = \bigcap_{j \in J} \varphi^{-1}(\mathrm{pr}_j|_G^{-1}(U_j)).$$

Now recall that for a fixed $j \in J$,

$$\mathrm{pr}_j|_G^{-1}(U_j) = G \cap \prod_{i \in I} \begin{cases} G_i & ; \text{ if } i \neq j \\ U_i & ; \text{ if } i = j \end{cases}.$$

Recall that $\varphi\colon G\times G\to G$ is actually the restriction of applying $\varphi_i\colon G_i\times G_i\to G_i$ (the group law of G_i) pointwise in the product $\prod_{i\in I}G_i$ onto $G\times G$, so

$$\begin{aligned}\varphi^{-1}(\mathrm{pr}_j|_G^{-1}(U_j)) &= \varphi^{-1}(G)\cap\prod_{i\in I}\begin{cases}\varphi_i^{-1}(G_i) & ;\text{ if } i\neq j\\ \varphi_i^{-1}(U_i) & ;\text{ if } i=j\end{cases}\\ &= (G\times G)\cap\prod_{i\in I}\begin{cases}G_i\times G_i & ;\text{ if } i\neq j\\ \varphi_i^{-1}(U_i) & ;\text{ if } i=j\end{cases}\\ &= (G\times G)\cap {\mathrm{pr}'_j}^{-1}(\varphi_i^{-1}(U_j))\end{aligned}$$

where pr'_j is the projection $\prod_{i\in I}G_i\times G_i\to G_j\times G_j$. Since $\varphi_i^{-1}(U_j)\subseteq G_j\times G_j$ is finite, it is open and so its preimage under the projection is also open. Since $G\times G$ is open, we have $\varphi^{-1}(\mathrm{pr}_j|G^{-1}(U_j))$ is open. Since J is finite, $\varphi^{-1}(V)$ is a finite intersection of open sets, hence open. This proves that the preimage of any member of the usual basis is open in $G\times G$, so $\varphi\colon G\times G\to G$ is continuous.

Now, let us show that $\iota\colon G\to G$ is also continuous. Using the same strategy, let $V:=\left(\bigcap_{j\in J}\mathrm{pr}_j^{-1}(U_j)\right)\cap G$ with $J\subseteq I$ finite and $U_j\subseteq G_j$ for each $j\in J$. We have

$$\iota^{-1}(V)=\bigcap_{j\in J}\iota^{-1}(\mathrm{pr}_j|_G^{-1}(U_j)),$$

but as before, we also have

$$\iota^{-1}(\mathrm{pr}_j|_G^{-1}(U_j))=G\cap\mathrm{pr}_j^{-1}(\iota^{-1}(U_j))$$

which is clearly open by the same argument as before, so ι is continuous. □

Lemma 13. *Let G be a profinite group and let $g\in G$. If $H\leq G$ is open, then gH is open. If $H\leq G$ is closed, then gH is closed.*

Proof. Since a profinite group is a topological group, we see that the group laws $\varphi\colon G\times G\to G$ and $\iota\colon G\to G$ are continuous.

(Taken from [2, Section 4.3]) For any $x\in G$, the map $L_x\colon G\to G$ defined by $y\mapsto xy$ is continuous because it is a composition of continuous functions

$$\begin{aligned}G\to G\times G &\xrightarrow{\varphi} G\\ y\ \mapsto (x,y)\ &\mapsto xy.\end{aligned}$$

Moreover, L_x is bijective since G is a group, we see that $L_x^{-1}=L_{x^{-1}}$ and so L_x^{-1} is continuous, hence L_x is a homeomorphism from G to itself.

This means for any $g\in G$, if $H\leq G$ is open, then $gH=L_g(H)$ is also open since L_g is a homeomorphism. Same for the case that $H\leq G$ is closed. □

Lemma 14. *Let G be a profinite group. If $H\leq G$ is closed and has finite index in G, then H is open.*

Proof. If $H\leq G$ is closed, then the cosets gH are closed for all $g\in G$, by the previous lemma. The complement $G\setminus H$ is a disjoint union of those cosets except H, so if $[G\colon H]$ is finite then the union is a finite union of closed sets, hence closed. □

Lemma 15. *Suppose $\bar{K}/K$ is a Galois extension. If L/K is a finite extension such that $\bar{K}/L$ is a Galois extension, then $\mathrm{Gal}(\bar{K}/L)$ is an open subgroup of the profinite group $\mathrm{Gal}(\bar{K}/K)$.*

Proof. First, let us show that $\mathrm{Gal}(\bar{K}/L)$ is a closed subgroup of $\mathrm{Gal}(\bar{K}/K)$. To do this, it is quite easy because once we list the family $\{G_i\}_{i\in I}$ as $\{\mathrm{Gal}(K_i/K)\}_{i\in I}$ in the infinite product for the profinite group $\mathrm{Gal}(\bar{K}/K)$, we see that

$$\begin{aligned}\mathrm{Gal}(\bar{K}/L) &= \left\{(\sigma_i)_{i\in I}\in\mathrm{Gal}(\bar{K}/K)\colon \sigma_i|_L=\mathrm{id}_L,\forall i\in I\right\}\\ &= \left\{(\sigma_i)_{i\in I}\in\mathrm{Gal}(\bar{K}/K)\colon \sigma_i\in\mathrm{Gal}(K_i/L)\forall i\in I\right\}\\ &= \mathrm{Gal}(\bar{K}/K)\cap\prod_{i\in I}\mathrm{Gal}(K_i/L),\end{aligned}$$

which is an intersection of closed sets (the product is an infinite product of closed sets, hence closed), so $\mathrm{Gal}(\bar{K}/L)$ is closed.

Next, let us show that if we furthermore assume that L/K is finite, then $\mathrm{Gal}(\bar{K}/L)$ is open. Apply 11 to see that $[L\colon K] = [\mathrm{Gal}(\bar{K}/K)\colon \mathrm{Gal}(\bar{K}/L)]$, so the index of $\mathrm{Gal}(\bar{K}/L)$ is finite in $\mathrm{Gal}(\bar{K}/K)$. Now apply 14 to see that $\mathrm{Gal}(\bar{K}/L)$ is open. □

Proposition 16. *Let G be $\mathrm{Gal}(\bar{\mathbb{F}}_q/\mathbb{F}_q)$ where $q = p^n$ with p prime and $n \in \mathbb{N}^*$. Let x be an element of $\bar{\mathbb{F}}_q$, then the map*

$$\Phi_x\colon \begin{cases} \mathrm{Gal}(\bar{\mathbb{F}}_q/\mathbb{F}_q) & \to \bar{\mathbb{F}}_q \\ \sigma & \mapsto \sigma(x) \end{cases}$$

is continuous.

Proof. It is enough to show that for all $y \in \bar{\mathbb{F}}_q$, $\Phi_x^{-1}(\{y\})$ is open. If this is the case then the preimage of an open set is the union of preimages of each element of that set, which is a union of open set, hence open. This would prove the continuity of Φ_x. Now if $\Phi_x^{-1}(\{y\})$ is empty, then it is open. Suppose $\Phi_x^{-1}(\{y\}) \neq \emptyset$. Take $\sigma_0 \in \Phi_x^{-1}(\{y\})$ to be an arbitrary element. Then

$$\begin{aligned} \Phi_x^{-1}(\{y\}) &= \{\sigma \in \mathrm{Gal}(\bar{K}/K)\colon \sigma(x) = \sigma_0(x)\} \\ &= \{\sigma \in \mathrm{Gal}(\bar{K}/K)\colon (\sigma^{-1} \circ \sigma_0)(x) = x\} \\ &= \{\sigma \in \mathrm{Gal}(\bar{K}/K)\colon (\sigma^{-1} \circ \sigma_0) \in \mathrm{Gal}(\bar{K}/K(x))\} \\ &= \sigma_0 \mathrm{Gal}(\bar{K}/K(x)). \end{aligned}$$

Apply 15 to see that $\mathrm{Gal}(\bar{K}/K(x))$ is open. Now apply 13 to see that $\sigma_0\mathrm{Gal}(\bar{K}/K(x))$ is also open. This proves the continuity of Φ_x. □

1.2.2 Isogenies and Separability

This section aims to recall the notion of isogeny and separability, and to prove some basic results about separability. The main reference of this section is [22, II.2, II.4, III.4, III.5].

Morphisms and Separability

We take for granted the following results from [22, II.2].

Proposition 17. *Let $C \subseteq \mathbb{P}^n$ be a curve and let $V \subseteq \mathbb{P}^n$ be a variety. If C is smooth, then any rational map $\phi\colon C \to V$ is a morphism.*

Proof. See [22, II.2.1]. □

Theorem 18. *A morphism between curves is either constant or surjective.*

Proof. See [22, II.2.3]. □

Now, let C/K be a smooth curve and let $\mathscr{F}_K(C, \mathbb{P}^1)$ denote the set of rational maps from C to $\mathbb{P}^1$ defined over K, then the map

$$\Phi\colon \begin{cases} K(C) \cup \{\infty\} & \to \mathscr{F}_K(C, \mathbb{P}^1) \\ f & \mapsto [f, 1] \\ \infty & \mapsto [1, 0] \end{cases}$$

is a bijection, where $\Phi(f)$ can be given explicitly as

$$P \mapsto \begin{cases} [f(P), 1] & \text{if } f \text{ is regular at } P \\ [1, 0] & \text{otherwise} \end{cases}$$

for every $f \in K(C)$. This allows us to define the induced injection of function fields as follows.

Definition (Induced injection of function fields)**.** *Let C_1/K and C_2/K be curves and let $\phi\colon C_1 \to C_2$ be a nonconstant rational map defined over K. We write ϕ^* for the induced injection of function fields, defined by*

$$\phi^*\colon \begin{cases} K(C_2) & \to K(C_1) \\ f & \mapsto \Phi^{-1}(\Phi(f) \circ \phi). \end{cases}$$

Definition (Separability of a map)**.** *Let C_1/K and C_2/K be curves and let $\phi\colon C_1 \to C_2$ be a nonconstant rational map defined over K. We say that ϕ is separable, inseparable, or purely inseparable depending on the corresponding separability of the field extension $K(C_1)/\phi^*K(C_2)$. Also, we define the degree $\deg\phi$ to be the index $[K(C_1)\colon \phi^*K(C_2)]$ of the field extension.*

Proposition 19 ([22, II.2.6b])**.** *Let $\phi\colon C_1 \to C_2$ be a nonconstant map of smooth curves. Then for all but finitely many $Q \in C_2$, $\#\phi^{-1}(Q) = \deg_s(\phi)$.*

Proof. See [22, II.2.6b]. □

Differentials

Definition (Differential forms)**.** *Let C be a curve. For every $f \in \bar{K}(C)$, we attach a symbol d to construct $\mathrm{d}f$. This gives the set*

$$\{\mathrm{d}f\colon f \in \bar{K}(C)\}.$$

Now, we quotient out this set by the following rules of equivalence:

(i) $\mathrm{d}(x+y) \sim \mathrm{d}x + \mathrm{d}y$ *for all* $x, y \in \bar{K}(C)$.

(ii) $\mathrm{d}(xy) \sim x\mathrm{d}y + y\mathrm{d}x$ *for all* $x, y \in \bar{K}(C)$.

(iii) $\mathrm{d}a \sim 0$ *for all* $a \in \bar{K}$.

The remaining quotient set is called the space of (meromorphic) differential forms, and is denoted by Ω_C. It is a $\bar{K}(C)$-vector space.

We define $\mathrm{Der}_k(K, E)$ as the space of functions $D\colon K \to E$ satisfying

- For all $f, g \in K$, $D(f+g) = D(f) + D(g)$.
- For all $f, g \in K$, $D(fg) = fD(g) + gD(f)$.
- For all $a \in k$, $D(a) = 0_E$.

Proposition 20 (edited from [4])**.** $\mathrm{Hom}_K(\Omega_K, E) \cong \mathrm{Der}_k(K, E)$

Proof. Define $\phi\colon \mathrm{Der}_k(K,E) \to \mathrm{Hom}_K(\Omega_K, E)$ as $\phi(D) = \mathrm{d}f \mapsto D(f)$ for any $D \in \mathrm{Der}_k(K,E)$. It remains to check that $\phi(D)$ is a vector space homomorphism, i.e. linear over k. This is true because $\phi(D)(\mathrm{d}(f+g)) = D(f+g) = D(f) + D(g) = \phi(D)(\mathrm{d}f) + \phi(D)(\mathrm{d}g)$ and $\phi(D)(\lambda\mathrm{d}f) = \phi(D)(\mathrm{d}(\lambda f)) = D(\lambda f) = \lambda D(f) = \lambda\phi(D)(\mathrm{d}f)$, hence linear. □

Lemma 21 (edited from [18, II.3.4])**.** *If K' is a finite separable extension of K, and K is a finite extension of k, then the restriction $\mathrm{Der}_k(K', E) \to \mathrm{Der}_k(K, E)$ is bijective.*

Proof. By the primitive element theorem, $K' = K(\alpha)$ for some $\alpha \in K'$ and there exists a separable minimal polynomial $P \in K[x]$ such that $P(\alpha) = 0$. We write $P(x) = \sum_{i=0}^n a_i x^i$ and see that $P(\alpha) = 0$ gives $D(P(\alpha)) = 0$ for any $D \in \mathrm{Der}_k(K', E)$. This means that if $\bar{D}$ is an extension of D from K to K', then

$$\begin{aligned}\bar{D}(P(\alpha)) &= \bar{D}\left(\sum_{i=0}^n a_i\alpha^i\right)\\ &= \sum_{i=0}^n \bar{D}(a_i\alpha^i)\\ &= \sum_{i=1}^n (a_i\bar{D}(\alpha^i) + \alpha^i\bar{D}(a_i)) + \bar{D}(a_0).\end{aligned}$$

But $\bar{D}(x^i) = x\bar{D}(x^{i-1}) + x^{i-1}\bar{D}(x)$ for all $i \in \mathbb{N}_{\geq 2}$, so by simple induction we have $\bar{D}(x^i) = ix^{i-1}\bar{D}(x)$ for all $i \in \mathbb{N}_{\geq 1}$ and for all $x \in K'$. This gives

$$\bar{D}(P(\alpha)) = \sum_{i=1}^n \left(a_i i\alpha^{i-1}\bar{D}(\alpha) + \alpha^i\bar{D}(a_i)\right) + D(a_0) = 0.$$

Since P is separable, $P' \neq 0$, so $\sum_{i=1}^n a_i i x^{i-1} \neq 0$. Hence, we have

$$\bar{D}(\alpha) = \frac{-\sum_{i=0}^n \alpha^i\bar{D}(a_i)}{\sum_{i=1}^n a_i i\alpha^{i-1}}.$$

But $a_i \in K$ for all i so $\bar{D}(a_i)$ is actually $D(a_i)$. Hence,

$$\bar{D}(\alpha) = \frac{-\sum_{i=0}^{n} \alpha^i D(a_i)}{\sum_{i=1}^{n} a_i i \alpha^{i-1}}.$$

Now $\bar{D}$, exists uniquely at α. This actually extends to all of K' since each element x of K' can be written as

$$\sum_{i=0}^{[K':K]-1} \lambda_i \alpha^i$$

where $\lambda_i \in K$ for all i. This defines $\bar{D}(x)$ as

$$\sum_{i=1}^{[K':K]-1} \lambda_i i \alpha^{i-1} \bar{D}(\alpha) + \sum_{i=0}^{[K':K]-1} \alpha^i D(\lambda_i).$$

(as usual) which is uniquely defined depending on the value of $\bar{D}(\alpha)$. This completes the proof. □

Proposition 22 (from [22, II.4.2a])**.** *Let C be a curve. Then Ω_C is a 1-dimensional $\bar{K}(C)$-vector space.*

Proof. Let P be a nonsingular point on C and pick any uniformizer t of P in $\bar{K}(C)$. Observe that $\bar{K}(C)$ is a finite separable extension of $\bar{K}(t)$ ([22], II.1.4; the argument still holds when we replace K by $\bar{K}$ in the proof), so $\mathrm{Der}_{\bar{K}}(\bar{K}(C), E) \leftrightarrow \mathrm{Der}_{\bar{K}}(\bar{K}(t), E)$ by the previous lemma. Now, consider the map $\phi\colon \mathrm{Der}_{\bar{K}}(\bar{K}(t), E) \to E$ defined by $\phi(D) = D(t)$ for all D. This is injective because if $D_1(t) = D_2(t)$ then for all $y \in \bar{K}(t)$, y can be written as a rational function of polynomials over t (with coefficients and constants in $\bar{K}$), which, by the derivative rule, $D(y)$ can be defined in terms of $D(t)$ and so this gives $D_1(y) = D_2(y)$ for all $y \in \bar{K}(t)$ hence $D_1 = D_2$. Now clearly it is surjective since one can define $D(t) = \lambda$ for any $\lambda \in E$. Hence,

$$\mathrm{Der}_{\bar{K}}(\bar{K}(C), E) \leftrightarrow \mathrm{Der}_{\bar{K}}(\bar{K}(t), E) \leftrightarrow E$$

and observe that the bijection from $\mathrm{Der}_{\bar{K}}(\bar{K}(C), E)$ to E is linear, so $\dim_{\bar{K}} \mathrm{Der}_{\bar{K}}(\bar{K}(C), E) = \dim E = 1$. But $\mathrm{Hom}_{\bar{K}(C)}(\Omega_C, E) \cong \mathrm{Der}_{\bar{K}}(\bar{K}(C), E)$ so it has dimension one also. Now suppose $E = \bar{K}(C)$ and so this means $\dim_{\bar{K}(C)} \Omega_C = 1$. (Recall the basic fact that if V is a finite dimensional A-vector space, then $\mathrm{Hom}_A(V, A) \cong V$.) □

Proposition 23 (from [22, II.4.2b])**.** *Let C be a curve. Let $x \in \bar{K}(C)$. Then $\mathrm{d}x$ is a $\bar{K}(C)$-basis for Ω_C if and only if $\bar{K}(C)/\bar{K}(x)$ is a finite separable extension.*

Proof. ($\Rightarrow$) Suppose $\mathrm{d}x$ is a basis, then for all $\mathrm{d}f \in \Omega_C$, $\mathrm{d}f = \lambda \mathrm{d}x$ where $\lambda \in \bar{K}(C)$. As seen previously, $\bar{K}(C)/\bar{K}(t)$ is a finite separable extension, so it suffices to show that $\bar{K}(t)/\bar{K}(x)$ is a finite separable extension. Since $\bar{K}(C)/\bar{K}(t)/\bar{K}(x)/\bar{K}$, $\bar{K}(C)$ is finitely generated over $\bar{K}$, and has transcendental degree one, $\bar{K}(t)/\bar{K}(x)$ is a finite algebraic extension. We're left with showing that it is separable.

By contradiction suppose it is not separable, then there exists an element $a \in \bar{K}(t)$ with minimal polynomial $f \in \bar{K}(x)[y]$ such that $f(a) = 0$ and $\partial_y f(a) = 0$. For the sake of clarity, let us denote $L := \bar{K}(x)$ so that we can write $f \in L[y]$ peacefully. Suppose $f(y) = \sum_{i=0}^{n} l_i y^i$, where $l_i \in L$ and $l_n = 1$ (monic).

Consider $a \in L \subseteq \bar{K}(C)$ so $\mathrm{d}(a) = \lambda \mathrm{d}x$ for some $\lambda \in \bar{K}(C)$. And for each $i \in \{0 \dots n\}$ we have $l_i \in L$ so there exists $\mu_i \in \bar{K}(C)$ such that $\mathrm{d}(l_i) = \mu_i \mathrm{d}x$. Then,

$$0 = f(a) = \sum_{i=0}^{n} l_i a^i$$

implies

$$a^n = -\sum_{i=0}^{n-1} l_i a^i$$

so

$$\begin{aligned}
\mathrm{d}(a^n) &= \mathrm{d}\left(-\sum_{i=0}^{n-1} l_i a^i\right) \\
n a^{n-1}\mathrm{d}(a) &= -\sum_{i=0}^{n-1} \mathrm{d}(l_i a^i) \\
&= -\sum_{i=0}^{n-1} (l_i \mathrm{d}(a^i) + a^i \mathrm{d}(l_i)) \\
&= -\left(\mathrm{d}(a) \underbrace{\sum_{i=1}^{n-1} l_i i a^{i-1}}_{0 \text{ because } \partial_y f(a)=0} + \sum_{i=0}^{n-1} \mathrm{d}(l_i) a^i\right) \\
&= -\sum_{i=0}^{n-1} \mathrm{d}(l_i) a^i.
\end{aligned}$$

Now, by our hypothesis, suppose $\mathrm{d}(a) = \mu \mathrm{d}x$ and suppose $\mathrm{d}(l_i) = \lambda_i \mathrm{d}x$, where $\mu, \lambda_0, \ldots, \lambda_{n-1} \in \bar{K}(C)$, so that we arrive with

$$\left(n a^{n-1}\mu + \sum_{i=0}^{n-1} \lambda_i a^i\right) \mathrm{d}x = 0.$$

Since $\mathrm{d}x \neq 0$, we have

$$n a^{n-1}\mu + \sum_{i=0}^{n-1} \lambda_i a^i = 0. \qquad (\star)$$

Now, consider the claim that for any $f \in \bar{K}(x)$, if $\mathrm{d}f = \lambda \mathrm{d}x$ then $\lambda \in \bar{K}(x)$. We can see this by considering $P \in \bar{K}[x]$ first, where $P = \sum_{i=0}^{r} b_i x^i$ where $b_0, \ldots, b_r \in \bar{K}$. We have

$$\mathrm{d}(P) = \sum_{i=0}^{r} (b_i \mathrm{d}(x^i) + \mathrm{d}(b_i) x^i) = \sum_{i=1}^{r} b_i i x^{i-1} \mathrm{d}x + \sum_{i=0}^{r} \underbrace{\mathrm{d}(b_i)}_{0} x^i = \mathrm{d}x \sum_{i=1}^{r} b_i i x^{i-1}.$$

Suppose $d(P) = \lambda \mathrm{d}x$ then since $\mathrm{d}x \neq 0$, $\lambda = \sum_{i=1}^{r} b_i i x^{i-1} \in \bar{K}(x)$. This proves that $\lambda \in \bar{K}(x)$ for any $P \in \bar{K}[x]$. Now, if $P \in \bar{K}(x)$, $0 = d(1) = d(P\frac{1}{P}) = Pd\left(\frac{1}{P}\right) + \frac{1}{P}d(P)$ so $d\left(\frac{1}{P}\right) = -\frac{1}{P^2}d(P) \in \bar{K}(x)$. Now it is easy to see that for any fraction $\frac{P}{Q} \in \bar{K}(x)$ where $P, Q \in \bar{K}[x]$, $d\left(\frac{P}{Q}\right) = Pd\left(\frac{1}{Q}\right) + \frac{1}{Q}d\left(\frac{1}{P}\right) \in \bar{K}(x)$. This proves the claim.

Going back to our equation $(\star)$, we now know that $\mu, \lambda_0, \ldots, \lambda_{n-1} \in \bar{K}(x) = L$, so we obtain a polynomial with degree $n-1$ that vanishes at a. This contradicts the assumption that the minimum polynomial f is of degree n, and hence proves that $\bar{K}(t)/\bar{K}(x)$ is separable, therefore $\bar{K}(C)/\bar{K}(x)$ is a finite separable extension.

($\Leftarrow$) Suppose $\bar{K}(C)/\bar{K}(x)$ is a finite separable extension. Then, by the primitive element theorem, we may write $\bar{K}(C) = \bar{K}(x, \alpha)$ for some $\alpha \in \bar{K}(C)$, and that the minimal polynomial for α in $L[y]$ (where $L = \bar{K}(x)$) exists and is separable. So one can write $\sum_{i=0}^{n} l_i \alpha^i = 0$, and this gives

$$\underbrace{\sum_{i=1}^{n} l_i i \alpha^{i-1}}_{\neq 0 \text{ due to separability}} \mathrm{d}(\alpha) + \sum_{i=0}^{n} \underbrace{\mathrm{d}(l_i)}_{\text{can be written as a multiple of } \mathrm{d}x} \alpha^i = 0$$

so $\mathrm{d}(\alpha)$ can also be written as a multiple of $\mathrm{d}x$. Now, for each $f \in \bar{K}(C)$, we can write f as an algebraic combination of objects in $\bar{K}[x, \alpha]$, so $\mathrm{d}f$ is an algebraic combination of objects in the form $\lambda \mathrm{d}x$ for some $\lambda \in \bar{K}(C)$, hence the sum is also in the form $\lambda' \mathrm{d}x$ for maybe some other $\lambda' \in \bar{K}(C)$. This means every element in Ω_C can be written as a $\bar{K}(C)$-multiple of $\mathrm{d}x$, hence $\mathrm{d}x$ is a basis of Ω_C. This completes the proof. □

Now, if $\phi\colon C_1 \to C_2$ is a nonconstant map of curves, then $\phi^*\colon \bar{K}(C_2) \to \bar{K}(C_1)$ induces another map which is also denoted by ϕ^*, defined by

$$\begin{cases} \Omega_{C_2} & \to \Omega_{C_1} \\ \sum f_i \mathrm{d}x_i & \mapsto \sum (\phi^* f_i) \mathrm{d}(\phi^* x_i). \end{cases}$$

Proposition 24 (from [22, II.4.2c]). *Let $\phi\colon C_1 \to C_2$ be a nonconstant map of curves. Then ϕ is separable if and only if the map $\phi^*\colon \Omega_{C_2} \to \Omega_{C_1}$ is injective (equivalently, nonzero).*

Proof. See [22, II.4.2c]. □

On Elliptic Curves

Let us now go back into the context of elliptic curves.

Definition (Isogeny). *Let E_1 and E_2 be elliptic curves. A morphism $\phi\colon E_1 \to E_2$ is said to be an isogeny if $\phi(O) = O$. The set of isogenies from E_1 to E_2 is denoted by $\operatorname{Hom}(E_1, E_2)$.*

Remark. *Nonzero isogenies are nonconstant isogenies since if ϕ is zero then it is also constant in particular, and if it is constant, by definition that $\phi(O) = O$, it must therefore be zero.*

Theorem 25. *Let $\phi\colon E_1 \to E_2$ be a nonzero isogeny. If ϕ is separable then $\#\ker\phi = \deg\phi$.*

Proof. This proof follows the proof of [22, III.4.10a]. First, let us show that $\#\phi^{-1}(Q)$ are equal for all $Q \in E_2$. Well pick any $Q, Q' \in E_2$ and let R be a point in $\phi^{-1}(Q' - Q)$ (which is nonempty since ϕ is surjective because it is nonconstant and due to 18). Now the map

$$\begin{cases} \phi^{-1}(Q) & \to \phi^{-1}(Q') \\ P & \mapsto P + R \end{cases}$$

is clearly injective. Now if $P' \in \phi^{-1}(Q')$ then $P' - R \in \phi^{-1}(Q)$ because $\phi(P' - R) = \phi(P') - \phi(R) = Q' - (Q' - Q) = Q$. So for each $P' \in \phi^{-1}(Q')$, $P' - R$ is a preimage, hence the map is surjective. This gives a bijection between $\phi^{-1}(Q)$ and $\phi^{-1}(Q')$ for all $Q, Q' \in E_2$ and so $\#\phi^{-1}(Q)$ are all equal. By 19, all (except finite number of points) $Q \in E_2$ satisfies $\#\phi^{-1}(Q) = \deg_s(\phi)$ so at least one satisfies this in particular. This means $\#\phi^{-1}(Q) = \deg_s(\phi)$ is actually true for all $Q \in E_2$, and since ϕ is separable, $\deg_s(\phi) = \deg(\phi)$. This proves that $\#\ker\phi = \#\phi^{-1}(O) = \deg_s(\phi) = \deg(\phi)$. □

We now define something called the invariant differential.

Definition. *Let E/K be an elliptic curve given by the usual Weierstrass equation*

$$E\colon y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6.$$

We define the invariant differential

$$\omega := \frac{\mathrm{d}x}{2y + a_1x + a_3} \in \Omega_E.$$

When E is clear from context, we just write ω to denote the invariant differential for E.

Theorem 26 ([22, III.5.2]). *Let E, E' be elliptic curves, let ω be the invariant differential on E, and let $\phi, \psi\colon E \to E'$ be isogenies. Then*

$$(\phi + \psi)^*\omega = \phi^*\omega + \psi^*\omega.$$

Proof. See [22, III.5.2]. □

Theorem 27 (special case of [22, III.5.5]). *Let $E/\mathbb{F}_q$ be an elliptic curve over a finite field $\mathbb{F}_q$ of characteristic p. Let $\phi\colon E \to E$ be the Frobenius morphism sending (x, y) to (x^q, y^q). Then $(1 - \phi)\colon E \to E$ is a separable map.*

Remark. *The elliptic curve $E/\mathbb{F}_q$ can be written as*

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

with $a_1, a_2, a_3, a_4, a_6 \in \mathbb{F}_q$. Then the Frobenius morphism $\phi\colon E \to E^{(q)}$ is actually defined from E to $E^{(q)}$ where

$$E^{(q)}\colon \operatorname{Frob}_q(y)^2 + a_1\operatorname{Frob}_q(x)\operatorname{Frob}_q(y) + a_3\operatorname{Frob}_q(y) = \operatorname{Frob}_q(x)^3 + a_2\operatorname{Frob}_q(x)^2 + a_4\operatorname{Frob}_q(x) + a_6.$$

But $a_1, a_2, a_3, a_4, a_6 \in \mathbb{F}_q$ so it is fixed by Frob_q, i.e., the equation $E^{(q)}$ is equivalent to

$$\operatorname{Frob}_q(y^2 + a_1xy + a_3y) = \operatorname{Frob}_q(x^3 + a_2x^2 + a_4x + a_6).$$

Since Frob_q is injective, we conclude that $E^{(q)} = E$, and so ϕ is well-defined as an endomorphism from E to itself and so we can actually write $\phi\colon E \to E$.

We delay the proof of 27 because we will need some results on the Frobenius map first.

1.2.3 The Frobenius Map on Curves

Generalizing the remark after the theorem 27, we discuss some basic results of the Frobenius map on curves here. If K is a field with characteristic $p \neq 0$. Let $q = p^r$ for some $r \in \mathbb{N}^*$. For any polynomial $f \in K[X_0, \ldots, X_n]$, we denote by $f^{(q)}$ the polynomial obtained by raising all coefficients to the power of q. Then, for any curve C/K, the projective variety generated by the ideal generated by the set

$$\{f^{(q)} \colon f \in I(C)\}$$

is a curve, denoted by $C^{(q)}/K$. We define the Frobenius map on curves as a morphism given by

$$\phi \colon \begin{cases} C & \to C^{(q)} \\ [x_0, \ldots, x_n] & \mapsto [x_0^q, \ldots, x_n^q] \end{cases}.$$

We state a useful result from [22, II.2.11] as follows.

Proposition 28. *Let K be a field of characteristic $p > 0$, let $q = p^r$ for some $r \in \mathbb{N}^*$, let C/K be a curve, and let $\phi \colon C \to C^{(q)}$ be the Frobenius morphism. Then ϕ is inseparable and $\deg \phi = q$.*

Proof. See [22, II.2.11]. □

Proof of Theorem 27. Note that we follow [22, III.5.5] but simplify for only this special case. Recall 24 and let us prove that the map $(1 - \phi)^* \colon \Omega_E \to \Omega_E$ is nonzero. By 26,

$$(1 - \phi)^*(\omega) = \omega - \phi^*(\omega).$$

Since ϕ is inseparable by 28, $\phi^*(\omega) = 0$, so $(1 - \phi)^*(\omega) = \omega$, i.e. $(1 - \phi)^*$ is not the zero map. □

1.2.4 The Dual Isogeny

Later on, we rely on the existence of a dual for each isogeny to prove some useful results. Consider the following theorem.

Theorem 29 (Dual isogeny)**.** *Let E_1 and E_2 be elliptic curves with an isogeny $\phi \colon E_1 \to E_2$. Then there exists a unique isogeny $\hat{\phi} \colon E_2 \to E_1$ satisfying $\hat{\phi} \circ \phi = [\deg \phi]$. We call this $\hat{\phi}$ the* dual isogeny *of ϕ and later on we use the symbol $\hat{}$ to denote the dual of any isogeny.*

Proof. See [22, III.6.1a]. An explicit construction is given by [22, III.6.1b]. □

Theorem 30. *For all $m \in \mathbb{Z}$, on any elliptic curve, $\deg[m] = m^2$.*

Proof. See [22, III.6.2]. □

1.2.5 A form of The Cauchy–Schwarz Inequality

In this context, we use the generalized notion of a "positive definite quadratic form" using the following definition.

Definition. *Let A be an abelian group. A map $d \colon A \to \mathbb{Z}$ is said to be a quadratic form if*

(i) *$d(\alpha) = d(-\alpha)$ for all $\alpha \in A$;*

(ii) *The pairing*

$$(\cdot, \cdot) \colon \begin{cases} A \times A & \to \mathbb{R} \\ (\alpha, \beta) & \mapsto d(\alpha + \beta) - d(\alpha) - d(\beta) \end{cases}$$

is bilinear.

Furthermore, it is said to be positive definite if it satisfies:

(iii) *$d(\alpha) \geq 0$ for all $\alpha \in A$;*

(iv) *$d(\alpha) = 0$ if and only if $\alpha = 0$.*

Theorem 31. *Let E_1 and E_2 be elliptic curves. The degree map*

$$\deg \colon \mathrm{Hom}(E_1, E_2) \to \mathbb{Z}$$

is a positive definite quadratic form.

Proof. See [22, III.6.3]. □

Lemma 32 (a form of Cauchy–Schwarz, [22, V.1.2])**.** *Let A be an abelian group and let $d\colon A \to \mathbb{Z}$ be a positive definite quadratic form. Then $|d(\psi-\phi)-d(\psi)-d(\phi)| \leq 2\sqrt{d(\phi)d(\psi)}$ for all $\psi, \phi \in A$.*

Proof. The following is edited from [22, V.1.2]. Let L be the bilinear pairing defined by

$$\begin{cases} A \times A & \to R \\ (\alpha, \beta) & \mapsto d(\alpha-\beta)-d(\alpha)-d(\beta) \end{cases}.$$

Since d is positive definite, for all $m, n \in \mathbb{Z}$, for all $\psi, \phi \in A$, we have

$$0 \leq d(m\psi - n\phi) = L(m\psi, n\phi) + d(m\psi) + d(n\phi) = mnL(\psi,\phi) + m^2 d(\psi) + n^2 d(\phi).$$

Now take $m = -L(\psi,\phi)$ and $n = 2d(\psi)$ to see that

$$0 \leq -2d(\psi)L(\psi,\phi)^2 + d(\psi)L(\psi,\phi)^2 + 4d(\psi)^2 d(\phi) = d(\psi)(4d(\psi)d(\phi) - L(\psi,\phi)^2).$$

Well if $d(\psi) = 0$ then $\psi = 0$ and $|d(\psi-\phi)-d(\psi)-d(\phi)| = |d(-\phi)-d(\phi)| = 0 = 2\sqrt{d(\phi)d(\psi)}$ proves the inequality. Otherwise $d(\psi) > 0$ and so $0 \leq 4d(\psi)d(\phi) - L(\psi,\phi)^2$. This proves the inequality. □

1.2.6 The Proof

We piece together the supporting claims to prove the following theorem.

Theorem 33 (Hasse's bound)**.** *Let $E/\mathbb{F}_q$ be an elliptic curve defined over a finite field. Then*

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

Proof. The following proof tries to simplify the one given in [22, V.1.1]. Choose a Weierstrass equation for E with coefficients in $\mathbb{F}_q$, and let $\phi\colon E \to E$ be the Frobenius map $(x,y) \mapsto (x^q, y^q)$. This is well-defined by the remark after Theorem 27. Now, recall the basic facts from Galois theory [9] that for any point $P = (x_P, y_P) \in E(\bar{\mathbb{F}}_q)$, $P \in E(\mathbb{F}_q)$ if and only if x_P and y_P are fixed by all of $\mathrm{Gal}(\bar{\mathbb{F}}_q/\mathbb{F}_q)$. Now suppose $x \in \bar{\mathbb{F}}_q$ and let Φ be the continuous map (16 proves the continuity)

$$\begin{cases} \mathrm{Gal}(\bar{\mathbb{F}}_q/\mathbb{F}_q) & \to \bar{\mathbb{F}}_q \\ \sigma & \mapsto \sigma(x) \end{cases}.$$

If $\psi \in \mathrm{Gal}(\bar{\mathbb{F}}_q/\mathbb{F}_q)$ is a map that permutes x away from itself then $\psi(x) \neq x$, so $\psi \in \Phi^{-1}(\bar{\mathbb{F}}_q \setminus \{x\})$, which is open by continuity of Φ. By 9, the set $\Phi^{-1}(\bar{\mathbb{F}}_q \setminus \{x\})$ hits ϕ^k for some $k \in \mathbb{N}^*$. This means $\phi^k(x) \neq x$ and so it is impossible that $\phi(x) = x$. What we've shown so far is that if ϕ fixes x then surely every one of $\mathrm{Gal}(\bar{\mathbb{F}}_q/\mathbb{F}_q)$ fixes x also. Conversely, if $x \in \mathbb{F}_q$ then obviously all of $\mathrm{Gal}(\bar{\mathbb{F}}_q/\mathbb{F}_q)$ fix x. This allows us to see that

$$P \in E(\mathbb{F}_q) \text{ if and only if } \phi(P) = P \text{ if and only if } (1-\phi)(P) = O.$$

So $E(\mathbb{F}_q) = \ker(1-\phi)$. The result from 27 tells us that $(1-\phi)$ is separable. Now apply 25 to see that $\#\ker(1-\phi) = \deg(1-\phi)$. This proves that $\#E(\mathbb{F}_q) = \deg(1-\phi)$. Now we apply 31 to see that deg is a positive definite quadratic form and then apply 32 to see that

$$|\deg(1-\phi) - \deg([1]) - \deg(\phi)| \leq 2\sqrt{\deg([1])\deg(\phi)}.$$

Using the fact that $\deg([1]) = 1$ (trivial by definition) and $\deg(\phi) = q$ (by 28), we conclude that

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

This completes the proof. □

1.3 Integer Factorization

We continue to an important unsolved problem in computer science: Can integer factorization be done in polynomial time? There is a very interesting line of research trying to improve and analyze integer factorization algorithms. It is interesting in its own right, but we will not be focusing on that here. Instead, we will give a short introduction and consider how elliptic curves may be used to help with integer factorization.

1.3.1 The RSA Cryptosystem

The RSA cryptosystem is one of the first asymmetric cryptosystem. We begin with the method for the key generation.

1. Choose two prime numbers p and q. Let $n = pq$ and $m = (p-1)(q-1)$.
2. Choose an integer e such that $1 < e < m$ and $\gcd(e, m) = 1$.
3. Compute d such that $1 < d < m$ and $de \equiv 1 \pmod{m}$. (This requires proving that such d exists and can be computed easily, which will be supplied next.)

Proof that d exists and can be computed easily. Since $\gcd(e, m) = 1$, by Bézout's identity, there exists a solution $(x, y) \in \mathbb{Z} \times \mathbb{Z}$ to $ex + my = 1$. Such a solution can be taken so that $0 \leq x < m$. Algorithmically, this can be done using the extended Euclidean algorithm on (e, m) in polynomial time. Once we found such x, then we take $d := x \bmod m$ so that $1 < d < m$ and $de = 1 - my'$ for some $y' \in \mathbb{Z}$ means $de \equiv 1 \pmod{m}$. □

The public key is the pair (n, e), and the private key is d. Suppose Alice generates the key and Bob wants to send some information to Alice, then first Alice publishes the pair (n, e). Now suppose Bob wants to encrypt a message M, which is an integer such that $0 \leq M < n$. Bob then computes $c := M^e \bmod n$ and sends c to Alice.

Now suppose Alice wants to decrypt the ciphertext c. Alice can compute $c^d \bmod n$. Since $c^d = (M^e)^d = M^{de} \equiv M \pmod{n}$, we see that $c^d \bmod n = M$. This allows the Alice to recover Bob's message. Note that the congruence $M^{de} \equiv M \pmod{n}$ requires a little bit of elementary number theory: Since $de \equiv 1 \pmod{m}$ suppose $de - 1 = mk$ for some $k \in \mathbb{Z}$. Then $M^{de} = M \cdot M^{mk}$. But $M^m = M^{(p-1)(q-1)} \equiv 1 \pmod{n}$. This is due to Euler's theorem and the fact that $\varphi(n) = \varphi(pq) = (p-1)(q-1)$. This completes the description of the cryptosystem.

The hard part is that if Eve, an eavesdropper, wants to recover Bob's message knowing only c, e, n, is it possible to deduce M? It turns out that it is actually possible, but practically time-consuming, to recover M. This is called the RSA problem.

Problem (RSA Problem)**.** *Given $c, e, n \in \mathbb{Z}$ such that $0 \leq c < n$, $1 < e < m$, $\gcd(e, m) = 1$. Suppose that there exists (unknown) primes p, q such that $n = pq$, and there exists (unknown) integer message $0 \leq M < n$ such that $M^e \bmod n = c$. The problem is to recover such M.*

It is unknown to this day of writing this, whether there exists an algorithm to solve the RSA problem in polynomial time or not. However, if integer factorization can be done in polynomial time, then one can use such algorithm to factorize n into p and q very quickly. Knowing p and q, one can easily compute m and d, and if Eve manages to do this, Eve can recover the message using Alice's decryption algorithm directly.

Remark. *A common misleading oversimplification is the quote "RSA relies on the hardness of integer factorization". In fact, if one can solve integer factorization in polynomial time, then the RSA problem can be solved in polynomial time. We actually don't know about the other way around, and it is reasonably suggested that the RSA problem might actually be easier than integer factorization, since we only need to factor semiprimes (product of two primes).*

In the next subsections, we will consider how to factorize integers more efficiently than brute-forcing all the possible factors. Note that all of these algorithms are still called "inefficient" because they are not done in polynomial time.

1.3.2 Pollard's $p - 1$ algorithm

A very basic, obvious, and naive algorithm to factorize an integer $N \in \mathbb{N}$ is to loop over all integers $i = 1, 2, \ldots, N$ and check if $N \bmod i = 0$. A slight improvement is to loop only over primes, and loop only at most $\sqrt{N}$ (since if $d \mid N$ then $\frac{N}{d} \mid N$ also). These improvements leave the algorithm to be run in $O(\sqrt{N})$ (which is not bounded by any polynomial of $\log_2 N$). In this section, we describe Pollard's $p - 1$ algorithm, which runs successfully with high probability, in $O(L \log L \log^2 N)$ with a conveniently chosen value of L.

Consider the following algorithm (see Algorithm 1).

Let us analyze the algorithm. If at some point of the inner for-loop, $1 < F < N$, then by the definition, since $F = \gcd(A - 1, N)$, F divides N and so it is a nontrivial factor of N. Suppose we ignore the line $A \leftarrow A^i \bmod N$ and replace it with just $A \leftarrow A^i$ first. Now there are two other cases ([6, Section 5.4]):

Algorithm 1 Pollard's $p-1$ Algorithm

Require: $N \geq 2$, $L \in \mathbb{N}^*$
Ensure: F is a prime factor of N, or failure
  $a \leftarrow$ a random integer between 2 and $N-1$
  $F \leftarrow \gcd(a, N)$
  if $1 < F < N$ **then** success: return F and halt
  end if
  $A \leftarrow a$
  for $i = 1, \ldots, L$ **do**
    $A \leftarrow A^i \bmod N$
    $F \leftarrow \gcd(A-1, N)$
    if $1 < F < N$ **then** success: return F and halt
    else if $F = N$ **then** failure: restart the algorithm by selecting a again
    end if
  end for
  If no success has been returned yet, then failure.

1. $F = N$. This means $\gcd(A-1, N) = N$, i.e. N divides $A-1$. But at this point $A = a^{i!}$, so $N \mid a^{i!} - 1$. Even if we increment i, we see that $a^{(i+1)!} - 1 = (a^{i!})^{i+1} - 1$, which is a multiple of $a^{i!} - 1$, which is already a multiple of N. This means there is no point in increasing i: the later $\gcd(A-1, N)$ will always be N and $F = N$ will be a trivial factor of N.

2. $F = 1$. This means $\gcd(A-1, N) = 1$. Recall that at this point $A = a^{i!}$, so $\gcd(a^{i!}-1, N) = 1$. And since $a^{i!} - 1$ is a factor of $a^{(i+1)!} - 1$, we hope that by incrementing i, extra factors would appear, so we don't stop at this point but rather increase i.

Now we see that even if we replace back that line with $A \leftarrow A^i \bmod N$, the arguments still work in the same way.

Consider the integer $N \in \mathbb{N}^*$ which is to be factorized. Suppose there is a prime p which is a factor of N. Write

$$p - 1 = \prod_{i=1}^{t} q_i^{e_i}.$$

and let $L := \max_{1 \leq j \leq t} e_j q_j$. We claim [22, XI.2.1] that the case $F = 1$ will eventually be gone. This is captured in the following proposition.

Proposition 34. *Let $N \in \mathbb{N}^*$ be the input integer, which is divisible by a prime p where*

$$p - 1 = \prod_{i=1}^{t} q_i^{e_i}.$$

Let

$$L := \max_{i \leq j \leq t} e_j q_j.$$

Then the algorithm 1 will not fail on the second case, i.e., in the end it is either the case of $1 < F < N$ (success) or $F = N$ (failure), but never $F = 1$.

Proof. Consider at $i = L$, we have $A = a^{L!} \bmod N$. It follows that by the definition of L, the quantity $q_j^{e_j}$ divides $L!$ for each $1 \leq j \leq t$, so $p-1 \mid L!$. By Fermat's little theorem, $a^{p-1} \equiv 1 \bmod p$ so $a^{L!} \equiv 1 \bmod p$ also. This means p is a factor of $a^{L!} - 1$. But by assumption, we knew that p is a factor of N also, so $\gcd(a^{L!} - 1, N) \neq 1$ for sure, and $\gcd(a^{L!} - 1, N) = \gcd((a^{L!} \bmod N) - 1, N)$, so the case of $F = 1$ cannot happen at $i = L$. □

Now, the only case of failure is when $F = N$. For most values of N, this is unlikely. Here we attach an experimental evidence in figure 1.1 to show that this is not always true, but for a random N, it looks fine.

In particular, for $N = 2047 = 23 \times 89$, the probability of failure can be up to ≥ 0.78. In the next section, we will discuss the ECM, which can be thought of a generalization of Pollard's $p-1$ algorithm but instead of using the group $\mathbb{F}_p^\times$ we use the groups given by elliptic curves instead.

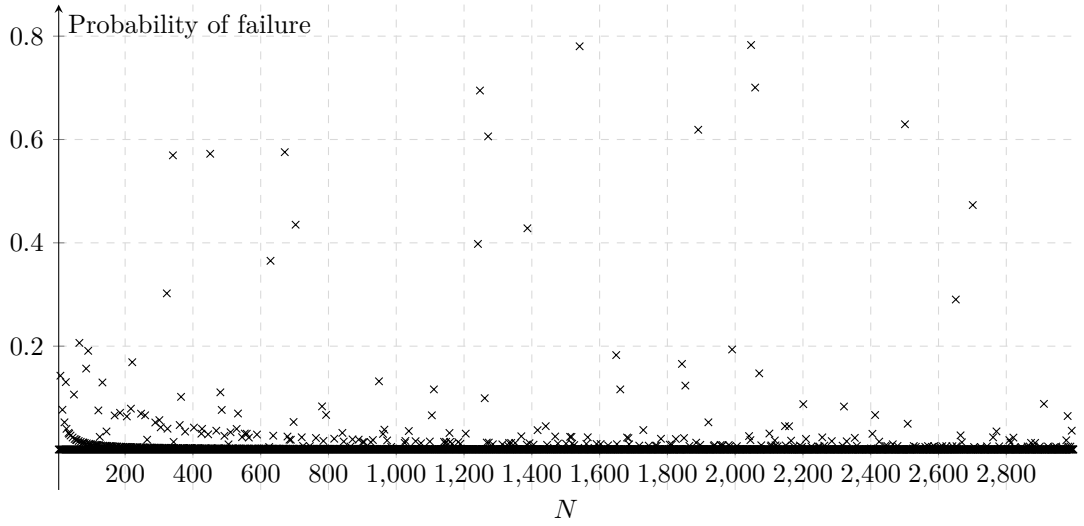

Figure 1.1: Probability of failure of Pollard's $p-1$ algorithm on composite numbers N (upon randomizing $a \in \{2, \ldots, N-1\}$)

Algorithm 2 Lenstra's elliptic curve factorization

Require: $N \geq 2$, $L \in \mathbb{N}^*$
Ensure: F is a prime factor of N, or failure
  $E \leftarrow$ an elliptic curve over $\mathbb{Z}/N\mathbb{Z}$
  $P \leftarrow$ a point on $E(\mathbb{Z}/N\mathbb{Z})$
  $Q \leftarrow P$
  for $i = 2, \ldots, L$ **do**
    $Q \leftarrow [i]Q$
    if during the computation of $[i]Q$, an inverse of a noninvertible element of $a \in \mathbb{Z}/N\mathbb{Z}$ is needed **then**
      $F \leftarrow \gcd(a, N)$
      if $1 < F < N$ **then** success: return F and halt
      else if $F = N$ **then** failure: restart the algorithm by selecting E, P again
      end if
    end if
  end for
  If no success has been returned yet, then failure.

1.3.3 Lenstra's elliptic curve factorization algorithm (ECM)

Instead of working over $\mathbb{F}_p^\times$, we work over $E(\mathbb{F}_p)$ instead. This is captured by the following algorithm (Algorithm 2).

As before, we see that if there is a prime factor p of N such that $\#E(\mathbb{F}_p) = \prod_{j=1}^{t} q_t^{e_t}$ so that $L \geq \max_{1\leq j\leq t} q_t e_t$, then $\#E(\mathbb{F}_p) \mid L!$, so the computation of $[L!]P$ can be factored as

$$\left[\frac{L!}{\#E(\mathbb{F}_p)}\right]([\#E(\mathbb{F}_p)]P) = \left[\frac{L!}{\#E(\mathbb{F}_p)}\right](O) = O.$$

This means, for such high enough L, there cannot be failure by succeeding in computing $[L!]P$; during the loop $i = 2, \ldots, L$ there must be at least once that $[i]Q$ requires inverting a noninvertible element of $\mathbb{Z}/N\mathbb{Z}$. So the only failure is when, upon seeing that noninvertible element a and letting $F = \gcd(a, N)$, we have $F = N$. Such case is improbable as of figure 1.1, but in this Lenstra's method, when this happens we can completely change the initial curve E. Even if $p \mid N$ is fixed, we're allowed to vary E so that there are multiple possibilities for $\#E(\mathbb{F}_p)$. This gives a heuristic time complexity of $\exp((\sqrt{2}+o(1))\sqrt{\log p \log\log p})$ [13]. However, almost all of the arguments can be made rigorous [6, Section 7.4.1, pg. 339] except the following, which is still a conjecture given by [13, (2.9)].

Conjecture 1. *For a real number $x > e$, define*

$$L(x) = \exp(\sqrt{\log x \log\log x}).$$

Let α be a strictly positive real number, then the probability that a random integer s inside the interval $(x + 1 - \sqrt{x}, x + 1 + \sqrt{x})$ has all its prime factors bounded from above by $L(x)^\alpha$ is $L(x)^{-\frac{1}{2\alpha}+o(1)}$, for $x \to +\infty$.

Remark. *For the implementation details, choosing a random elliptic curve followed by a point on it is actually a bit complicated. We proceed with choosing an integer $0 \leq A < N$ first and a point $P = (x_0, y_0)$ where $0 \leq x_0, y_0 < N$. Then let $B = y_0^2 - x_0^3 - Ax_0$, so that E can be written as $y^2 = x^3 + Ax + B$ over $\mathbb{Z}/N\mathbb{Z}$.*

Also note that $\mathbb{Z}/N\mathbb{Z}$ is not necessarily a field, so the general theory of elliptic curves doesn't work, but we don't need to worry too much. In fact, the computations of multiplication-by-m map can be done using the same usual formula for elliptic curves, where we would handle division by a noninvertible element in the computation.

Let us end with the remark that ECM is still currently not the best factorization algorithm, but is very good in practice for factorizing smooth integers. There are other factorization algorithms that would be better suited for the general case, such as the quadratic sieve (QS) and the general number field sieve (GNFS). However, they are beyond the scope of this thesis.

1.4 The Discrete Logarithm Problem

1.4.1 The ElGamal Cryptosystem

We fix a group $(G, \cdot)$ and let us consider the following description.

1. Alice and Bob agree on an element $C \in G$.
2. Alice select $a \in \mathbb{N}$, which is the secret key, and computes $A := C^a$, which is the public key.
3. Bob wants to encrypt a message $M \in G$. Bob choose a random $k \in \mathbb{N}$.
4. Bob computes
$$B_1 = C^k \qquad \text{and} \qquad B_2 = MA^k.$$
5. Bob sends (B_1, B_2) through an insecure communication line.
6. Seeing that $B_1^a = (C^k)^a = C^{ak} = (C^a)^k = A^k$, one can compute M as $M = MA^k(A^k)^{-1} = B_2(B_1^a)^{-1}$. Alice can then compute $B_2(B_1^a)^{-1}$ to recover M.

Now we see that this is obviously correct. But if Eve wants to recover M knowing only (C, A, B_1, B_2), then a must be recovered. This information can be deduced from knowing A and C. One can check the sequence $C, C^2, \ldots$ until $C^e = A$ for some e. Now even if e may not be equal to a, but $A = C^e$ would imply that $B_1^e = (C^k)^e = (C^e)^k = A^k$ and so $B_2(B_1^e)^{-1}$ is still equal to M. Hence, the problem of Eve is to find an integer e such that $C^e = A$. This gives the problem in the next section.

1.4.2 Discrete Logarithm Problem (DLP) and ECDLP

Consider the discrete logarithm problem, which is related to Eve's problem in the previous description of the ElGamal cryptosystem.

Problem (Discrete logarithm problem (DLP)). *Given a group G and two elements $x, y \in G$ such that $y \in \langle x \rangle$, then there exists $m \in \mathbb{N}$ such that $y = x^m$. The problem is to recover such m (even if there are multiple possibilities, any such m is fine), given (x, y), knowing G beforehand.*

Following this, we will denote by DLP_G the DLP for the group G, and write $m = \mathrm{DLP}_G(x, y)$ as an answer to the DLP.[4]

In fact, we've proved that if Eve can solve DLP in polynomial time, then the knowledge (C, A, B_1, B_2) can be decrypted in polynomial time into M too as we can compute $e = \mathrm{DLP}_G(C, A)$ so that $M = B_2(B_1^e)^{-1}$. The converse is not yet known. In order to be precise, let us also define another problem.

Problem (Diffie–Hellman problem (DHP)). *Knowing the group G, suppose $g \in G$ and that there exists $x, y \in \mathbb{N}$ such that $A = g^x$ and $B = g^y$. Knowing the values of A, B and g, but not x nor y, the problem is to find $g^{xy} \in G$.*

Following this, we will denote by DHP_G the DHP for the group G, and write $C = \mathrm{DHP}_G(A, B, g)$ as the answer to the DHP.[5]

Indeed, in the ElGamal cryptosystem, Eve's problem is actually a variant of DHP: the problem of knowing (C, A, B_1, B_2) and wanting to recover M is actually equivalent to $\mathrm{DHP}_G(A, B_1, C)$, because once we know the answer D to $\mathrm{DHP}_G(A, B_1, C)$, we see that since $B_1 = C^k$ and $A = C^a$, then D must be C^{ak}, so $B_2 D^{-1} = MA^k C^{-ak} = M$. What we've proved is that if we can easily solve DLP, then we can also easily solve DHP and the converse is not yet known. That being said, if one could solve DHP easily, then the ElGamal cryptosystem could've been broken as easily. In other words, we say that the ElGamal cryptosystem relies on the hardness of DHP, which, as same as RSA, has not yet been proved nor disproved as of now.

However, one can look at the problem in a more special cases of G. Consider the case where G is an additive group $(\mathbb{Z}, +)$. DHP says that suppose we know $g \in \mathbb{Z}$ and $xg, yg \in \mathbb{Z}$, find $xyg \in \mathbb{Z}$. This can be easily solved using basic division: compute $x = \frac{xg}{g}$ and $y = \frac{yg}{g}$, then compute the product xyg. Hence, the problem $\mathrm{DHP}_{(\mathbb{Z},+)}$ is easy, i.e., all can be done in $O(\log^C(xyg))$ time for a constant C (depending on multiplication and division algorithms, but surely $C \leq 4$).

The same can be said when G is the additive group $(\mathbb{Z}/q\mathbb{Z}, +)$ when $q = p^r$ is a prime power. Now things become a lot more complicated when one looks at $\mathbb{F}_q^\times$. However, this is still much easier than the general case. There is an algorithm called index calculus that could be used to solve DLP (and therefore DHP) for $\mathbb{F}_q^\times$ in subexponential time. This is unfortunately also beyond the scope of the thesis, see [1] for more information.

Now, since G can be any group, why not let G be the group of an elliptic curve? In fact, by trying $G = E(\mathbb{F}_q)$, it is much harder than the previous cases. As of now, there is no algorithm better than $O(\sqrt{q})$ to solve DLP for elliptic curves in general. There are, however, algorithms to reduce some instances into $\mathbb{F}_{q^d}^\times$. See [14] for MOV algorithm. This is why the ElGamal cryptosystem with elliptic curves as the base group are considered for applications in cryptography. Note that applying ElGamal cryptosystem directly are too unsafe for cryptographic attacks, so in practice, one would use a combined method such as the Integrated Encryption Scheme (IES). Note that we call the DLP over $E(\mathbb{F}_q)$ as "Elliptic curve discrete logarithm problem (ECDLP)" and the DHP over $E(\mathbb{F}_q)$ as "Elliptic curve Diffie–Hellman problem (ECDHP)".

Now we go back to the general case and let us describe the $O(\sqrt{\mathrm{ord}(x)})$ algorithm to solve $\mathrm{DLP}_G(x, y)$. Indeed, an obvious naive way to compute this in $O(\mathrm{ord}(x))$ is to repeatedly compute $x, x^2, x^3, \ldots$ until we see y at x^m and return m. However, if we know the order of the subgroup generated by x beforehand (denote it by n), then an $O(\sqrt{n})$ algorithm is possible as shown in the following Algorithm 3.

Note that if the list A has no match with B, then the assumption is false: y cannot be x^m for some $1 \leq m < n$. This is because once we assume $y = x^m$, we can write $m = i + jN$ by the division algorithm and the corresponding match must be found. This proves the correctness of the algorithm, and if the matching step between A and B is handled using efficient data structures such as hash table, then the matching step can be done in $O(N) = O(\sqrt{n})$ steps. Therefore this algorithm uses $O(\sqrt{n})$ memory and time. In fact, there is a better algorithm with nearly the same

[4] Note that this is not a standard notation.

[5] This is also not a standard notation.

Algorithm 3 Shanks' Babystep–Giantstep algorithm

Require: A group G, elements $x, y \in G$, and $n = \mathrm{ord}(x)$
Ensure: $m \in \mathbb{N}$ such that $x^m = y$
  $N \leftarrow \lceil \sqrt{n} \rceil$
  make a list $A \leftarrow [x, x^2, \ldots, x^N]$
  $z \leftarrow (x^N)^{-1}$
  make a list $B \leftarrow [yz, yz^2, \ldots, yz^N]$
  if an element of A matches an element of B, say, $x^i = yz^j$ **then**
    $m \leftarrow i + jN$ ▷ Since $x^i = yz^j = y(x^{-N})^j$, we have $y = x^{i+jN}$.
  end if

time complexity but much less memory due to Pollard called Pollard's ρ algorithm. The detailed analysis can be found in [15].

We continue with a few more applications of the apparent hardness of DLP and DHP.

1.4.3 Diffie–Hellman Key Exchange

Once we know the DHP, the Diffie–Hellman key exchange is very easy. Consider the scenario where Alice and Bob want to obtain a shared secret but the communication is done in an insecure channel. This may look impossible at first but it is actually given by the following method.

1. Alice and Bob agree on a group G and an element $g \in G$ such that $n = \mathrm{ord}(g)$ is finite. Everyone knows this information, including the eavesdroppers like Eve.
2. Alice picks a secret integer $1 < a < n$. Bob picks a secret integer $1 < b < n$.
3. Alice computes $A = g^a$. Bob computes $B = g^b$.
4. Alice publishes A. Bob publishes B. This can be done in an insecure channel.
5. Alice receives B. Bob receives A.
6. Alice computes B^a. Bob computes A^b. This is equal, and this becomes the shared secret.

During the exchange of A and B, Eve may know every public information: G, g, A, B (recall that $A = g^a$ and $B = g^b$), but what Eve won't be able to figure out easily is g^{ab}. This is precisely the DHP. By assuming the hardness of DHP, it is implied that the key exchange is safe.

The method is explained in a general settings for any group G and any base element $g \in G$ such that $\mathrm{ord}(g)$ is finite. However, since we've seen that DHP on elliptic curves in general is expected to be hard, we can take G to be $E(\mathbb{F}_q)$ for an agreed curve E over an agreed finite field $\mathbb{F}_q$, with a predefined point $P \in E(\mathbb{F}_q)$. This specialization of the Diffie–Hellman key exchange is known as the "Elliptic Curve Diffie–Hellman (ECDH)" protocol, and can be used for secret key agreement. Typical uses include the case of making a shared secret for another underlying symmetric cryptographic protocol.

1.4.4 Elliptic Curve Digital Signature Algorithm (ECDSA)

The goal of the digital signature algorithm is for a person (say, Alice) to digitally sign a document in a way such that any other person (say, Bob) can verify that the document is indeed signed by Alice. Here we follow the implementation given in [22].

1. Alice and Bob agree on a finite field $\mathbb{F}_p$, an elliptic curve $E/\mathbb{F}_p$, and a point $P \in E(\mathbb{F}_p)$ of prime order N. Everyone, including Eve, knows this information.
2. Alice selects a secret integer $1 < a < N$ and compute $A = [a]P$.
3. Alice publishes A. This is the public verification key. The secret a is the private signing key.
4. Suppose Alice wants to sign a document $d \bmod N$. Alice chooses a random integer $0 \leq k < N$ and computes $[k]A$. Let s_1 be the x-coordinate (modulo N) of $[k]P$. Compute $(d + as_1)k^{-1}$ in $\mathbb{F}_p$ and bring its integer representative between 0 and $p - 1$ to reduce under modulo N, and call this s_2. That is, $s_2 := (d + as_1)k^{-1} \bmod N$. Alice then publishes (s_1, s_2) as the digital signature of the document d.

5. Suppose Bob wants to verify (s_1, s_2). Bob computes $v_1 := ds_2^{-1} \bmod N$ and $v_2 := s_1 s_2^{-1} \bmod N$. Bob then computes $[v_1]P + [v_2]A \in E(\mathbb{F}_p)$ and check the x-coordinate of this point. If it is congruent to s_1 in modulo N, then the signature is valid. Otherwise, the signature is invalid.

Since

$$[v_1]P + [v_2]A = [ds_2^{-1}]P + [s_1 s_2^{-1}][a]P = [s_2^{-1}(d + as_1)]P = [k]P,$$

we see that if (s_1, s_2) is correct, then the x-coordinate of $[k]P$ (which is s_1) must be equal to the x-coordinate of $[v_1]P + [v_2]A$.

1.5 Schoof's algorithm to count $\#E(\mathbb{F}_q)$

In this chapter, we end with Schoof's algorithm to count the number of rational points over an elliptic curve over a finite field.

An interesting question is: given an elliptic curve $E/\mathbb{F}_q$, how many rational points are there, i.e. find $\#E(\mathbb{F}_q)$. A straightforward algorithm loops over all $x \in \mathbb{F}_q$ and checks for valid $y \in \mathbb{F}_q$. This runs in $O(q^2)$, but by substituting x in the Weierstrass equation for E, we're left with a quadratic equation over y, which has either $0, 1$, or 2 solutions, depending on the discriminant. Suppose we want to solve for $y \in \mathbb{F}_q$ for the equation

$$Ay^2 + By + C = 0,$$

then the quadratic formula requires computing $\sqrt{B^2 - 4AC}$ over $\mathbb{F}_q$. If $B^2 - 4AC = 0$, there is exactly one solution for y. If $B^2 - 4AC$ is a square in $\mathbb{F}_q$, then there are two solutions for y. Otherwise, $B^2 - 4AC$ is nonzero and not a square, so there is no solution for y. By computing x and assuming that we can check whether a value in $\mathbb{F}_q$ is a square or not in $O(\log q)$, we obtain an $O(q \log q)$ algorithm to count the number of rational points.

This, however, is still exponential in $\log q$. Here let us consider the following algorithm given by Schoof, which solves the problem in $O(\log^8 q)$. Recall 33: $\#E(\mathbb{F}_q) \in [q+1-2\sqrt{q}, q+1+2\sqrt{q}]$. The idea is to find $\#E(\mathbb{F}_q) \bmod k$ for different values of k and patch them together in the interval using Chinese remainder theorem. Before going to Schoof's algorithm, let us consider a few prerequisites here.

1.5.1 The m-Torsion Points

An important result that gives us the structure of $E[m]$ is the following theorem, which can be proved using 30, 25, and the characterization of finite abelian groups.

Theorem 35. *Let E/K be an elliptic curve and let $m \in \mathbb{Z}^*$. If $m \neq 0$ in K, i.e. $\mathrm{char}(K) = 0$ or $\mathrm{char}(K) \nmid m$, then*

$$E[m] \cong \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}.$$

Proof. See [22, III.6.4b]. □

1.5.2 The Tate Module

Consider an isogeny $\phi\colon E_1 \to E_2$. By restricting it to the set of ℓ-torsion points, it gives $\phi\colon E_1[\ell] \to E_2[\ell]$. The image is also an ℓ-torsion point because $[\ell]\phi(P) = \phi([\ell]P) = \phi(O) = O$ for all $P \in E_1[\ell]$.

If we take the inverse limit $\varprojlim_n E_1[\ell^n]$ where ℓ is a prime, with respect to maps $E_1[\ell^{n+1}] \xrightarrow{[\ell]} E_1[\ell^n]$, we retrieve a structure which is a subset of $\prod_{n=1}^{\infty} E_1[\ell^n]$. We call this the ℓ-adic Tate module and denote by $T_\ell(E_1)$. This is a $\mathbb{Z}_\ell$-module because a pointwise multiplication by elements from $\mathbb{Z}_\ell \subseteq \prod_{n=1}^{\infty} \mathbb{Z}/\ell^n\mathbb{Z}$ gives the necessary properties for it to be a module. Similarly, we construct $T_\ell(E_2)$.

Now (we temporarily forget about E_1 and E_2 and just consider a general E), since $E[\ell] \cong \mathbb{Z}/\ell\mathbb{Z} \times \mathbb{Z}/\ell\mathbb{Z}$ for all prime $\ell \neq \mathrm{char}(K)$, we see that the inverse limit is actually

$$T_\ell(E) := \varprojlim_n E[\ell^n] \cong \varprojlim_n \mathbb{Z}/\ell^n\mathbb{Z} \times \mathbb{Z}/\ell^n\mathbb{Z}.$$

Passing the limit through the multiplication, we would be left with

$$T_\ell(E) \cong \mathbb{Z}_\ell \times \mathbb{Z}_\ell,$$

which is a $\mathbb{Z}_\ell$-module, so it is a free module of rank 2 in particular. This allows us, in the future, to think of endomorphisms of $T_\ell(E)$ as a matrix of entries in $\mathbb{Z}_\ell$ upon choosing a $\mathbb{Z}_\ell$-basis for $T_\ell(E)$.

Since ϕ induces $\phi\colon E_1[\ell] \to E_2[\ell]$ for all ℓ, it works for $\ell, \ell^2, \dots$ for ℓ prime in particular. This induces a map

$$\phi_\ell\colon \prod_{n=1}^{\infty} E_1[\ell^n] \to \prod_{n=1}^{\infty} E_2[\ell^n].$$

Now we can restrict this map to $T_\ell(E_1)$ and we see that the image is also in $T_\ell(E_2)$, so we have

$$\phi_\ell\colon T_\ell(E_1) \to T_\ell(E_2).$$

Furthermore, this map is $\mathbb{Z}_\ell$-linear. This allows us to do linear algebra on isogenies. To capture, the whole process is a conversion from an isogeny to a $\mathbb{Z}_\ell$-linear map between Tate modules, in other words, our construction gives

$$\mathrm{Hom}(E_1, E_2) \to \mathrm{Hom}(T_\ell(E_1), T_\ell(E_2))$$

for any prime ℓ.

1.5.3 The Weil Pairing

Theorem 36. *For any elliptic curve E/K, there exists a function*

$$e_m\colon E[m] \times E[m] \to \boldsymbol{\mu}_m$$

where $\boldsymbol{\mu}_m$ is the group of m-th root of unity in $\bar{K}$, satisfying the following properties:

1. *Bilinear:*

$$e_m(S_1 + S_2, T) = e_m(S_1, T)e_m(S_2, T),$$
$$e_m(S, T_1 + T_2) = e_m(S, T_1)e_m(S, T_2)$$

   for all $S, T, S_1, S_2, T_1, T_2 \in E[m]$.

2. *Alternating:*

$$e_m(T, T) = 1$$

   for all $T \in E[m]$.

3. *Nondegenerate: For a fixed $T \in E[m]$, if*

$$e_m(S, T) = 1$$

   for all $S \in E[m]$ then $T = O$.

4. *Galois-invariant:*

$$\sigma(e_m(S, T)) = e_m(S^\sigma, T^\sigma)$$

   for all $S, T \in E[m]$ and $\sigma \in \mathrm{Gal}(\bar{K}/K)$.

5. *Compatible:*

$$e_{mm'}(S, T) = e_m([m']S, T)$$

   for all $S \in E[mm']$ and $T \in E[m]$.

Proof. We construct such a function explicitly, called the Weil pairing. See [22, III.8] for an explicit construction. See [22, III.8.1] for the proof that it satisfies these properties. □

Now, once we have such e_{ℓ^n} for all $n \in \mathbb{N}$ on a fixed prime ℓ, we have the maps

$$e_{\ell^n}\colon E[\ell^n] \times E[\ell^n] \to \boldsymbol{\mu}_{\ell^n}.$$

Taking inverse limit with respect to maps

$$E[\ell^{n+1}] \xrightarrow{[\ell]} E[\ell^n] \qquad \text{and} \qquad \boldsymbol{\mu}_{\ell^{n+1}} \xrightarrow{x \mapsto x^\ell} \boldsymbol{\mu}_{\ell^n},$$

we claim the induced map

$$e\colon T_\ell(E) \times T_\ell(E) \to T_\ell(\boldsymbol{\mu})$$

where $T_\ell(\boldsymbol{\mu})$ is the Tate module for roots of unity, i.e.

$$T_\ell(\boldsymbol{\mu}) = \varprojlim_n \boldsymbol{\mu}_{\ell^n}.$$

Before finishing the claim, we need to prove that the image of $T_\ell(E) \times T_\ell(E)$ by e actually lands on $T_\ell(\boldsymbol{\mu})$. Well suppose $P \in T_\ell(E)$ is represented by $(P_n)_{n\in\mathbb{N}}$ with $P_i \in E[\ell^i]$ for all $i \in \mathbb{N}$, and similarly for $Q \in T_\ell(E)$ with $(Q_n)_{n\in\mathbb{N}}$. Apply e_{ℓ^i} pointwise to the sequences $(P_n)_{n\in\mathbb{N}}$ and $(Q_n)_{n\in\mathbb{N}}$ to obtain $(e_{\ell^n}(P_n, Q_n))_{n\in\mathbb{N}}$. Let us check that this sequence satisfies the inverse system for $T_\ell(\boldsymbol{\mu})$, i.e.,

$$e_{\ell^{n+1}}(P_{n+1}, Q_{n+1})^\ell = e_{\ell^n}(P_n, Q_n).$$

Since $P_n = [\ell]P_{n+1}$ and $Q_n = [\ell]Q_{n+1}$ by the inverse system of $T_\ell(E)$, we see that it is enough to check that

$$e_{\ell^{n+1}}(S, T)^\ell = e_{\ell^n}([\ell]S, [\ell]T)$$

for any $S, T \in E[\ell^{n+1}]$. By the property 36, we see immediately that this is true, applying bilinearity and compatibility. This proves that

$$e\colon T_\ell(E) \times T_\ell(E) \to T_\ell(\boldsymbol{\mu})$$

is well-defined. Moreover, it inherits the properties from 36, making it bilinear, alternating, non-degenerate, Galois-invariant, and compatible.

Now, we consider an isogeny of E, i.e. $\phi \in \mathrm{Hom}(E, E)$. Then, we can induce it into $\phi_\ell \in \mathrm{Hom}(T_\ell(E), T_\ell(E))$. Fix a $\mathbb{Z}_\ell$-basis of $T_\ell(E)$, so that one can write the corresponding matrix of ϕ_ℓ as

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

with $a, b, c, d \in \mathbb{Z}_\ell$. This gives a well-defined determinant and trace for ϕ_ℓ with respect to this basis. We claim that they are invariant over different bases.

Theorem 37. *Let E/K be an elliptic curve. Let $\phi \in \mathrm{Hom}(E, E)$ be an isogeny, and let $\phi_\ell \in \mathrm{Hom}(T_\ell(E), T_\ell(E))$ be its induced enodomorphism of $T_\ell(E)$. Then, for any fixed $\mathbb{Z}_\ell$-basis of $T_\ell(E)$, we have*

$$\det(\phi_\ell) = \deg(\phi) \qquad \textit{and} \qquad \mathrm{tr}(\phi_\ell) = 1 + \deg(\phi) - \deg(1 - \phi).$$

In particular, this also shows that $\det(\phi_\ell)$ and $\mathrm{tr}(\phi_\ell)$ are in $\mathbb{Z}$ and don't depend on ℓ.

Proof. See [22, III.8.6]. □

1.5.4 The Characteristic Polynomial

What we've shown so far is a development towards the important result that forms the basis of Schoof's algorithm. By Hasse's bound, we've seen that for any elliptic curve $E/\mathbb{F}_q$, the number of rational points $\#E(\mathbb{F}_q)$ is in $[q + 1 - 2\sqrt{q}, q + 1 + 2\sqrt{q}]$. Let us denote by a the quantity $q + 1 - \#E(\mathbb{F}_q)$, and let us denote by ϕ the usual q-th power Frobenius map defined on $E \to E$. The main claim is the following.

Theorem 38. *The equation*

$$\phi^2 - a\phi + q = 0$$

is satisfied in $\mathrm{Hom}(E, E)$.

Proof. (Simplifying the one given by [22, V.2.3.1].) By the results earlier, we see that for any prime $\ell \neq p$,

$$\det(\phi_\ell) = \deg(\phi) = q$$

and

$$\mathrm{tr}(\phi_\ell) = 1 + \deg(\phi) - \deg(1 - \phi) = 1 + q - \#E(\mathbb{F}_q) = a.$$

By basic linear algebra, the characteristic polynomial for ϕ_ℓ is

$$\chi(T) = \det(TI_2 - \mathrm{Mat}(\phi_\ell)) = T^2 - \mathrm{tr}(\phi_\ell)T + \det(\phi_\ell) = T^2 - aT + q.$$

By Cayley–Hamilton theorem, ϕ_ℓ kills the polynomial, so

$$\phi_\ell^2 - a\phi_\ell + q = 0.$$

Now $0 = \det(\phi_\ell^2 - a\phi_\ell + q) = \deg(\phi^2 - a\phi + q)$ so we see that $\phi^2 - a\phi + q = 0$ in $\mathrm{Hom}(E, E)$. □

1.5.5 The Algorithm of Schoof

We go back to the question of counting $\#E(\mathbb{F}_q)$ by the idea of patching the solutions modulo small primes over the range $[q+1-2\sqrt{q}, q+1+2\sqrt{q}]$ using Chinese remainder theorem. To do that, we use the equation

$$\phi^2 - a\phi + q = 0$$

and try to compute $a \bmod \ell$ for some small set of ℓ's. In fact, if $\phi^2 - a\phi + q = 0$ is true, then it must be true upon evaluating at a point P, i.e.

$$\phi^2(P) - [a]\phi(P) + [q]P = O.$$

Writing $P = (x, y)$, we have

$$[a](x, y) = (x^{q^2}, y^{q^2}) + [q](x, y).$$

If furthermore, P is assumed to have order ℓ, then we can write $a = b\ell + c$ with $0 \leq c < \ell$ using the basic division algorithm so that

$$[a](x, y) = [b\ell + c](x, y) = [b]\underbrace{[\ell](x, y)}_{O} + [c](x, y) = [c](x, y).$$

In other words, by considering all ℓ-torsion points (x, y) of E, one can precompute

$$Q_{(x,y)} = (x^{q^2}, y^{q^2}) + [q](x, y)$$

at first, and then loop over $O, (x, y), [2](x, y), \ldots, [\ell-1](x, y)$ and check if their coordinates match. Let $S_{(x,y)}$ be the set of indices k such that $[k](x, y) = Q_{(x,y)}$, then it is obvious that $\bigcap_{P \in E[\ell]} S_P$ is a singleton $\{k_\star\}$. Such $k_\star$ is $a \bmod \ell$. Once we can do this for enough number of ℓ's, we can pinpoint a using the Chinese remainder theorem.

Instead of working on the points $P \in E[\ell]$, we work on all of them simultaneously using the technology of division polynomials.

Theorem 39. *For a Weierstrass equation*

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

there exists a sequence of polynomials $(\psi_n)_{n \in \mathbb{N}}$ with $\psi_n \in \mathbb{Z}[a_1, a_2, a_3, a_4, a_6, x, y]$ such that ψ_n vanishes on precisely the x-coordinates of the nonzero n-torsion points of E and nowhere else for all $n \in \mathbb{N}$.

Proof. Define

$$\begin{aligned} b_2 &= a_1^2 + 4a_2 \\ b_4 &= 2a_4 + a_1a_3 \\ b_6 &= a_3^2 + 4a_6 \\ b_8 &= a_1^2a_6 + 4a_2a_6 - a_1a_3a_4 + a_2a_3^2 - a_4^2, \end{aligned}$$

and,

$$\begin{aligned} \psi_1 &= 1, \\ \psi_2 &= 2y + a_1x + a_3, \\ \psi_3 &= 3x^4 + b_2x^3 + 3b_4x^2 + 3b_6x + b_8, \\ \psi_4 &= (\psi_2)(2x^6 + b_2x^5 + 5b_4x^4 + 10b_6x^3 + 10b_8x^2 + (b_2b_8 - b_4b_6)x + (b_4b_8 - b_6^2)), \end{aligned}$$

then inductively by the formulas

$$\begin{aligned} \psi_{2m+1} &= \psi_{m+2}\psi_m^3 - \psi_{m-1}\psi_{m+1}^3 \text{ for } m \geq 2, \\ \psi_2\psi_{2m} &= \psi_{m-1}^2\psi_m\psi_{m+2} - \psi_{m-2}\psi_m\psi_{m+1}^2 \text{ for } m \geq 3. \end{aligned}$$

One can (tediously) check that this sequence of polynomials satisfies the required property. This is taken from [22, Exercise 3.7]. We call this particular sequence of polynomials "the division polynomials". Notice that the computation is done rather simply, i.e., computing ψ_ℓ would take $O(\ell^c)$ for some constant c. □

Now, going back to the algorithm, instead of computing $(x^{q^2}, y^{q^2}) + [q](x, y)$ for each point (x, y), we treat x and y as symbols and compute it in the ring

$$R_\ell = \mathbb{F}_q[x, y]/(\psi_\ell(x), y^2 + a_1xy + a_3y - x^3 - a_2x^2 - a_4x - a_6).$$

That is, we compute using a polynomial representative in this ring, and when that polynomial has a nonlinear higher degree terms of y, we replace y^2 with $x^3 + a_2x^2 + a_4x + a_6 - a_1xy - a_3y$ so that the terms of y can be at most degree one. Also, when a polynomial has degree higher than ψ_ℓ, we divide that polynomial by ψ_ℓ and take its remainder. This allows us to bound the number of terms of a representative in R_ℓ to be at most ℓ^2. This completes the algorithm. We present it in the following pseudocode of Algorithm 4.

Algorithm 4 Schoof's algorithm

$A \leftarrow 1$
$\ell \leftarrow 3$
while $A < 4\sqrt{q}$ **do**
  $Q \leftarrow (x^{q^2}, y^{q^2}) + [q](x, y) \in E(R_\ell)$
  $P \leftarrow O$
  for $n = 0, \ldots, \ell - 1$ **do**
    if $Q = P$ **then**
      $n_\ell \leftarrow n$ and break out of the loop
    end if
    $P \leftarrow P + (x^q, y^q)$
  end for
  $A \leftarrow \ell A$
  $\ell \leftarrow$ the next smallest prime greater than ℓ
end while
use the Chinese remainder theorem to find a such that $a \equiv n_\ell \pmod{\ell}$ for all used ℓ
return $\#E(\mathbb{F}_q) = q + 1 - a$

Remark. *It would be nice to add another section describing computation of the Weil pairing using Miller's algorithm. However, this is already very long, and unfortunately it has to be omitted. However, one can still find the implementation of this algorithm in the repository given in the appendix.*

Chapter 2

The Mordell–Weil Theorem for $E/\mathbb{Q}$

In this chapter, we aim to study an important property of $E/\mathbb{Q}$, that is, the group $E(\mathbb{Q})$ is finitely generated. This is the celebrated Mordell–Weil theorem which we aim to give a prove. Recall that $\mathrm{char}(\mathbb{Q}) = 0$, so that we may work on Weierstrass equations of the form

$$E\colon y^2 = x^3 + Ax + B$$

with $A, B \in \mathbb{Q}$ instead of the more general equation.

2.1 The Lutz–Nagell Theorem

Before moving onto the Mordell–Weil theorem, it would be nice to consider a result concerning the structure of the torsion points of $E(\mathbb{Q})$. This section aims to give a simple algorithm to calculate the torsion subgroup

$$E(\mathbb{Q})_{\mathrm{tors}} = \bigcup_{n\geq 1} E(\mathbb{Q})[n]$$

of $E(\mathbb{Q})$.

For an elliptic curve $E/\mathbb{Q}$, we consider a special case where its Weierstrass equation has the form

$$E\colon y^2 = x^3 + Ax + B,$$

with $A, B \in \mathbb{Z}$. In such cases, we specifically say that E is defined over the integers, and write $E/\mathbb{Z}$. Moreover, if $r \geq 1$ is an integer, we write E_r for the subset

$$\{(x, y) \in E(\mathbb{Q})\colon \nu_p(x) \leq -2r, \nu_p(y) \leq -3r\} \cup \{O\}$$

of points on E. Note that this depends on p, but we will often write just E_r and will keep in mind that we agree on a fixed prime p beforehand.

We now take and simplify [24, 8.1] in the following finite sequence of claims.

Proposition 40 ([24, 8.1(2)])**.** *Let $E/\mathbb{Z}\colon y^2 = x^3 + Ax + B$. Let p be prime. If $(x, y) \in E(\mathbb{Q})$, then $\nu_p(x) < 0$ if and only if $\nu_p(y) < 0$. In such case, there exists an integer $r \geq 1$ such that $\nu_p(x) = -2r$ and $\nu_p(y) = -3r$.*

Proof. Suppose $(x, y) \in E(\mathbb{Q})$. We have the equality

$$2\nu_p(y) = \nu_p(y^2) = \nu_p(x^3 + Ax + B) = \min(\nu_p(x^3), \nu_p(Ax), \nu_p(B)) = 3\nu_p(x),$$

where the last equality holds if $\nu_p(x) < 0$, but if $\nu_p(x) \geq 0$ then $\nu_p(y) \geq 0$ as well. This proves that $\nu_p(x) < 0$ if and only if $\nu_p(y) < 0$, and in such case, $2\nu_p(y) = 3\nu_p(x)$. This completes the proof. □

Now, let us define

$$\tilde{\lambda}_r\colon \begin{cases} E_r & \to \mathbb{Z}_{(p)} \\ (x, y) & \mapsto p^{-r}x/y \\ O & \mapsto 0 \end{cases}$$

where $\mathbb{Z}_{(p)} \subseteq \mathbb{Q}$ is defined by

$$\left\{\frac{a}{b} \in \mathbb{Q}\colon a, b \in \mathbb{Z}; p \nmid b\right\}.$$

By 40, if $(x,y) \in E_r$ then $\nu_p(x) = -2r'$ and $\nu_p(y) = -3r'$ for some integer $r' \geq r$. This means $\nu_p(x/y) = r' \geq r$ and so $\nu_p(p^{-r}x/y) = r' - r \geq 0$, i.e. $p^{-r}x/y$ lands on $\mathbb{Z}_{(p)}$, so the map is well-defined.

Proposition 41 ([24, 8.1(4)])**.** *Let $E/\mathbb{Z}\colon y^2 = x^3 + Ax + B$. Let p be prime. If $(x,y) \in E_r$ but $(x,y) \notin E_{r+1}$ then $\tilde{\lambda}_r(x,y) \notin p\mathbb{Z}_{(p)}$.*

Proof. Consider the set $E_r \setminus E_{r+1}$, which is precisely

$$\{(x,y) \in E(\mathbb{Q})\colon \nu_p(x) = -2r, \nu_p(y) = -3r\},$$

so its image under $\tilde{\lambda}_r$ is $p^{-r}\left(\frac{k}{p^{2r}}\right) / \left(\frac{\ell}{p^{3r}}\right)$ for some $k, \ell \in \mathbb{Z}$ not divisible by p. Simplifying this expression, we have its image as $\frac{k}{\ell}$. Since $k, \ell \in \mathbb{Z}$ not divisible by p, we see that this cannot lie in $p\mathbb{Z}_{(p)}$. □

Now, the goal is to show that the composition $E_r \xrightarrow{\tilde{\lambda}_r} \mathbb{Z}_{(p)} \xhookrightarrow{\eta} \mathbb{Z}_{(p)}/p^{4r}\mathbb{Z}_{(p)}$ (where η is the natural modulo map) is a homomorphism of groups, and so we can apply the fundamental homomorphism theorem to see that

$$E_r/\ker(\eta \circ \tilde{\lambda}_r) \cong \operatorname{im}(\eta \circ \tilde{\lambda}_r)$$

is an isomorphism of groups.

Consider the following change of coordinates.

Lemma 42 ([24, 8.3])**.** *Let $E/\mathbb{Z}\colon y^2 = x^3 + Ax + B$. Let p be prime. We write a change of coordinates*

$$t = \frac{x}{y} \qquad \text{and } s = \frac{1}{y}.$$

Then $(x,y) \in E_r$ if and only if $\nu_p(s) \geq 3r$. If $\nu_p(s) \geq 3r$ then $\nu_p(t) \geq r$.

Proof. Well $\nu_p(s) = \nu_p\left(\frac{1}{y}\right) = -\nu_p(y)$. If $(x,y) \in E_r$ then $\nu_p(y) \leq -3r$ by definition, and so $\nu_p(s) = -\nu_p(y) \geq 3r$. Conversely, if $\nu_p(s) \geq 3r$ then $\nu_p(y) = -\nu_p(s) \leq -3r$. Apply 40 to see that $\nu_p(x) \leq -2r$ as well, so $(x,y) \in E_r$. This finishes the proof of the first statement.

Now, if $\nu_p(s) \geq 3r$, then $\nu_p(y) \leq -3r$ and $\nu_p(x) \leq -2r$. Apply 40 and suppose $\nu_p(y) = -3r'$ and $\nu_p(x) = -2r'$ with $r' \geq r \geq 1$. Then

$$\nu_p(t) = \nu_p\left(\frac{x}{y}\right) = \nu_p(x) - \nu_p(y) = -2r' - (-3r') = r' \geq r.$$

This finishes the proof. □

Now, suppose $P_1, P_2 \in E_r$. We consider the line passing through P_1 and P_2. If P_3 is another point on that line, then $P_1 + P_2 + P_3 = O$ by the group law on the elliptic curve. If we can manage to show that $P_3 \in E_r$ also, then it follows that E_r is a subgroup, because for any points $P, Q \in E_r$, we have $P + Q = -(-(P+Q) + O)$ and then we can apply the claim twice to see that $P + Q \in E_r$.

Consider the line passing through $P_1 \in E_r$ and $P_2 \in E_r$ hitting P_3, which is

$$ax + by + d = 0$$

for some $a, b, d \in \mathbb{Q}$. But without loss of generality one can assume further that $a, b, d \in \mathbb{Z}$. Dividing by y gives

$$at + b + ds = 0,$$

so we may work in this (s,t)-coordinate system instead.

Lemma 43 ([24, 8.4])**.** *Let $c \in p\mathbb{Z}_{(p)}$. Then working in $\mathbb{R}^2$, the line $t = c$ may hit the curve $s = t^3 + As^2t + Bs^3$ at a point (s,t) such that $s \in p\mathbb{Z}_{(p)}$ at most once. The line is not tangent at the point of intersection.*

Proof. Let us prove that it can intersect at most once. Suppose there are two points $s_1, s_2 \in p\mathbb{Z}_{(p)}$. Then (s_1, t) and (s_2, t) lie on the curve, so $s_1 - s_2 = A(s_1^2 - s_2^2)t + B(s_1^3 - s_2^3)$. Since we've assumped $s_1, s_2 \in p\mathbb{Z}_{(p)}$, write $s_1 = ps_1'$ and $s_2 = ps_2'$ for $s_1', s_2' \in \mathbb{Z}_{(p)}$. Suppose $\nu_p(s_1 - s_2) = k$, then

$\nu_p(ps_1' - ps_2') = k$ so $\nu_p(s_1' - s_2') = k-1$. This means $\nu_p({s_1'}^2 - {s_2'}^2) = \nu_p(s_1' - s_2') + \nu_p(s_1' + s_2') \geq k-1$ and similarly, $\nu_p({s_1'}^3 - {s_2'}^3) = \nu_p(s_1' - s_2') + \nu_p({s_1'}^2 + s_1's_2' + {s_2'}^2) \geq k-1$. We have

$$\begin{aligned} k = \nu_p(s_1 - s_2) &= \nu_p(A(s_1^2 - s_2^2)t + B(s_1^3 - s_2^3)) \\ &\geq \min(\nu_p(A(s_1^2 - s_2^2)t), \nu_p(B(s_1^3 - s_2^3))) \\ &= \min(\nu_p(A(p^2{s_1'}^2 - p^2{s_2'}^2)t), \nu_p(B(p^3{s_1'}^3 - p^3{s_2'}^3))) \\ &\geq \min(k-1+2, k-1+3) \\ &\geq k+1, \end{aligned}$$

which can be possible only if $k = +\infty$. This proves that $s_1 - s_2 = 0$.

Now, let us prove that the line is not tangent. By implicit differentiation on $s = t^3 + As^2t + Bs^3$, we see that

$$\frac{\mathrm{d}s}{\mathrm{d}t} = 3t^2 + As^2 + 2Ast\frac{\mathrm{d}s}{\mathrm{d}t} + 3Bs^2\frac{\mathrm{d}s}{\mathrm{d}t},$$

so

$$\frac{\mathrm{d}s}{\mathrm{d}t} = \frac{3t^2 + As^2}{1 - 2Ast - 3Bs^2}.$$

If $t = c$ is tangent to the curve, then the denominator must be zero (the slope must be infinity), that is, $1 - 2Ast - 3Bs^2 = 0$. This means $3Bs^2 + 2Ast = 1$, but the left hand side belongs to $p\mathbb{Z}_{(p)}$ and the right hand side does not. Hence a contradiction, so it is not tangent to the curve. □

Lemma 44. *Going back to the main situation, i.e., $E/\mathbb{Z}$: $y^2 = x^3 + Ax + B$, p is prime, $r \geq 1$ is an integer, assuming the change of coordinates*

$$t = \frac{x}{y} \qquad \text{and } s = \frac{1}{y}.$$

Assume $P_1, P_2 \in E_r$ and that the line passing through the points is

$$at + b + ds = 0,$$

then $d \neq 0$.

Proof. Suppose $d = 0$ then the line is in the form $at + b = 0$, i.e. $t = -b/a$. At P_1, we see that $at_1 + b = 0$ but $t_1 \in p^r\mathbb{Z}_{(p)}$, so $\nu_p(b) = \nu_p(a) + \nu_p(t_1)$, i.e. $\nu_p(-b/a) = \nu_p(t_1) \geq r$. In particular, $-b/a \in p\mathbb{Z}_{(p)}$, and so we can apply the previous lemma to see that it hits points with s-coordinate in $p\mathbb{Z}_{(p)}$ at most once. This implies $P_1 = P_2$, and so the line must be tangent at the point of intersection, which contradicts the lemma. This proves that $d \neq 0$. □

Dividing by $d \neq 0$, such line of intersection becomes

$$\frac{a}{d}t + \frac{b}{d} + s = 0,$$

or, more conveniently,

$$s = -\frac{a}{d}t - \frac{b}{d}.$$

Lemma 45 ([24, 8.5])**.** *Within this scenario (see previous lemma), we name the quantity $-\frac{a}{d}$ as α and name the quantity $-\frac{b}{d}$ as β, so that the line of intersection becomes $s = \alpha t + \beta$, then we have*

$$\alpha = \frac{t_2^2 + t_1t_2 + t_1^2 + As_2^2}{1 - A(s_1 + s_2)t_1 - B(s_2^2 + s_1s_2 + s_1^2)}.$$

Proof. If $t_1 \neq t_2$, we have $\alpha = \frac{s_2 - s_1}{t_2 - t_1}$. Since $s_i = t_i^3 + AS_i^2t_i + Bs_i^3$ for $i = 1, 2$, we have

$$\begin{aligned} &(s_2 - s_1)(1 - A(s_1 + s_2)t_1 - B(s_2^2 + s_1s_2 + s_1^2)) \\ &= (s_2 - s_1) - A(s_2^2 - s_1^2)t_1 - B(s_2^3 - s_1^3) \\ &= (s_2 - As_2^2t_2 - Bs_2^3) - (s_1 - As_1^2t_1 - Bs_1^3) + As_2^2(t_2 - t_1) \\ &= t_2^3 - t_1^3 + As_2^2(t_2 - t_1) \\ &= (t_2 - t_1)(t_2^2 + t_1t_2 + t_1^2 + As_2^2). \end{aligned}$$

This completes the proof in this case. Now if $t_1 = t_2$, then $s_1 \neq s_2$ implies $d = 0$ which contradicts the previous lemma, so $s_1 = s_2$. This means the line $s = \alpha t + \beta$ is the tangent at $P_1 = P_2$. (Not to be confused with the line $t = c$ in the previous lemma, which cannot be tangent to the curve!) We do another implicit differentiation on $s = t^3 + Ats^2 + Bs^3$ to see that

$$\frac{\mathrm{d}s}{\mathrm{d}t} = 3t^2 + As^2 + 2Ast\frac{\mathrm{d}s}{\mathrm{d}t} + 3Bs^2\frac{\mathrm{d}s}{\mathrm{d}t},$$

i.e.,

$$\frac{\mathrm{d}s}{\mathrm{d}t} = \frac{3t^2 + As^2}{1 - 2Ast - 3Bs^2}$$

as before, which is exactly in the form

$$\alpha = \frac{t_2^2 + t_1t_2 + t_1^2 + As_2^2}{1 - A(s_1 + s_2)t_1 - B(s_2^2 + s_1s_2 + s_1^2)}.$$

when $t_1 = t_2$ and $s_1 = s_2$. □

We see further that by this lemma, we have the following claim.

Lemma 46. *Following the same scenario (see previous lemma), we have*

$$\nu_p(\alpha) \geq 2r \qquad \text{and} \qquad \nu_p(\beta) \geq 3r.$$

Proof. By direct computation,

$$\begin{aligned}
\nu_p(\alpha) &= \nu_p\left(\frac{t_2^2 + t_1t_2 + t_1^2 + As_2^2}{1 - A(s_1 + s_2)t_1 - B(s_2^2 + s_1s_2 + s_1^2)}\right) \\
&= \nu_p(t_2^2 + t_1t_2 + t_1^2 + As_2^2) - \nu_p(1 - \underbrace{A(s_1 + s_2)t_1 - B(s_2^2 + s_1s_2 + s_1^2)}_{\in p\mathbb{Z}_{(p)}}) \\
&= \nu_p(t_2^2 + t_1t_2 + t_1^2 + As_2^2) \\
&\geq \min(2\nu_p(t_2), \nu_p(t_1) + \nu_p(t_2), 2\nu_p(t_1), \nu_p(A) + 2\nu_p(s_2)) \\
&\geq 2r.
\end{aligned}$$

Moreover,

$$\begin{aligned}
\nu_p(\beta) &= \nu_p(s_1 - \alpha t_1) \\
&\geq \min(\nu_p(s_1), \nu_p(\alpha) + \nu_p(t_1)) \\
&\geq 3r.
\end{aligned}$$

This completes the proof. □

This is enough to deduce the group structure.

Proposition 47 ([24, 8.1(1)])**.** *For an elliptic curve $E/\mathbb{Z}$ and any integer $r \geq 1$, the set E_r is a subgroup of $E(\mathbb{Q})$.*

Proof. Let us show that if $P_1, P_2 \in E_r$ has the line $at + b + ds = 0$ passing through them (where $a, b, d \in \mathbb{Z}$ and $d \neq 0$) hitting another point $P_3 \in E(\mathbb{Q})$, then $P_3 \in E_r$ also. Well, we write the line $at + b + ds = 0$ as $s = \alpha t + \beta$ with α found using 45. Then we consider the set of points in the intersection of the line and the elliptic curve, which satisfies the system

$$\begin{cases} s &= \alpha t + \beta \\ s &= t^3 + As^2t + Bs^3 \end{cases}.$$

By eliminating s by substituting $s \leftarrow \alpha t + \beta$ in the equation $s = t^3 + As^2t + Bs^3$, we see that it becomes

$$\alpha t + \beta = t^3 + A(\alpha t + \beta)^2 t + B(\alpha t + \beta)^3.$$

Rearranging the terms, we have

$$0 = t^3 + \frac{2A\alpha\beta + 3B\alpha^2\beta}{1 + B\alpha^3 + A\alpha^2}t^2 + \dots,$$

so we can apply Vieta's formula to see that

$$t_1 + t_2 + t_3 = -\frac{2A\alpha\beta + 3B\alpha^2\beta}{1 + B\alpha^3 + A\alpha^2}.$$

Apply ν_p to both sides, we have

$$\nu_p(t_1 + t_2 + t_3) \geq 5r,$$

with the knowledge that $\nu_p(1 + B\alpha^3 + A\alpha^2) = 0$ and that $\nu_p(\alpha) \geq 2r$ and $\nu_p(\beta) \geq 3r$ by the previous lemma. Since $t_1, t_2 \in p\mathbb{Z}_{(p)}$, we have $t_3 \in p\mathbb{Z}_{(p)}$ also (otherwise $\nu_p(t_1 + t_2 + t_3) < 1$, a contradiction). Now, $s_3 = \alpha t_3 + \beta$ so $\nu_p(s_3) \geq 3r$. Apply the characterization 42 to see that $P_3 \in E_r$. This proves the group structure. □

Now we revisit the map $\eta \circ \tilde{\lambda}_r$ which we wanted to prove that it is a homomorphism.

Proposition 48 (partial result of [24, 8.1(3)])**.** *For an elliptic curve $E/\mathbb{Z}$, a prime p, and an integer $r \geq 1$, the map*

$$\eta \circ \tilde{\lambda}_r \colon \begin{cases} E_r & \to \mathbb{Z}_{(p)}/p^{4r}\mathbb{Z}_{(p)} \\ (x, y) & \mapsto p^{4r}\mathbb{Z}_{(p)} + p^{-r}x/y \end{cases}$$

is a group homomorphism.

Proof. It is enough to use the same notion of $P_1, P_2, P_3 \in E_r$ on a line and show that $(\eta \circ \tilde{\lambda}_r)(P_1) + (\eta \circ \tilde{\lambda}_r)(P_2) + (\eta \circ \tilde{\lambda}_r)(P_3) = p^{4r}\mathbb{Z}_{(p)} + 0$, because if this is shown, then for any $P, Q \in E_r$,

$$(\eta \circ \tilde{\lambda}_r)(P) + (\eta \circ \tilde{\lambda}_r)(Q) + (\eta \circ \tilde{\lambda}_r)(-(P + Q)) = p^{4r}\mathbb{Z}_{(p)} + 0$$

but for any $(x, y) \in E_r$, $(\eta \circ \tilde{\lambda}_r)(-(x, y)) = p^{-r}x/(-y) = -p^{-r}x/y = -(\eta \circ \tilde{\lambda}_r)(x, y)$, so this would imply

$$(\eta \circ \tilde{\lambda}_r)(P) + (\eta \circ \tilde{\lambda}_r)(Q) - (\eta \circ \tilde{\lambda}_r)(P + Q) = p^{4r}\mathbb{Z}_{(p)} + 0$$

, i.e.

$$(\eta \circ \tilde{\lambda}_r)(P + Q) = (\eta \circ \tilde{\lambda}_r)(P) + (\eta \circ \tilde{\lambda}_r)(Q).$$

Now, let us prove the claim. Write $P_i = (x_i, y_i)$ as (s_i, t_i) with the usual change of coordinates

$$t_i = \frac{x_i}{y_i} \qquad \text{and } s_i = \frac{1}{y_i}$$

for $i = 1, 2, 3$. We have

$$(\eta \circ \tilde{\lambda}_r)(P_i) = p^{4r}\mathbb{Z}_{(p)} + p^{-r}(x_i/y_i) = p^{4r}\mathbb{Z}_{(p)} + p^{-r}t_i.$$

So that the sum becomes

$$\sum_{i=1}^{3}(\eta \circ \tilde{\lambda}_r)(P_i) = p^{4r}\mathbb{Z}_{(p)} + p^{-r}(t_1 + t_2 + t_3),$$

but we've shown that $\nu_p(t_1 + t_2 + t_3) \geq 5r$, so $\nu_p(p^{-r}(t_1 + t_2 + t_3) \geq 4r$, that is, $p^{-r}(t_1 + t_2 + t_3)$ belongs to $p^{4r}\mathbb{Z}_{(p)}$. This proves that

$$\sum_{i=1}^{3}(\eta \circ \tilde{\lambda}_r)(P_i) = p^{4r}\mathbb{Z}_{(p)}$$

which completes the proof. □

Now we can study the kernel of the map.

Proposition 49 (full result of [24, 8.1(3)])**.** *For an elliptic curve $E/\mathbb{Z}$, a prime p, and an integer $r \geq 1$, the map*

$$\eta \circ \tilde{\lambda}_r \colon \begin{cases} E_r & \to \mathbb{Z}_{(p)}/p^{4r}\mathbb{Z}_{(p)} \\ (x, y) & \mapsto p^{4r}\mathbb{Z}_{(p)} + p^{-r}x/y \end{cases}$$

has its kernel as E_{5r}, therefore, the map

$$\lambda_r \colon \begin{cases} E_r/E_{5r} & \to \mathbb{Z}_{(p)}/p^{4r}\mathbb{Z}_{(p)} \\ E_{5r} + (x, y) & \mapsto p^{4r}\mathbb{Z}_{(p)} + p^{-r}x/y \end{cases}$$

is well-defined. Furthermore, it is an injective homomorphism of additive abelian groups.

Proof. Let us study the kernel of $\eta \circ \tilde{\lambda}_r$. We have

$$\begin{aligned}\ker(\eta \circ \tilde{\lambda}_r) &= (\eta \circ \tilde{\lambda}_r)^{-1}(p^{4r}\mathbb{Z}_{(p)}) \\ &= \tilde{\lambda}_r^{-1}(p^{4r}\mathbb{Z}_{(p)}) \\ &= \{(x,y) \in E_r \colon p^{-r}x/y \in p^{4r}\mathbb{Z}_{(p)}\} \\ &= \{(x,y) \in E_r \colon \nu_p(x/y) \geq 5r\} \\ &= E_{5r},\end{aligned}$$

where 40 is used to deduce the last equality. This determines the kernel. Now we apply the fundamental homomorphism theorem to see that the composition

$$E_r/E_{5r} \cong \operatorname{im}(\eta \circ \tilde{\lambda}_r) \hookrightarrow \mathbb{Z}_{(p)}/p^{4r}\mathbb{Z}_{(p)}$$

is precisely λ_r. Since it is an inclusion of an isomorphism, it is an injective homomorphism. □

Corollary 50 ([24, 8.6])**.** *Let $E/\mathbb{Z}\colon y^2 = x^3 + Ax + B$, p be a prime, $r \geq 1$ be an integer. If $n > 1$ is an integer that is not a power of p, then E_1 contains no point of exact order n.*

Proof. Suppose $P \in E_1$ has order n. Since n is not a power of p, we can write $n = p^{\nu_p(n)}k$ for some k coprime to p. The point $Q = [p^{\nu_p(n)}]P$ has order k. Let r be the largest integer such that $Q \in E_r$. In particular, $Q \notin E_{5r}$ so

$$k\lambda_r(E_{5r} + Q) = \lambda_r(E_{5r} + kQ) = \lambda_r(E_{5r} + O) = p^{4r}\mathbb{Z}_{(p)}.$$

Since k is coprime to p, in particular, $\nu_p(k) = 0$, so

$$\lambda_r(E_{5r} + Q) = p^{4r}\mathbb{Z}_p,$$

that is, $E_{5r} + Q \in \ker(\lambda_r)$, so $Q \in E_{5r}$, which is a contradiction. Therefore, P does not exist. □

We are now ready to state and prove the Lutz–Nagell theorem.

Theorem 51 (Lutz–Nagell)**.** *Let $E/\mathbb{Z}\colon y^2 = x^3 + Ax + B$. Let $P = (x,y) \in E(\mathbb{Q})$. If P has finite order, then $x, y \in \mathbb{Z}$. If furthermore $y \neq 0$, then $y^2 \mid 4A^3 + 27B^2$.*

Proof. (Edited from [24, 8.7]) Suppose x or y is not in $\mathbb{Z}$, then there is a prime p dividing the denominator of one of them. Fixing this prime p, by 40, $P \in E_r$ for some integer $r \geq 1$. Let n be the order of P and let ℓ be a prime dividing n. Let $Q = [n/\ell]P \in E_r$, then Q has order ℓ. By 50, since $Q \in E_r \subseteq E_1$ has exact order ℓ, then ℓ must be a prime power of p, but it is prime, so $\ell = p$. Let j be the maximum integer such that $Q \in E_j$, so that $Q \notin E_{j+1}$. Apply 41 to see that $\tilde{\lambda}_j(Q) \notin p\mathbb{Z}_{(p)}$. But

$$p\lambda_j(Q) = \lambda_j([p]Q) = \lambda_j(O) = p^{4j}\mathbb{Z}_{(p)},$$

so $\tilde{\lambda}_j(Q) \in p^{4j-1}\mathbb{Z}_{(p)}$. This contradicts the claim that $\tilde{\lambda}_j(Q) \notin p\mathbb{Z}_{(p)}$. It follows that $x, y \in \mathbb{Z}$.

Now further assume $y \neq 0$, then $[2]P \neq O$. Write $[2]P = (x_2, y_2)$. Since $[2]P$ has finite order, by the first part of the theorem we see that $x_2, y_2 \in \mathbb{Z}$. Now we recall the explicit duplication law for E [22, III.2.3]. Since we're using $E\colon y^2 = x^3 + Ax + B$, we have $a_1 = 0, a_2 = 0, a_3 = 0, a_4 = A, a_6 = B$ in $E\colon y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$. By direct computation, we have

$$\begin{aligned} b_2 &= a_1^2 + 4a_2 &&= 0 \\ b_4 &= 2a_4 + a_1a_2 &&= 2A \\ b_6 &= a_3^2 + 4a_6 &&= 4B \\ b_8 &= a_1^2a_6 + 4a_2a_6 - a_1a_3a_4 + a_2a_3^2 - a_4^2 &&= -A^2, \end{aligned}$$

so that the duplication formula

$$x([2](x,y)) = \frac{x^4 - b_4x^2 - 2b_6x - b_8}{4x^3 + b_2x^2 + 2b_4x + b_6}$$

becomes

$$x_2 = \frac{x^4 - 2Ax^2 - 8Bx + A^2}{4x^3 + 4Ax + 4B} = \frac{x^4 - 2Ax^2 - 8Bx + A^2}{4y^2}$$

for the point $P = (x, y)$. This implies that

$$y^2 \mid x^4 - 2Ax^2 - 8Bx + A^2.$$

Since

$$(3x^2 + 4A)(x^4 - 2Ax^2 - 8Bx + A^2) - (3x^3 - 5Ax - 27B)(x^3 + Ax + B) = 4A^3 + 27B^2,$$

and y^2 divides both $x^4 - 2Ax^2 - 8Bx + A^2$ and $x^3 + Ax + B$, we conclude that y^2 divides $4A^3 + 27B^2$. □

Corollary 52. *Let $E/\mathbb{Q}$ be an elliptic curve. Then $E(\mathbb{Q})_{\text{tors}}$ is finite.*

Proof. We claim that there exists a change of variable, i.e., there exists $E'/\mathbb{Z}$ isomorphic to $E/\mathbb{Q}$, so that the rational torsion points of E' corresponds bijectively to the rational torsion points of E. Let us give that such change of variable. Let $E/\mathbb{Q}$ be given by

$$E\colon y^2 = x^3 + Ax + B, \qquad A, B \in \mathbb{Q}.$$

Then write $A = \frac{a}{c}$ and $B = \frac{b}{d}$ with $a, b, c, d \in \mathbb{Z}; c, d \neq 0$ to get

$$E\colon y^2 = x^3 + \frac{a}{c}x + \frac{b}{d}.$$

Now multiply everything by c^6d^6 to see that

$$E\colon c^6d^6y^2 = c^6d^6x^3 + ac^5d^6x + bc^6d^5.$$

Construct the following change of variable:

$$s \leftarrow c^3d^3y \qquad \text{and } t \leftarrow c^2d^2x.$$

We have

$$E\colon s^2 = t^3 + ac^3d^4t + bc^6d^5.$$

So the curve

$$E'\colon y^2 = x^3 + ac^3d^4x + bc^6d^5$$

is an elliptic curve over $\mathbb{Z}$ which is isomorphic to E. Apply Lutz–Nagell theorem to E' to see that $E'(\mathbb{Q})_{\text{tors}}$ is finite, and so $E(\mathbb{Q})_{\text{tors}}$ is finite as well. □

This does not only show that $E(\mathbb{Q})_{\text{tors}}$ is finite, but also that since we can do a change of variables on E into another elliptic curve over $\mathbb{Z}$, we can loop over all y such that y^2 divides $4A^3 + 27B^2$ and check all possible x to enumerate all the torsion points, and maps them back into $E/\mathbb{Q}_{\text{tors}}$. All can be done in finite (albeit exponential) time with respect to the size of the input, i.e., $\log_2(A) + \log_2(B)$. See the following Algorithm 5 for the pseudocode.

In fact, Lutz–Nagell theorem tells us a list of potential torsion points of E. However, the procedure of verifying whether a potential point is actually really a torsion point or not deserves an explanation. The final loop $P \in T$ checks whether a potential point P is actually a torsion point or not. If it were to be a torsion point, then any multiple $[n]P$ of P would also be a torsion point, but we know that all torsion points lie in T, so $[n]P \in T$ in particular. This allows us to bound the exact order of P by $\#T + 1$. Furthermore, if $[n]P = [m]P$ for some $1 \geq n < m \leq \#T$, then $[m - n]P = O$, meaning that there exists an integer k (in this case it is $m - n$) such that $[k]P = O$ with $1 \leq k \leq \#T$. In other words, the list

$$P, [2]P, [3]P, \ldots, [\#T + 1]P$$

has more than $\#T$ terms, so either some term here lies outside T (which tells us that P is not a torsion point), or that two terms are equal (which tells us that there exists $k \in \mathbb{N}^*$ such that $[k]P = O$).

Furthermore, there are other better algorithms to do this but unfortunately it is out of the scope of the thesis. See Doud's method [24, Section 9.6] for reference.

Algorithm 5 A direct algorithm to enumerate the torsion points

Require: $E/\mathbb{Q}\colon y^2 = x^3 + Ax + B$
Ensure: $S = E(\mathbb{Q})_{\text{tors}}$
  if $A, B \notin \mathbb{Z}$ **then**
    flag $\leftarrow$ `true`
    convert E into $E/\mathbb{Z}$ using the change of variable in 52
  end if
  $y \leftarrow 0$, $T \leftarrow \emptyset$
  solve for x in $0 = x^3 + Ax + B$ and add the solutions $(x, 0)$ to the set T
  for $y = 1, 2, 3, \ldots, 4A^3 + 27B^2$ **do**
    if $y^2 \mid 4A^3 + 27B^2$ **then**
      solve for x in $y^2 = x^3 + Ax + B$ and add the solutions $(x, y), (-x, y)$ to the set T
    end if
  end for
  $S \leftarrow \emptyset$
  for $P \in T$ **do**
    $Q \leftarrow O$
    for $n = 1, \ldots, \#T + 1$ **do**
      $Q \leftarrow Q + P$
      if $Q \notin T$ **then**
        erase P from T and continue the outer loop onto the next $P \in T$
      end if
      if $Q = O$ **then**
        $Q = [n]P = O$, so add P to S and break this loop
      end if
    end for
  end for
  if flag is `true` **then** convert S back using the change of variable in 52
  end if

2.2 The Weak Mordell–Weil Theorem

Remark. *To avoid confusion, in this thesis, if $(G, \times)$ is an abelian group, then instead of denoting by G^2 the group of squares, we denote by G^{sq} the group of squares. It is defined as*

$$G^{\text{sq}} = \{g^2 \colon g \in G\}.$$

Observe that it is a normal subgroup of G. This notation is made in order to avoid confusion between the direct product $G \times G$ (which is also denoted by G^2) and this particular group of squares G^{sq}. In the following sections, we write G^n to denote the direct product $\underbrace{G \times G \times \cdots \times G}_{n}$.

Now we move to the main part. In order to prove the Mordell–Weil theorem, the main idea consists of two steps:

1. Weak Mordell–Weil theorem: the group $E(\mathbb{Q})/2E(\mathbb{Q})$ is finite;
2. Descent: assuming that $E(\mathbb{Q})/2E(\mathbb{Q})$ is finite, introduce the notion of height and show that there exists a finite set of points under a constant height, such that all other points can descend into this set in a finite number of steps.

In this section we will try to prove the Weak Mordell–Weil theorem first. Normally one would introduce the idea in the case where $E[2] \in E(\mathbb{Q})$ with a particular map $\phi\colon E(\mathbb{Q}) \to (\mathbb{Q}^\times/\mathbb{Q}^{\times\text{sq}})^3$, and then prove that the kernel of this map is $2E(\mathbb{Q})$ and that the image of the map is finite. However, to keep it shorter, we will directly generalize this idea into the case without the assumption $E[2] \subseteq E(\mathbb{Q})$, and work with the generalized map (using the same idea) on a field K which would be a finite extension of $\mathbb{Q}$. To continue, we first introduce the basic notions of algebraic number theory.

2.2.1 Ring of Integers in an Algebraic Number Field

This subsection aims to explain the following figure 2.1. We will eventually work with K and $\mathcal{O}_K$. This section utilizes [16, §2] as the main reference.

Remark. *The standard notation uses $\bar{A}$ to denote the integral closure of A. However, the bar, which normally means "closure", can mean algebraic closure, integral closure, or topological closure, depending on usage. To avoid confusion, we explicitly write* $\mathrm{IntCl}_B(A)$ *to denote the integral closure of A in B, and drop B when it is clear which B we're working on, writing just* $\mathrm{IntCl}(A)$*. Note that this is not a standard notation.*

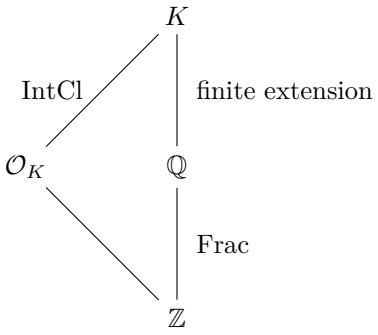

Figure 2.1: Ring of integers in an algebraic number field

Definition (Algebraic number field)**.** *A field K which makes a finite extension $K/\mathbb{Q}$ is said to be an algebraic number field.*

Remark. *Since $\mathbb{Q}$ is a perfect field, such a finite extension K is always separable.*

Definition. *Let $A \subseteq B$ be commutative rings, then an element $b \in B$ is said to be integral over A if there exists a nonconstant monic polynomial $P \in A[T]$ such that $P(b) = 0$. The ring B is said to be integral over A if every element is integral over A.*

Remark. *Such polynomial has to be monic, i.e. the leading coefficient is* 1 *(multiplicative identity of A), and this is important, because otherwise it is just the definition of algebraic elements over a ring.*

Lemma 53 ([16, (2.3)])**.** *Working with matrices over a ring (instead of a field), we recall the following result from linear algebra which still holds: Let A be an $n \times n$ matrix over a ring R. Let A^* be the adjoint matrix, i.e. the transpose of the cofactor matrix C, where $C_{i,j}$ is $(-1)^{i+j}M_{i,j}$, where $M_{i,j}$ is the determinant of the $(n-1) \times (n-1)$ matrix obtained by deleting the i-th row and the j-th column from A. Then,*

$$AA^* = A^*A = \det(A)I_n.$$

Proof. The proof follows the standard proof of this statement for matrices over a field (i.e. using Laplace expansion to compute the determinant, then evaluate the cofactor to see that it equals the expressions from Laplace expansion), but we see that we are not using any properties of the base field which are not present for a ring. This allows the result to hold for matrices over a ring. □

Proposition 54. *We work over a square matrix A of dimension $n \times n$ over a ring R. Let x be a vector over that ring, i.e. $x \in R^n$, then $Ax = 0$ implies $(\det A)x = 0$, where the latter is just the multiplication of the vector x by the scalar* $\det A$*.*

Proof. Using the previous lemma, we see that from

$$A^*A = \det(A)I_n,$$

we have

$$0 = A^*0 = A^*Ax = \det(A)I_nx = \det(A)x.$$

This completes the proof. □

Proposition 55 ([16, (2.2)])**.** *Finitely many elements $b_1, \ldots, b_n \in B$ are all integral over A if and only if the ring $A[b_1, \ldots, b_n]$ (viewed as an A-module) is finitely generated.*

Proof. (This follows the one given in [16, (2.2)].) Let us first prove that if $b \in B$ is integral over A then $A[b]$ is a finitely generated A-module. Since b is integral over A, there exists a monic polynomial $f \in A[T]$ such that $f(b) = 0$. Write it as

$$f(T) = \sum_{i=0}^{n} a_i T^i$$

with $a_n = 1$ and $a_0, a_1, \ldots, a_n \in A$ where $n \geq 1$. Now let $c \in A[b]$ be any element. By definition, there exists $g \in A[T]$ such that $g(b) = c$. By the division algorithm (since f is monic, we don't require A to be a field), one can write $g = fq + r$ with $r = 0$ or $\deg(r) < \deg(f) = n$, so that $g(b) = f(b)q(b) + r(b)$. But $f(b) = 0$, so if

$$r(T) = \sum_{i=0}^{\deg(r)} d_i T^i,$$

then

$$c = g(b) = r(b) = \sum_{i=0}^{\deg(r)} d_i b^i.$$

This proves that the set $\{1, b, b^2, \ldots, b^{n-1}\}$ generates the A-module $A[b]$. Now we may do by induction to prove the general case. Suppose $b_1, \ldots, b_n \in B$ are integral over A. Then $A[b_1, \ldots, b_{n-1}]$ is finitely generated. Write $R = A[b_1, \ldots, b_{n-1}]$, then b_n is integral over R in particular. This means $R[b_n]$ is a finitely generated R-module. Suppose $\omega_1, \ldots, \omega_r$ is a generating set of $R[b_n]$, where $\omega_1, \ldots, \omega_r \in R[b_n]$. This means every element $x \in R[b_n]$ can be written as $\sum_{i=1}^{r} \lambda_i \omega_i$ for some $\lambda_i \in R$. Since R is a finitely generated A-module, suppose $\zeta_1, \ldots, \zeta_s \in R$ is an A-generating set of R. Then one can write $\lambda_i = \sum_{j=1}^{s} \mu_j \zeta_j$, for some $\mu_j \in A$ for each $1 \leq j \leq s$. This means such an element $x \in R[b_n]$ can be written as

$$x = \sum_{i=1}^{r} \lambda_i \omega_i = \sum_{i=1}^{r} \left(\sum_{j=1}^{s} (\mu_j \zeta_j) \right) \omega_i = \sum_{i,j} \mu_j \zeta_j \omega_i.$$

So the set $\{\zeta_j \omega_i\}_{\substack{1 \leq j \leq s \\ 1 \leq i \leq r}}$ is an A-generating set for $R[b_n]$, which is finite. This proves that $R[b_n]$ is a finitely generated A-module. This completes the forward implication.

Now we look at the converse. Suppose $A[b_1, \ldots, b_n]$ is finitely generated, i.e. the elements $\omega_1, \ldots \omega_r \in A[b_1, \ldots, b_n]$ generates $A[b_1, \ldots, b_n]$. Then for any $b \in A[b_1, \ldots, b_n]$, we also have $b\omega_i \in A[b_1, \ldots, b_n]$, so that for each i there exists $a_{i,1}, \ldots, a_{i,r} \in A$ such that

$$b\omega_i = \sum_{j=1}^{r} a_{i,j} \omega_j.$$

Define an $r \times r$ matrix A with $A_{i,j} = a_{i,j}$. Then the above equation is equivalent to

$$(bI_r - A)x = 0$$

where

$$x = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_r \end{pmatrix}.$$

Apply the previous proposition to see that $\det(bI_r - A)x = 0$ also. Now since $1 \in A[b_1, \ldots, b_n]$, we have

$$1 = \sum_{i=1}^{r} c_i \omega_i$$

for some $c_1, \ldots, c_r \in A$. Let

$$v = \begin{pmatrix} c_1 & c_2 & \cdots & c_r \end{pmatrix}$$

and multiply the equation $\det(bI_r - A)x = 0$ by v from the left to see that

$$v \det(bI_r - A)x = \det(bI_r - A)vx = 0,$$

but $vx = 1$, so $\det(bI_r - A) = 0$. The polynomial $\chi(T) = \det(TI_r - A)$ is a monic polynomial in $A[T]$ where $\chi(b) = 0$. This proves that b is integral over A. This is true for any $b \in A[b_1, \ldots, b_n]$, hence, this completes the proof. □

Theorem 56. *The sum of integral elements in B over A are integral over A. The product of integral elements in B over A are integral over A. In particular, the set of all integral elements in B over A forms a ring.*

Proof. By the previous proposition, if $b_1, b_2, \ldots, b_n \in B$ are integral over A, then any element $b \in A[b_1, \ldots, b_n]$ are also integral over A since $A[b_1, \ldots, b_n, b] = A[b_1, \ldots, b_n]$ which is a finitely generated A-module. In particular, if b_1, b_2 are integral over A, then since $b_1 + b_2$ and $b_1 b_2$ are in $A[b_1, b_2]$, they are in fact integral over A. □

Proposition 57 ([16, (2.4)])**.** *If $A \subseteq B \subseteq C$ are rings such that C is integral over B and B is integral over A, then C is integral over A.*

Proof. Using the same argument as in the converse direction of 55, let $c \in C$ be an element. Pick a monic polynomial $f \in B[T]$ killing c. Write

$$f(T) = \sum_{i=0}^{n} b_i T^i.$$

Since B is integral over A, the elements $b_0, \ldots, b_n$ are integral over A, so $R = A[b_0, \ldots, b_n]$ is a finitely generated A-module. Write $\zeta_1, \ldots, \zeta_s \in R$ as an A-generating set for R. Since c is integral over R, $R[c]$ is a finitely generated R-module. Write $\omega_1, \ldots, \omega_r \in R[c]$ as an R-generating set for $R[c]$. Then for any $x \in R[c]$ there exists $\lambda_1, \ldots, \lambda_r \in R$ such that

$$x = \sum_{i=1}^{r} \lambda_i \omega_i.$$

Further, for each $1 \leq i \leq r$, since $\lambda_i \in R$, there exists $\mu_{i,1}, \ldots, \mu_{i,s} \in A$ such that

$$\lambda_i = \sum_{j=1}^{s} \mu_{i,j} \zeta_j.$$

This gives

$$x = \sum_{i=1}^{r} \sum_{j=1}^{s} \mu_{i,j} \zeta_j \omega_i,$$

the existence of $\mu_{i,j}$ holds for every $x \in R[c]$, and so $\{\omega_i \zeta_j\}_{\substack{1 \leq i \leq r \\ 1 \leq j \leq s}}$ is an A-generating set for $R[c]$. Therefore $R[c] = A[b_0, \ldots, b_n, c]$ is a finitely generated A-module, which means c is integral over A. This is true for all $c \in C$, so C is integral over A. □

This allows us to properly define

Definition. *For rings $A \subseteq B$, we denote by $\mathrm{IntCl}_B(A)$ the set of elements of B which are integral over A.*

Now, a very common situation arises when one picks a base ring A, and let K be its field of fractions, and let L/K be a finite field extension, then let $B := \mathrm{IntCl}_L(A)$. This gives the following figure 2.2.

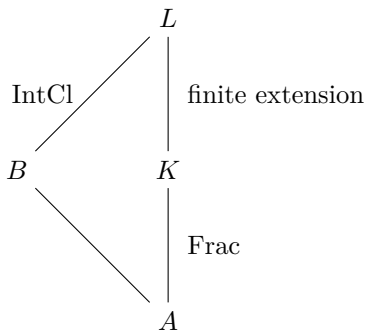

Figure 2.2: Ring of integers in an algebraic number field

Once we plug in $A \leftarrow \mathbb{Z}$, and rename L to K, which will be a finite extension of $\mathbb{Q}$ (called algebraic number field), we have the diagram shown in Figure 2.1. Such B will be denoted by $\mathcal{O}_K$ and called the ring of integers in K. We consider the following definition.

Definition. *Let K be an algebraic number field. Let S be a finite set of prime ideals of $\mathcal{O}_K$, then we define*

$$\mathcal{O}_{K,S} := \{x \in K : \text{there exists } a \in \bigcap_{\mathfrak{p} \notin S} \mathcal{O}_K \setminus \mathfrak{p} \text{ such that } ax \in \mathcal{O}_K\} = \left\{ \frac{a}{b} : a \in \mathcal{O}_K, b \in \bigcap_{\mathfrak{p} \notin S} \mathcal{O}_K \setminus \mathfrak{p} \right\}.$$

We state a few results that are required to prove the weak Mordell–Weil theorem here, but the proofs require basic knowledge in the theory of algebraic number theory.

Theorem 58 ([10, 8.8])**.** *Let K be an algebraic number field. Let S be a finite set of prime ideals of $\mathcal{O}_K$. Then there exists a finite set $S' \supseteq S$ such that the ring $\mathcal{O}_{K,S'}$ is a PID.*

Theorem 59. *Let K be an algebraic number field and let S be a finite set of prime ideals of $\mathcal{O}_K$. Then the group $\mathcal{O}_{K,S}^\times$ is finitely generated.*

Further properties can be proved, but are outside the scope of the thesis. We move to the main map in the next section.

2.2.2 The Map $\phi\colon E(K) \to (K^\times/K^{\times\text{sq}})^3$

Suppose $E/\mathbb{Q}$ is an elliptic curve in the form

$$E\colon y^2 = (x - e_1)(x - e_2)(x - e_3)$$

with $e_1, e_2, e_3 \in K$ where K is an algebraic number field, one defines the map $\phi\colon E(K) \to (K^\times/K^{\times\text{sq}})^3$ as

$$\begin{aligned}
(x, y) &\mapsto (x - e_1, x - e_2, x - e_3) \text{ if } y \neq 0 \\
O &\mapsto (1, 1, 1) \\
(e_1, 0) &\mapsto ((e_1 - e_2)(e_1 - e_3), e_1 - e_2, e_1 - e_3) \\
(e_2, 0) &\mapsto (e_2 - e_1, (e_2 - e_1)(e_2 - e_3), e_2 - e_3) \\
(e_3, 0) &\mapsto (e_3 - e_1, e_3 - e_2, (e_3 - e_1)(e_3 - e_2)).
\end{aligned}$$

First, we see that for any curve $E/\mathbb{Q}$, our usual Weierstrass equation is $y^2 = x^3 + Ax + B$ for $A, B \in \mathbb{Q}$. However, in order for it to be a valid elliptic curve, it must be smooth, i.e. the equation $0 = x^3 + Ax + B$ cannot have multiple roots in $\bar{\mathbb{Q}}$, because if there were to be multiple roots at $(x_\star, 0)$, then the polynomial $f(x, y) = y^2 - x^3 - Ax - B$ has $f(x_\star, 0) = 0$ and $\frac{\partial f}{\partial x}(x_\star, 0) = \frac{\partial f}{\partial y}(x_\star, 0) = 0$, so the point $(x_\star, 0)$ becomes singular and so the curve is not an elliptic curve. This means, in $\bar{\mathbb{Q}}$, one can factorize $x^3 + Ax + B$ into $(x - e_1)(x - e_2)(x - e_3)$ with $e_1, e_2, e_3 \in \bar{\mathbb{Q}}$ all distinct from each other. Now, we don't need to extend fully to $\bar{\mathbb{Q}}$, since e_1, e_2, e_3 are the roots of $f(x, 0)$, we see that the field extension $K = \mathbb{Q}(e_1, e_2, e_3)$ has e_1, e_2, e_3 already. Moreover, we simply see that $K/\mathbb{Q}$ is a finite extension, so K is an algebraic number field. This allows us to define the map ϕ for any elliptic curve $E/\mathbb{Q}$ with a convenient choice of K. Further, we see that K is separable, but it might not be normal, so we keep adding a finite number of roots of minimal polynomials that are not in K yet. This gives a number field K which makes $K/\mathbb{Q}$ a finite Galois extension.

The goal of this section is to show that this map ϕ is a homomorphism, has the kernel of $2E(K)$, and the image of this map is finite. Once this is done, by the fundamental homomorphism theorem for groups we would obtain

$$E(K)/2E(K) \cong \text{im}(\phi)$$

so that $E(K)/2E(K)$ would be finite. Furthermore, with another little result, we can claim that $E(\mathbb{Q})/2E(\mathbb{Q})$ is also finite.

Proposition 60. *The map ϕ is a homomorphism of groups.*

Proof. (Edited from the first part of [24, 8.14] and a part of [10, 8.11]) We use the same trick as before: instead of proving that $\phi(P + Q) = \phi(P)\phi(Q)$ for all $P, Q \in E(K)$, let us show that for any $P, Q, R \in E(K)$ lying on the same line, $\phi(P)\phi(Q)\phi(R) = 1$ in $(K^\times/K^{\times\text{sq}})^3$.

Consider the simple case where at least one of P, Q, R is equal to O. Without loss of generality let that be R, then $P = -Q$, but $\phi(-Q) = \phi(Q)$ because the map $Q \mapsto -Q$ fixes the x-coordinate and negates the y-coorinate, giving the same value in the definition of ϕ. This means

$$\phi(P)\phi(Q)\phi(R) = \phi(-Q)\phi(Q) \cdot 1 = \phi(Q)^2 = 1$$

in $K^{\times \mathrm{sq}})^3$. Now assume that none of the P, Q, R coincide with O. This means the line passing through three of them cannot be a vertical line. So we assume such line is given by

$$y = \lambda x + \mu.$$

Since the points P, Q, R lie on the intersection of this line and the curve $y^2 = (x-e_1)(x-e_2)(x-e_3)$, we see that x_P, x_Q, x_R satisfies

$$(\lambda x + \mu)^2 = \underbrace{(x-e_1)(x-e_2)(x-e_3)}_{x^3+Ax+B}.$$

Do a change of variable $u \leftarrow x - e_i$ to see that the equation

$$(\lambda(u+e_i)+\mu)^2 = (u+e_i)^3 + A(u+e_i) + B$$

has three roots: $x_P - e_i, x_Q - e_i, x_R - e_i$. Expanding the equation, one sees that

$$\lambda^2(u^2 + 2ue_i + e_i^2) + 2\lambda(u+e_i)\mu + \mu^2 = u^3 + 3u^2e_i + 3ue_i^2 + e_i^3 + Au + Ae_i + B.$$

Apply Vieta's formula for the product of roots to see that

$$(x_P - e_i)(x_Q - e_i)(x_R - e_i) = \lambda^2 e_i^2 + 2e_i\mu + \mu^2 - e_i^3 - Ae_i - B = \underbrace{(\lambda e_i + \mu)^2}_{\in K^{\times \mathrm{sq}}} - \underbrace{(e_i^3 + Ae_i + B)}_{0} = 1.$$

Each component of $\phi(P)\phi(Q)\phi(R)$ is 1, so $\phi(P)\phi(Q)\phi(R)$ equals 1. This completes the proof. □

Proposition 61. *The kernel* $\ker(\phi)$ *is* $2E(K)$.

Proof. (This follows the proof given in the second part of [24, 8.14].)

($\subseteq$) We see that O is in both sets. Suppose $P \in \ker(\phi)$ different from O, i.e. $P = (x_P, y_P)$ with $x_P - e_i = v_i^2$ for some $v_1, v_2, v_3 \in K^\times$. Then let us show that P can be written as $[2]Q$ for some $Q \in E(K)$. Since we've assumed E is in the form $y^2 = x^3 + Ax + B$, apply Vieta to see that $e_1 + e_2 + e_3 = 0$. Now, the points (e_i, v_i) for $i = 1, 2, 3$ denotes a set of points at different x-coordinates. Apply polynomial interpolation to see that there exists a unique quadratic polynomial f satisfying

$$f(e_i) = v_i \text{ for } i = 1, 2, 3.$$

Since the polynomial

$$f(T) = \sum_{\text{cyc}} v_i \frac{1}{(e_i - e_j)(e_i - e_k)}(T - e_j)(T - e_k)$$

satisfies the property, it must be this polynomial. Define $g(T) = x_P - T - f(T)^2$. Then $g(e_i) = 0$ for $i = 1, 2, 3$. Hence g is divisible by $(T-e_1)(T-e_2)(T-e_3) = T^3 + AT + B$. Working modulo T^3+AT+B, i.e., in the ring $K[T]/(T^3+AT+B)$, we see that g is zero in this ring, so $x_P - T - f(T)^2$ is zero in this ring. In other words, if $f(T) = u_0 + u_1T + u_2T^2$ then

$$x_P - T = (u_0 + u_1T + u_2T^2)^2$$

in the ring $K[T]/(T^3 + AT + B)$. Working in this ring, one can substitute $T^3 \leftarrow -AT - B$ and $T^4 \leftarrow -AT^2 - BT$. By direct computation in this ring,

$$\begin{aligned} x_P - T &= (u_0 + u_1T + u_2T^2)^2 \\ &= u_0^2 + 2u_0u_1T + (u_1^2 + 2u_0u_2)T^2 + 2u_1u_2T^3 + u_2^2T^4 \\ &= (u_0^2 - 2Bu_1u_2) + (2u_0u_1 - 2Au_1u_2 - Bu_2^2)T + (u_1^2 + 2u_0u_2 - Au_2^2)T^2. \end{aligned}$$

Since all the terms are in degree less than 3, the coefficients must equate, i.e.,

$$\begin{aligned} x_P &= u_0^2 - 2Bu_1u_2 \\ -1 &= 2u_0u_1 - 2Au_1u_2 - Bu_2^2 \\ 0 &= u_1^2 + 2u_0u_2 - Au_2^2. \end{aligned}$$

If $u_2 = 0$ then $u_1 = 0$ also, so $f(T)$ is constant which means $v_1 = v_2 = v_3$, so $x_P - e_i = v_i^2$ is constant, i.e. $e_1 = e_2 = e_3$, hence a contradiction. Now we see that $u_2 \neq 0$, so multiplying the last equation by u_1/u_2^3 to see that

$$0 = \frac{u_1^3}{u_2^3} + 2\frac{u_0 u_1}{u_2^2} - A\frac{u_1}{u_2}.$$

Multiply the second equation by $1/u_2^2$ to see that

$$-\frac{1}{u_2^2} = 2\frac{u_0 u_1}{u_2^2} - 2A\frac{u_1}{u_2} - B,$$

then subtract these equations to see that

$$\left(\frac{1}{u_2}\right)^2 = \left(\frac{u_1}{u_2}\right)^3 + A\frac{u_1}{u_2} + B.$$

Write $x_Q = u_1/u_2$ and $y_Q = 1/u_2$. The equation above tells us that $Q = (x_Q, y_Q)$ lies on $E(K)$. Then one can apply the duplication formula to see that $[2]Q = P$. Therefore, $\ker(\phi) \subseteq 2E(K)$.

($\supseteq$) If $P \in 2E(K)$ then $P = [2]Q$ for some $Q \in E(K)$. So $\phi(P) = \phi([2]Q) = \phi(Q+Q) = \phi(Q)^2 = 1$. That is, $P \in \ker(\phi)$. □

Proposition 62. *The image* $\mathrm{im}(\phi)$ *is finite.*

Proof. (Edited from one given in [10].) Let S be the set of prime ideals $\mathfrak{p}$ such that $\mathfrak{p}$ divides the ideal $2(4A^3 + 27B^2)\mathcal{O}_K$[1]. Since

$$4A^3 + 27B^2 = \mathrm{disc}(x^3 + Ax + B) = ((e_1 - e_2)(e_1 - e_3)(e_2 - e_3))^2,$$

we conclude that $2, e_1 - e_2, e_1 - e_3, e_2 - e_3$ are invertible in $\mathcal{O}_{K,S}$. Now we keep extending S until $\mathcal{O}_{K,S}$ becomes a PID using 58. Let $P \in E(K)$ and write $P = (u/t, v/s)$ with u, t coprime and v, s coprime. Plugging in the equation for E, we have

$$\left(\frac{v}{s}\right)^2 = \left(\frac{u}{t} - e_1\right)\left(\frac{u}{t} - e_2\right)\left(\frac{u}{t} - e_3\right),$$

which is

$$v^2 t^3 = s^2(u - e_1 t)(u - e_2 t)(u - e_3 t).$$

Since s and v are coprime, $s^2 \mid t^3$. Since t and u are coprime, $t^3 \mid s^2$. Without loss of generality, one can replace s and t with a suitable choice of s and t such that $t^3 = s^2$, and so there exists $d \in \mathcal{O}_{K,S}$ such that $t = d^2$ and $s = d^3$. Plugging in the solution, we reduce into

$$v^2 = (u - e_1 d^2)(u - e_2 d^2)(u - e_3 d^2).$$

Observe that the numbers $u - e_1 d^2, u - e_2 d^2, u - e_3 d^2$ are pairwise coprime. This is because if δ divides both $u - e_1 d^2$ and $u - e_2 d^2$, then it must also divide $e_2 u - e_1 e_2 d^2$ and $e_1 u - e_1 e_2 d^2$, and hence divides $(e_1 - e_2)u$. It would also divide $(e_1 - e_2)d^2$. Since $e_1 - e_2$ is a unit, this means δ divides u and d^2, hence also a unit. So they are indeed pairwise coprime. This means, for any prime ideal $\mathfrak{p}$, one valuates

$$2\nu_{\mathfrak{p}}(v) = \nu_{\mathfrak{p}}(v^2) = \nu_{\mathfrak{p}}(u - e_i d^2)$$

where i is either 1, 2 or 3. This proves that $\nu_{\mathfrak{p}}(u - e_i d^2)$ is even for each $i = 1, 2, 3$. Therefore, $u - e_i d^2$ can be written as a product of $\gamma_i \in \mathcal{O}_{K,S}^\times$ and $z_i \in {K^\times}^{\mathrm{sq}}$. This means

$$x_P - \alpha_i = \frac{u - e_i d^2}{t} = \gamma_i \underbrace{d^{-2} z_i^2}_{K^{\mathrm{sq}}}$$

lands in $\mathcal{O}_{K,S}^\times / {\mathcal{O}_{K,S}^\times}^{\mathrm{sq}}$. Now, by 59, we see that $\mathcal{O}_{K,S}^\times \cong (\mathcal{O}_{K,S}^\times)_{\mathrm{tors}} \times \mathbb{Z}^r$, and that ${\mathcal{O}_{K,S}^\times}^{\mathrm{sq}} \cong H \times \mathbb{Z}^r$ for some $H \leq (\mathcal{O}_{K,S}^\times)_{\mathrm{tors}}$. Since $(\mathcal{O}_{K,S}^\times)_{\mathrm{tors}}$ is finite, the quotient $\mathcal{O}_{K,S}^\times / {\mathcal{O}_{K,S}^\times}^{\mathrm{sq}}$ removes the factor $\mathbb{Z}^r$ and becomes finite. □

Proposition 63. *Suppose K is a field, and L/K is a Galois extension. Let $E/\mathbb{Q}$ be an elliptic curve. If $E(L)/2E(L)$ is finite then $E(K)/2E(K)$ is also finite.*

[1] For ideals, we say $\mathfrak{a}$ divides $\mathfrak{b}$ if $\mathfrak{b} \subseteq \mathfrak{a}$.

Proof. (Edited from one given in the beginning of [10, 8.11].) Consider the natural map $\eta\colon E(K)/2E(K) \to E(L)/2E(L)$. Let us show that its kernel is finite. Suppose $P \in \ker(\eta)$. This means there exists $Q \in E(L)$ such that $[2]Q = P$. We define a map c_P which takes each $\sigma \in \mathrm{Gal}(L/K)$ to $\sigma(Q) - Q$. We have $[2]c_P(\sigma) = [2](\sigma(Q) - Q) = \sigma([2]Q) - [2]Q = \sigma(P) - P = 0$. This means c_P is a map of $\mathrm{Gal}(L/K) \to E[2]$. Now we claim that $P \mapsto c_P$ is injective. If $c_P = c_{P'}$ then $\sigma(Q) - Q = \sigma(Q') - Q'$ for all $\sigma \in \mathrm{Gal}(L/K)$. This means $\sigma(Q' - Q) = Q' - Q$ for all $\sigma \in \mathrm{Gal}(L/K)$, i.e. $Q' - Q$ is fixed by the Galois group, so $Q' - Q \in E(K)$, so $P' - P \in 2E(K)$, i.e. $P = P'$ in $E(K)/2E(K)$. Now that we've proved that $P \mapsto c_P$ which is defined on $\ker(\eta) \to \mathscr{F}(\mathrm{Gal}(L/K), E[2])$, is injective, so

$$\#\ker(\eta) \leq \#\mathscr{F}(\mathrm{Gal}(L/K), E[2]) = (\#\mathrm{Gal}(L/K))^{\#E[2]}$$

which is finite. Now the kernel of η is finite, we can apply the fundamental homomorphism theorem to see that

$$(E(K)/2E(K))/\ker(\eta) \cong \mathrm{im}(\eta) \hookrightarrow E(L)/2E(L),$$

so $E(K)/2E(K)$ must be finite. □

We piece everything together in the following weak Mordell–Weil theorem

Theorem 64 (Weak Mordell–Weil)**.** *Let $E/\mathbb{Q}$ be an elliptic curve. Then $E(\mathbb{Q})/2E(\mathbb{Q})$ is finite.*

Proof. First we extend using a finite Galois extension $K/\mathbb{Q}$ which gives the base number field for all the theorems. Now we define the map $\phi\colon E(K) \to (K^\times/K^{\times\,\mathrm{sq}})^3$ as in the beginning of this subsection. By 60, ϕ is a homomorphism of groups. By 61, the kernel of ϕ is $2E(K)$. By 62, the image of ϕ is finite. This proves that the group

$$E(K)/2E(K)$$

is finite. By 63, since $K/\mathbb{Q}$ is a finite Galois extension, we see that

$$E(\mathbb{Q})/2E(\mathbb{Q})$$

is also finite. □

2.3 Mordell–Weil Theorem for $E/\mathbb{Q}$

Now that the weak Mordell–Weil theorem is finished, we are ready to descend. By the weak Mordell–Weil theorem, we can write $E(\mathbb{Q})/2E(\mathbb{Q})$ as the set of cosets

$$2E(\mathbb{Q}) + R_1, \ldots, 2E(\mathbb{Q}) + R_n$$

for some representatives $R_1, \ldots, R_n \in E(\mathbb{Q})$. The idea now is for any point $P \in E(\mathbb{Q})$, we try to write

$$P = R_{\alpha(1)} + [2]P_1$$

for some $\alpha(1) \in \{1, \ldots, n\}$. Then we continue

$$P_1 = R_{\alpha(2)} + [2]P_2$$

and so on, such that P_1 is in some way "less than" P, and P_2 is "less than" P_1, and so on, until P_m is "small enough". In this way, if we can show that there exists a constant, such that there are only finitely many points smaller than or equal to that constant, then we see that P can be written as a sum of points from that set and points from the set $\{R_i\}_{i=1}^n$. The measuring device will be called *height*, which will be introduced in the following subsection.

2.3.1 Height

For a number $\frac{a}{b} \in \mathbb{Q}$ with $\gcd(a, b) = 1$, we define

$$H\left(\frac{a}{b}\right) = \max(|a|, |b|) \qquad \text{and} \qquad h\left(\frac{a}{b}\right) = \log H\left(\frac{a}{b}\right).$$

In the context of elliptic curve $E/\mathbb{Q}$, if $(x, y) \in E(\mathbb{Q})$ then we define $h(x, y) := h(x)$ and $H(x, y) = H(x)$, with the special case of $H(O) = 1$ and $h(O) = 0$.

Further from that, we see that the height has some nice properties, but there can be a better function with even better properties.

Theorem 65 ([24, 8.18]). *Let $E/\mathbb{Q}$ be an elliptic curve. There exists a function $\hat{h}\colon E(\mathbb{Q}) \to \mathbb{R}_{\geq 0}$ with the following properties.*

1. *$\hat{h}(P) \geq 0$ for all $P \in E(\mathbb{Q})$.*
2. *There is a constant c_0 such that $|\frac{1}{2}h(P) - \hat{h}(P)| \leq c_0$ for all P.*
3. *Given a constant c, there are only finitely many points $P \in E(\mathbb{Q})$ with $\hat{h}(P) \leq c$.*
4. *$\hat{h}([m]P) = m^2\hat{h}(P)$ for all integers m and all points P.*
5. *$\hat{h}(P+Q) + \hat{h}(P-Q) = 2\hat{h}(P) + 2\hat{h}(Q)$ for all P, Q.*
6. *$\hat{h}(P) = 0$ if and only if P is a torsion point.*

In order to establish the proof of existence of such function, we go through a short sequence of lemmas.

Lemma 66 ([24, 8.20]). *Let $c_1, c_2, d_1, d_2 \in \mathbb{Z}$. Then*

$$\max(|c_1|, |d_1|)\max(|c_2|, |d_2|) \leq 2\max(|c_1c_2|, |c_1d_2 + c_2d_1|, |d_1d_2|).$$

Proof. One can do a direct case analysis, which gives an elementary proof. See [24, 8.20]. □

Lemma 67 ([24, 8.21]). *Let $c_1, c_2, d_1, d_2 \in \mathbb{Z}$ with $\gcd(c_1, d_1) = \gcd(c_2, d_2) = 1$. Then*

$$\gcd(c_1c_2, c_1d_2 + c_2d_1, d_1d_2) = 1.$$

Proof. Simple case analysis. See [24, 8.21]. □

Now the current goal is to bound the quantity

$$|h(P+Q) + h(P-Q) - 2h(P) - 2h(Q)|$$

by a constant. To do this, we bound $h(P+Q) + h(P-Q) - 2h(P) - 2h(Q)$ from above and from below.

Lemma 68 (from above). *For a fixed $E/\mathbb{Q}$, there exists a constant $c'' > 0$ such that*

$$h(P+Q) + h(P-Q) \leq 2h(P) + 2h(Q) + c''$$

for all $P, Q \in E(\mathbb{Q})$.

Proof. (Given by [24, 8.19].) Suppose $E/\mathbb{Z}$ is given by $y^2 = x^3 + Ax + B$ with $A, B \in \mathbb{Z}$. Let $P = (a_1/b_1, y_1)$ and $Q = (a_2/b_2, y_2)$ be points on $E(\mathbb{Q})$ so that $P + Q = (a_3/b_3, y_3)$ and $P - Q = (a_4/b_4, y_4)$ have $y_1, y_2, y_3, y_4 \in \mathbb{Q}$ with $a_1, a_2, a_3, a_4, b_1, b_2, b_3, b_4 \in \mathbb{Z}$ with $\gcd(a_1, b_1) = \gcd(a_2, b_2) = \gcd(a_3, b_3) = \gcd(a_4, b_4) = 1$. Let

$$\begin{aligned} g_1 &= 2(a_1b_2 + a_2b_1)(Ab_1b_2 + a_1a_2) + 4Bb_1^2b_2^2 \\ g_2 &= (a_1a_2 - Ab_1b_2)^2 - 4B(a_1b_2 + a_2b_1)b_1b_2 \\ g_3 &= (a_1b_2 - a_2b_1)^2. \end{aligned}$$

So that

$$\frac{a_3}{b_3} + \frac{a_4}{b_4} = \frac{g_1}{g_3} \qquad \text{and} \qquad \frac{a_3a_4}{b_3b_4} = \frac{g_2}{g_3}. \qquad (\star)$$

Apply 67 on a_3, a_4, b_3, b_4 to see that $\gcd(a_3a_4, a_3b_4 + b_3a_4, b_3b_4) = 1$. Therefore, there exists $x, y, z \in \mathbb{Z}$ such that

$$a_3a_4x + (a_3b_4 + b_3a_4)y + b_3b_4z = 1.$$

Multiplying both sides by g_3, we have

$$g_3(a_3a_4)x + g_3(a_3b_4 + b_3a_4)y + g_3(b_3b_4)z = g_3,$$

and substituting from $(\star)$, we have

$$g_2(b_3b_4)x + g_1(b_3b_4)y + g_3(b_3b_4)z = g_3.$$

Therefore, b_3b_4 divides g_3, so $|b_3b_4| \leq |g_3|$. Similarly, $|a_3a_4| \leq |g_2|$. Using $(\star)$ again, one sees that

$$|a_3b_4 + a_4b_3| \leq |g_1|.$$

This means (implicitly applying 66),

$$\begin{aligned} H(P+Q)H(P-Q) &= \max(|a_3|,|b_3|)\max(|a_4|,|b_4|) \\ &\leq 2\max(|a_3a_4|, |a_3b_4 + a_4b_3|, |b_3b_4|) \\ &\leq 2\max(|g_2|,|g_1|,|g_3|). \end{aligned}$$

Let $H_1 = \max(|a_1|,|b_1|)$ and $H_2 = \max(|a_2|,|b_2|)$, then

$$\begin{aligned} |g_1| &= |2(a_1b_2 + a_2b_1)(Ab_1b_2 + a_1a_2) + 4Bb_1^2b_2^2| \\ &\leq 2(H_1H_2 + H_2H_1)(|A|H_1H_2 + H_1H_2) + 4|B|H_1^2H_2^2 \\ &\leq 4(|A| + 1 + |B|)H_1^2H_2^2. \end{aligned}$$

Similarly,

$$|g_2| \leq ((1+|A|)^2 + 8|B|)H_1^2H_2^2, \qquad |g_3| \leq 4H_1^2H_2^2.$$

Therefore,

$$H(P+Q)H(P-Q) \leq CH_1^2H_2^2 = CH(P)^2H(Q)^2$$

for some constant $C > 0$. Taking logarithms gives the desired inequality. □

Lemma 69 (from below)**.** *For a fixed $E/\mathbb{Q}$, there exists a constant $c' > 0$ such that*

$$2h(P) + 2h(Q) - c' \leq h(P+Q) + h(P-Q)$$

for all $P, Q \in E(\mathbb{Q})$.

Proof. This follows [24, 8.22]. Fix an elliptic curve $E/\mathbb{Q}$. Let us first show that there exists a constant C_2 such that for all $R \in E$,

$$4h(R) \leq h([2]R) + C_2.$$

Write $R = (a/b, y)$ with $a, b \in \mathbb{Z}, y \in \mathbb{Q}$, and $\gcd(a,b) = 1$. Define

$$\begin{aligned} h_1 &= a^4 - 2Aa^2b^2 - 8Bab^3 + A^2b^4 \\ h_2 &= (4b)(a^3 + Aab^2 + Bb^3) \\ \Delta &= 4A^3 + 27B^2 \\ r_1 &= 12a^2b + 16Ab^3 \\ r_2 &= 27Bb^3 + 5Aab^2 - 3a^3 \\ s_1 &= 4\Delta a^3 - 4A^2Ba^2b + 4A(3A^3 + 22B^2)ab^2 + 12B(A^3 + 8B^2)b^3 \\ s_2 &= A^2Ba^3 + A(5A^3 + 32B^2)a^2b + 2B(13A^3 + 96B^2)ab^2 - 3A^2(A^3 + 8B^2)b^3, \end{aligned}$$

so that $r_1, r_2, s_1, s_2 \in \mathbb{Z}[a,b]$ (one can see them as polynomials in two variables, then plug a and b into the polynomials to retrieve the numbers) such that

$$\begin{aligned} 4\Delta b^7 &= r_1h_1 + r_2h_2 \\ 4\Delta a^7 &= s_1h_1 + s_2h_2. \end{aligned}$$

In this view, r_1, r_2, s_1, s_2 are homogeneous polynomials over a, b of degree 3. But for any homogeneous polynomial

$$p(x,y) = c_0x^3 + c_1x^2y + c_2xy^2 + c_3y^3,$$

we have

$$|p(a,b)| \leq (|c_0| + |c_1| + |c_2| + |c_3|)\max(|a|,|b|)^3.$$

Suppose $|b| \geq |a|$, then

$$\begin{aligned} |4\Delta||b|^7 &\leq |r_1(a,b)||h_1| + |r_2(a,b)||h_2| \\ &\leq C_1|b|^3\max(|h_1|,|h_2|), \end{aligned}$$

for some constant C_1 independent of R. Therefore,

$$|4\Delta||b|^4 \leq C_1 \max(|h_1|, |h_2|).$$

Let $d = \gcd(h_1, h_2)$. Then d divides $r_1h_1 + r_2h_2 = 4\Delta b^7$, and d divides $s_1h_1 + s_2h_2 = 4\Delta a^7$. Since $\gcd(a, b) = 1$, we have d divides 4Δ and so $d \leq |4\Delta|$. We have, by the duplication formula, that

$$\begin{aligned} H([2]R) &= H(x_{[2]R}) \\ &= H\left(\frac{x^4 - 2Ax^2 - 8Bx + A^2}{4x^3 + 4Ax + 4B}\right) \\ &= H\left(\frac{a^4 - 2Aa^2b^2 - 8Bab^3 + A^2b^4}{4b(a^3 + 4Aab^2 + 4Bb^3)}\right) \\ &= H(h_1/h_2) \\ &= \max\left(\frac{|h_1|}{d}, \frac{|h_2|}{d}\right). \end{aligned}$$

Therefore,

$$\begin{aligned} |4\Delta|H(R)^4 = |4\Delta||b|^4 &\leq C_1 \max(|h_1|, |h_2|) \\ &\leq C_1|4\Delta| \max\left(\frac{|h_1|}{d}, \frac{|h_2|}{d}\right) \\ &\leq C_1|4\Delta|H([2]R). \end{aligned}$$

Dividing by $|4\Delta|$ and taking logarithms gives

$$4h(R) \leq h([2]R) + C_2$$

for some constant C_2 independent of R. The case where $|a| \geq |b|$ is similar. Apply 68 for $P \leftarrow P+Q$ and $Q \leftarrow P - Q$ to see that

$$h([2]P) + h([2]Q) \leq 2h(P + Q) + 2h(P - Q) + c''.$$

Then we apply the claim above to see that

$$4h(P) + 4h(Q) - 2C_2 \leq 2h(P + Q) + 2h(P - Q) + c'',$$

so we pick $c' := C_2 + \frac{c''}{2}$ to see that

$$2h(P) + 2h(Q) - c' \leq h(P + Q) + h(P - Q).$$

This completes the proof. □

Lemma 70 ([24, 8.19])**.** *For a fixed $E/\mathbb{Q}$, there exists a constant $c_1 > 0$ such that*

$$|h(P + Q) + h(P - Q) - 2h(P) - 2h(Q)| \leq c_1$$

for all $P, Q \in E(\mathbb{Q})$.

Proof. Apply 69 and 68. □

Now we're ready to prove the main theorem 65.

Proof of Theorem 65. (This follows [24, 8.18].) We explicitly define

$$\hat{h}(P) := \frac{1}{2} \lim_{n\to\infty} \frac{1}{4^n} h([2^n]P).$$

To prove that the limit exists, we have

$$\lim_{n\to\infty} \frac{1}{4^n} h([2^n]P) = h(P) + \sum_{j=1}^{\infty} \frac{1}{4^j}(h([2^j]P) - 4h([2^{j-1}]P)).$$

By the main lemma 70, plugging in $Q \leftarrow P$ to see that

$$|h([2]P) - 4h(P)| \leq c_1$$

for all $P \in E(\mathbb{Q})$, and so the quantity $h([2^j]P) - 4h([2^{j-1}]P)$ is bounded by c_1. This means the infinite sum is bounded by sums of $\frac{c_1}{4^j}$, hence converges. Therefore, $\hat{h}(P)$ exists. We're left with just simply checking the properties of $\hat{h}$ defined in this way. See [24, 8.18] for full proof. □

2.3.2 Descent

Having a well-defined $\hat{h}$, we're ready to descend.

Theorem 71 (Mordell–Weil for $E/\mathbb{Q}$)**.** *For any $E/\mathbb{Q}$, the group $E(\mathbb{Q})$ is finitely generated.*

Proof. (This follows the one given in [24].) Let $R_1, \ldots, R_n \in E(\mathbb{Q})$ be the representatives for elements of $E(\mathbb{Q})/2E(\mathbb{Q})$. Let

$$c = \max_{1 \le i \le n} \hat{h}(R_i)$$

and let $\{Q_1, \ldots, Q_m\}$ be the set of points with $\hat{h}(Q_i) \le c$. This is a finite set by Theorem 65. Let G be the subgroup of $E(\mathbb{Q})$ generated by $R_1, \ldots, R_n, Q_1, \ldots, Q_m$. We claim that $G = E(\mathbb{Q})$. Suppose not, then $E(\mathbb{Q}) \setminus G$ is nonempty. Pick a random point in this set and we see that there are only finitely many points of height less than this point. Let P be the point with the smallest height in this collection of finitely many points. Observe that $P \in E(\mathbb{Q})$ so it can be projected into a unique class $2E(\mathbb{Q}) + R_i$ for some $1 \le i \le n$, so that $P - R_i \in 2E(\mathbb{Q})$. Write

$$P - R_i = [2]Q$$

for some $Q \in E(\mathbb{Q})$. By 65,

$$\begin{aligned} 4\hat{h}(Q) &= \hat{h}([2]Q) \\ &= \hat{h}(P - R_i) \\ &= 2\hat{h}(P) + 2\hat{h}(R_i) - \hat{h}(P + R_i) \\ &\le 2\hat{h}(P) + 2c + 0 \\ &< 2\hat{h}(P) + 2\hat{h}(P) = 4\hat{h}(P). \end{aligned}$$

This proves that Q has smaller height than P, so it must be in G. But $R_i \in G$ so $P = [2]Q + R_i \in G$ also, a contradiction. This proves that $E(\mathbb{Q}) = G$, and completes the proof of the Mordell–Weil theorem. □

Remark. *The actual Mordell–Weil theorem gives the result in a more general settings, namely for all elliptic curves over algebraic number fields. Unfortunately this is beyond the scope of the thesis, but we note that for a given number field K one can extend into another L such that L/K is a finite Galois extension, so that E can be written as*

$$E\colon y^2 = (x - e_1)(x - e_2)(x - e_3)$$

for some $e_1, e_2, e_3 \in L$. Then the weak Mordell–Weil theorem applies using the map

$$\phi\colon (L^\times / L^{\times \mathrm{sq}})^3.$$

The descent from the weak Mordell–Weil theorem to the full Mordell–Weil theorem requires a more general notion of height, defined on algebraic number fields.

Remark. *Once the Mordell–Weil theorem is proved. We apply the classification of finitely generated abelian groups to see that any elliptic curve $E/\mathbb{Q}$ has the group $E(\mathbb{Q})$ isomorphic to*

$$\mathbb{Z}^r \times E(\mathbb{Q})_{\mathrm{tors}}$$

with finite $E(\mathbb{Q})_{\mathrm{tors}}$. Lutz–Nagell theorem allows a computation for $E(\mathbb{Q})_{\mathrm{tors}}$. However, as of today there are still no known general algorithm to compute the integer r, which is called the rank *of the elliptic curve. This mysterious quantity plays a role in more important problems in the modern theory of elliptic curves, including the Birch–Swinnerton-Dyer conjecture. The theory behind the computations of the rank reamins interesting in its own right, and perhaps, worth discussing in another day.*

Appendix A

Implementation

The implementation mostly follows [22, Chapter IX]. Arithmetics of finite fields follows [11].

A.1 Arithmetics of finite fields

For the case of $\mathbb{F}_p$ with p prime, elements are kept as an integer representative between 0 to $p-1$. Addition and multiplication can be done modulo p directly. The additive inverse of $x \in \{0, \dots, p-1\}$ is $p-x$ if $x \neq 0$ and 0 otherwise. The multiplicative inverse of $x \in \{0, \dots, p-1\}$ is $x^{p-2} \bmod p$, which can be computed using at most $2 \log p$ multiplications in $\mathbb{F}_p$.

For the more general case of $\mathbb{F}_q$ with $q = p^r$, where p is prime and $r > 1$ is an integer, we use the following theorem.

Theorem 72. *For a prime number p and an integer $r \geq 1$, $\mathbb{F}_{p^r} \cong \mathbb{F}_p[T]/(\mu)$ for any irreducible polynomial μ of degree exactly r.*

Proof. We assume the basic fact from algebra that every finite fields of the same order are isomorphic. Hence we're left with proving that $\mathbb{F}_p[T]/(\mu)$ is a field of order p^n. This is actually simple. Recall the following results from algebra:

- The quotient of a PID by a maximal ideal gives a field.
- In a PID R, an element $x \in R$ is irreducible if and only if (x) is a maximal ideal.

Since $\mu \in \mathbb{F}_p[T]$ is irreducible, (μ) is maximal, so $\mathbb{F}_p[T]/(\mu)$ is a field. Now we enumerate the elements $\sum_{i=0}^{r-1} a_i T^i$ with $(a_i)_{i=0}^{r-1} \in \mathbb{F}_p^r$. No two elements of these coincide, and any other elements of $\mathbb{F}_p[T]$ must have degree at least r, which would coincide (using the identification from the quotient ring) with some element in this enumeration. This means $\#(\mathbb{F}_p[T]/(\mu)) = p^r$ and hence completes the proof. □

By this theorem, we compute objects in $\mathbb{F}_q$ using $\mathbb{F}_p[T]$ as a representative for each element, and we do computation modulo the ideal (μ) for a fixed irreducible $\mu \in \mathbb{F}_p[T]$. Addition and multiplication corresponds to polynomial addition (modulo μ). The additive inverse is the polynomial with each of the coefficients its additive inverse in $\mathbb{F}_p$. The multiplicative inverse is done by the Itoh–Tsujii algorithm [8] described below (following the details given by [17]).

A.1.1 Itoh–Tsujii Algorithm

Here we describe an algorithm to invert an element of $\mathbb{F}_q = \mathbb{F}_{p^n} \cong \mathbb{F}_p[T]/(\mu)$. In actual computation, the algorithm receives an input $f \in \mathbb{F}_p[T]$ and returns an output $g \in \mathbb{F}_p[T]$ such that $f(T)g(T) - 1 \in (\mu)$.

The idea is that if $\alpha \in \mathbb{F}_{p^n}^\times$ then $\alpha^{-1} = \frac{\alpha^{r-1}}{\alpha^r}$ for any $r \geq 1$. We choose an appropriate r, say $r = \sum_{i=0}^{n-1} p^i = \frac{p^n-1}{p-1}$. Then we claim that $\alpha^r \in \mathbb{F}_p$ for all $\alpha \in \mathbb{F}_{p^n}^\times$. Once this is done then α^{-1} is just the product of $\frac{1}{\alpha^r}$ (invert α^r in $\mathbb{F}_p$) and α^{r-1} (compute this using Frobenius endomorphisms). Let us prove the claim first.

Proposition 73. *For a fixed prime p and an integer $n \geq 2$, for any $\alpha \in \mathbb{F}_{p^n}^\times$, let $r = \frac{p^n-1}{p-1} \in \mathbb{N}$, then we have $\alpha^r \in \mathbb{F}_p$.*

Proof. Well $(\alpha^r)^{p-1} = \alpha^{p^n-1} = 1$ by Lagrange's theorem on $\mathbb{F}_{p^n}^\times$. So α^r is a root of $X^{p-1} - 1 = 0$. But consider $\mathbb{F}_p^\times \subseteq \mathbb{F}_{p^n}^\times$ and by Lagrange's theorem on $\mathbb{F}_p^\times$, $a^{p-1} = 1$ for all $a \in \mathbb{F}_p^\times$. This gives a

list of roots of $X^{p-1}-1=0$ in $\mathbb{F}_p$, which is precisely $\mathbb{F}_p^\times$. There are $p-1$ elements on this list, and the polynomial has degree $p-1$, so there cannot be other elements in $\mathbb{F}_{p^n}$ satisfying this equation. Since α^r satisfies the equation, it belongs to the list, i.e. $\alpha^r \in \mathbb{F}_p^\times \subseteq \mathbb{F}_p$. □

Now that the claim is finished, we consider the method (following [17]) to compute α^{r-1} from α. Consider the following map

$$\Phi_j : \begin{cases} \mathbb{F}_{p^n} & \to \mathbb{F}_{p^n} \\ \alpha & \mapsto \alpha^{p^j}. \end{cases}$$

In other words, the action of Φ_j is repeatedly applying Frob_p for j times. Now we consider that $r-1=\sum_{i=1}^{n-1} p^i$. We do the following process to define the sequence $(t_j)_{j\in\mathbb{N}}$ as

$$\begin{aligned} t_1 &= \Phi_1(\alpha) = \alpha^p \\ t_2 &= t_1\Phi_1(t_1) = \alpha^{p+p^2} \\ t_3 &= t_2\Phi_2(t_2) = \alpha^{p+p^2+p^3+p^4} \\ t_4 &= t_3\Phi_4(t_3) = \alpha^{p+p^2+p^3+p^4+\cdots+p^8} \\ &\vdots \quad \vdots \quad \vdots \end{aligned}$$

Concretely, construct $t_1=\Phi_1(\alpha)$ and for $i=2,\ldots,\lceil\log_2(r)\rceil$, construct

$$t_i := t_{i-1}\Phi_{2^{i-2}}(t_{i-1}).$$

And so we can recover $\alpha^{p+p^2+\cdots+p^{n-1}} = \alpha^{r-1}$ in no more than $2\log_2(r)$ multiplications within $(t_i)_{i=1}^{\lceil\log_2(r)\rceil}$. Consider the following pseudocode.

Algorithm 6 Itoh–Tsujii Algorithm

Require: $n\geq 2$, p prime, $f\in\mathbb{F}_p[T]$, μ monic and irreducible of degree n
Ensure: $f(T)g(T)-1\in(\mu)$
  $r \leftarrow \frac{p^n-1}{p-1}$
  $L \leftarrow \lceil\log_2(r)\rceil$
  $t_1 \leftarrow \Phi_1(f(T)) \bmod (\mu)$
  for $i=2,\ldots,L$ **do**
    $t_i \leftarrow t_{i-1}\Phi_{2^{i-2}}(t_{i-1}) \bmod (\mu)$
  end for
  $G \leftarrow n-1$ ▷ We will compute $f(T)^{p+p^2+\cdots+p^G}$.
  $g \leftarrow 1$
  while $G>0$ **do** ▷ Loop invariant: $g(T)f(T)^{p+p^2+\cdots+p^G} \bmod (\mu) = 1$.
    $B \leftarrow \lfloor\log_2(G)\rfloor + 1$ ▷ The quantity $B=\lfloor\log_2(G)\rfloor+1$ is the bit length of G.
    $g \leftarrow g\Phi_{G-2^{B-1}}(t_B) \bmod (\mu)$
    $G \leftarrow G-2^{B-1}$ ▷ The quantity 2^{B-1} denotes the most significant bit.
  end while

The algorithm given by Algorithm 6 is implemented to invert $\alpha\in\mathbb{F}_{p^n}^\times$, represented by $f\in\mathbb{F}_p[T]$ modulo (μ) into $g\in\mathbb{F}_p[T]$ so that $fg\equiv 1$ modulo (μ).

A.1.2 Shanks–Tonelli Algorithm

The goal of this subsection is to compute a square root (if exists) of an element in a finite field. We begin with Euler's criterion.

Theorem 74 (Euler's criterion)**.** *Let p be an odd prime and $n\in\mathbb{N}^*$. For a fixed $\alpha\in\mathbb{F}_{p^n}^\times$, there exists $\beta\in\mathbb{F}_{p^n}^\times$ such that $\alpha=\beta^2$ if and only if $\alpha^{\frac{p^n-1}{2}}=1$ in $\mathbb{F}_{p^n}$.*

Proof. We recall the fact that $\mathbb{F}_{p^n}^\times \cong \mathbb{Z}/(p^n-1)\mathbb{Z}$ as a group isomorphism, so by Lagrange's theorem, for any $\alpha\in\mathbb{F}_{p^n}^\times$, $\alpha^{p^n-1}=1$ in $\mathbb{F}_{p^n}$. Since p is an odd prime, $\frac{p^n-1}{2}$ is a well-defined integer so that $0=\alpha^{p^n-1}-1=(\alpha^{\frac{p^n-1}{2}}-1)(\alpha^{\frac{p^n-1}{2}}+1)$ in $\mathbb{F}_{p^n}$. Since $\mathbb{F}_{p^n}$ is a field, it is a domain in particular, so at least one of $\alpha^{\frac{p^n-1}{2}}-1$ and $\alpha^{\frac{p^n-1}{2}}+1$ must be zero.

Now, if there exists $\beta\in\mathbb{F}_{p^n}^\times$ such that $\beta^2=\alpha$, then $\alpha^{\frac{p^n-1}{2}}=\beta^{p^n-1}=1$ in $\mathbb{F}_{p^n}$ (the last equality follows from Lagrange's theorem once again), this proves the forward direction.

Observe that what we've just proved is that if an element is a nonzero square then it is a root of $X^{\frac{p^n-1}{2}} - 1 = 0$ in $\mathbb{F}_{p^n}$. Now we can list all the possible nonzero squares:

$$1^2, 2^2, \ldots, \left(\frac{p^n-1}{2}\right)^2, \left(\frac{p^n+1}{2}\right)^2, \ldots, (p^n-1)^2$$

but we're sure that $(p^n + x)^2 = x^2$ in $\mathbb{F}_{p^n}$, and furthermore, $(p^n - x)^2 = x^2$ in $\mathbb{F}_{p^n}$ also. This means, $1^2, \ldots, \left(\frac{p^n-1}{2}\right)^2$ is the complete list of nonzero squares in $\mathbb{F}_{p^n}$. Also if $x^2 = y^2$ then $(x - y)(x + y) = 0$, so these list has pairwise distinct elements since for any pair, the difference and the sum are both nonzero. This means we've identified the $\frac{p^n-1}{2}$ nonzero squares. By what we've proved earlier, these are roots of $X^{\frac{p^n-1}{2}} - 1 = 0$, which is a polynomial equation of degree $\frac{p^n-1}{2}$, giving at most $\frac{p^n-1}{2}$ roots. Since we gave exactly $\frac{p^n-1}{2}$ roots, we see that any other element outside the list fails to satisfy the equation $X^{\frac{p^n-1}{2}} - 1 = 0$. This means, for any $\alpha \in \mathbb{F}_{p^n}^\times$, if there is no $\beta \in \mathbb{F}_{p^n}^\times$ such that $\beta^2 = \alpha$, then α does not belong to the list of nonzero squares, and hence fails to satisfy the equation $X^{\frac{p^n-1}{2}} - 1 = 0$. So $\alpha^{\frac{p^n-1}{2}} \neq 1$. □

Remark. *Furthermore, for an element $\alpha \in \mathbb{F}_{p^n}^\times$, if $\alpha^{\frac{p^n-1}{2}} \neq 1$ in $\mathbb{F}_{p^n}$, then $\alpha^{\frac{p^n-1}{2}} = -1$ for sure, because $\alpha^{\frac{p^n-1}{2}}$ is a root of $X^2 - 1 = 0$ in $\mathbb{F}_{p^n}$, since $\alpha^{p^n-1} = 1$ in $\mathbb{F}_{p^n}$ by Lagrange's theorem on $\mathbb{F}_{p^n}^\times$.*

Now let us consider the Shanks–Tonelli algorithm, which will be used to compute a square root of α in $\mathbb{F}_{p^n}$, assuming that there exists one (we check using the value of $\alpha^{\frac{p^n-1}{2}}$ by Euler's criterion). First, we try to find an element $z \in \mathbb{F}_{p^n}^\times$ such that it is not a square, i.e. $z^{\frac{p^n-1}{2}} = -1$ in $\mathbb{F}_{p^n}$ by the previous remark. This is possible since $\frac{p^n-1}{2}$ elements are squares and 1 element is zero, so the other $\frac{p^n-1}{2}$ elements are nonsquares. By randomly picking an element from $\mathbb{F}_{p^n}^\times$, there is a $\frac{1}{2}$ chance of getting a nonsquare. Now that we assume that we've constructed a nonsquare $z \in \mathbb{F}_{p^n}^\times$, first we write $p^n - 1$ as $Q2^S$ where $S := \nu_2(p^n - 1)$. So Q is an odd integer. Now let $R := \alpha^{\frac{Q+1}{2}}$ then $R^2 = (\alpha^Q)(\alpha)$. If $\alpha^Q = 1$ then we're done: return R so that $R^2 = \alpha$. Otherwise, we have the following loop invariant (at $i = 0$ at first, and for $t = \alpha^Q$):

$$\begin{cases} R, t \in \mathbb{F}_{p^n}^\times; \\ R^2 = \alpha t; \\ t^{2^{S-i-1}} = 1. \end{cases}$$

(Note that for $i = 0$ and $t = \alpha^Q$, the last invariant holds because $t^{2^{S-i-1}} = \alpha^{\frac{Q2^S}{2}} = \alpha^{\frac{p^n-1}{2}}$ which equates to one by Euler's criterion.)

Now we will do a loop to refine R and t, i.e., find another R and t such that the loop invariant holds for $i+1$. Indeed, since we knew that $t^{2^{S-i-1}} = 1$, if $t^{2^{S-i-2}} = 1$ also then we don't have to do anything: the same R and t works for the next i. Otherwise, $t^{2^{S-i-2}}$ must be -1 since its square equals 1. The idea is to find a value $b \in \mathbb{F}_{p^n}^\times$ such that when we update $R \leftarrow Rb$ and $t \leftarrow tb^2$, the invariant holds. We can easily see that the invariant $R^2 = \alpha t$ holds no matter how we choose b to be. The important part is to choose b such that $(tb^2)^{2^{S-i-2}} = 1$, i.e.,

$$\underbrace{t^{2^{S-i-2}}}_{-1} b^{2^{S-i-1}} = 1,$$

that is, find b such that $b^{2^{S-i-1}} = -1$. This can be done using the knowledge of z. We knew that $z^{\frac{Q2^S}{2}} = z^{\frac{p^n-1}{2}} = -1$. If we let $B_0 = z^Q$ then $B_0^{2^{S-1}} = -1$. Then we do $B_1 = B_0^2$ to see that $B_1^{2^{S-2}} = -1$ and so on. Concretely, we can iteratively compute $B_i = B_{i-1}^2$ so that $B_i^{2^{S-i-1}} = -1$. And so in this stage, the problem of finding b become trivial–pick $b = B_i$. Continuing the main iteration like this, we end up at $i = S - 1$ and the loop invariant becomes

$$\begin{cases} R, t \in \mathbb{F}_{p^n}^\times; \\ R^2 = \alpha t; \\ t^{2^0} = 1. \end{cases}$$

In other words, $R^2 = \alpha$. This is concretely described by the following pseudocode of Algorithm 7.

Algorithm 7 Shanks–Tonelli Algorithm

Require: $n \in \mathbb{N}^*$, $p \geq 3$ prime, $\alpha \in \mathbb{F}_{p^n}^\times$, $\alpha^{\frac{p^n-1}{2}} = 1$
Ensure: $R^2 = \alpha$
  $z \leftarrow 1_{\mathbb{F}_{p^n}}$
  while $z^{\frac{p^n-1}{2}} \neq -1$ **do**
    $z \leftarrow$ random number in $\mathbb{F}_{p^n}^\times$
  end while
  $S \leftarrow \nu_2(p^n - 1)$
  $Q \leftarrow \frac{p^n-1}{S}$
  $B_0 \leftarrow z^Q$
  for $i = 1, \ldots, S-1$ **do**
    $B_i \leftarrow B_{i-1}^2$
  end for
  $R \leftarrow \alpha^{\frac{Q+1}{2}}$
  $t \leftarrow \alpha^Q$
  for $i = 1, \ldots, S-1$ **do**
    if $t^{2^{S-i-1}} = 1$ **then**
      $b \leftarrow 1$
    else
      $b \leftarrow B_{i-1}$
    end if
    $R \leftarrow Rb$
    $t \leftarrow tb^2$
  end for

A.2 Random points on an elliptic curve

An important implementation detail required to implement the algorithms (i.e. computing the Weil pairing) is "generate a random point from a given elliptic curve". For a given Weierstrass equation $E\colon y^2 = x^3 + Ax + B$, this can be done by randomly generating an integer $x \in \mathbb{F}_{p^n}$ then find the square root (if exists) of the number $x^3 + Ax + B \in \mathbb{F}_{p^n}$. If this exists, then we call it y and take (x, y) to be a random point on the curve. This is enough to find a point that doesn't coincide with a predefined set of points, by direct rejection sampling.

A.3 Source code

The source code is published at https://github.com/plumsirawit/comp-ec. The author regrets that the current implementation has two problems:

- The implementation for Schoof's algorithm isn't given. The difficulty lies in algebraic manipulation of the ring R_ℓ on elliptic curves, since not all elements are invertible. This is actually solved by giving a completely different method to compute $[q](x, y)$. The original paper of Schoof [21] has already given all the details for computation in R_ℓ. Also [12, Chapter II] gives useful results about computation with the division polynomials.
- The irreducibility test for polynomials is given only for polynomials of degree no more than three, since just "plugging in all possible roots to see if it factors" is enough to check the irreducibility. However, for a more general settings one has to use some stronger algorithms. See [11, Chapter 10] for a list of simple other options. Note that even if this irreducibility test is implemented, it loops over all elements of $\mathbb{F}_p$, hence taking exponential time in $\log p$.

Bibliography

[1] Leonard Adleman. A subexponential algorithm for the discrete logarithm problem with applications to cryptography. In *20th Annual Symposium on Foundations of Computer Science (sfcs 1979)*, pages 55–60, 1979.

[2] M. A. Armstrong. *Basic Topology*. Springer New York, 1983.

[3] E. Artin. Quadratische körper im gebiete der höheren kongruenzen. ii. analytischer teil. *Mathematische Zeitschrift*, 19(1):207–246, December 1924.

[4] Roman Bezrukavnikov. Algebraic geometry. Fall 2015. Massachusetts Institute of Technology: MIT OpenCouseWare, https://ocw.mit.edu/. License: Creative Commons BY-NC-SA.

[5] Kleber Carrapatoso. *MAA302: Topology and differential calculus*. 2023.

[6] Richard Crandall and Carl Pomerance. *Prime Numbers*. Springer, New York, NY, October 2010.

[7] Jyrki Lahtonen (https://math.stackexchange.com/users/11619/jyrki lahtonen). Extending homomorphism into algebraically closed field. Mathematics Stack Exchange. URL:https://math.stackexchange.com/q/898059 (version: 2014-08-15).

[8] Toshiya Itoh and Shigeo Tsujii. A fast algorithm for computing multiplicative inverses in GF(2^m) using normal bases. *Information and Computation*, 78(3):171–177, 1988.

[9] Diego Izquierdo. *MAA303: Algebra and arithmetic*. 2022.

[10] Diego Izquierdo. *MAT 562 : Introduction à la géométrie algébrique et courbes elliptiques*. 2024.

[11] John Kerl. Computation in finite fields. https://johnkerl.org/doc/ffcomp.pdf, 2004. [Accessed 13-02-2024].

[12] Serge Lang. *Elliptic Curves: Diophantine Analysis*. Springer Berlin Heidelberg, 1978.

[13] H. W. Lenstra. Factoring integers with elliptic curves. *The Annals of Mathematics*, 126(3):649, November 1987.

[14] A.J. Menezes, T. Okamoto, and S.A. Vanstone. Reducing elliptic curve logarithms to logarithms in a finite field. *IEEE Transactions on Information Theory*, 39(5):1639–1646, 1993.

[15] Stephen D. Miller and Ramarathnam Venkatesan. Spectral analysis of pollard rho collisions. In Florian Hess, Sebastian Pauli, and Michael Pohst, editors, *Algorithmic Number Theory*, pages 573–581, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

[16] Jürgen Neukirch. *Algebraic Number Theory*. Springer Berlin Heidelberg, 1999.

[17] Thomas Pornin. Itoh Tsuji algorithm. https://crypto.stackexchange.com/a/81143, 2020. [Accessed 13-02-2024].

[18] Alain Robert. *Elliptic Curves*. Springer Berlin Heidelberg, 1973.

[19] Joshua Ruiter. Infinite galois theory. https://users.math.msu.edu/users/ruiterj2/math/Documents/Notes%20and%20talks/Infinite%20Galois%20Theory.pdf, 2019. [Accessed 03-03-2024].

[20] Pitchayut Saengrungkongka. 18.785 lecture notes, 2023. Unpublished.

[21] René Schoof. Elliptic curves over finite fields and the computation of square roots mod p. *Mathematics of Computation*, 44(170):483–494, 1985.

[22] Joseph H. Silverman. *The Arithmetic of Elliptic Curves*. Springer New York, 2009.

[23] The Stacks project authors. The stacks project. https://stacks.math.columbia.edu, 2023.

[24] Lawrence C Washington. *Elliptic curves : number theory and cryptography – 2nd ed.* Discrete Mathematics and Its Applications. Chapman & Hall/CRC, 2008.