MSc Mathematics

Track: Geometry, Algebra and Number Theory

Master thesis

Isogeny-based cryptography

by

Alain Chavarri Villarello

July 20, 2022

Supervisor: Prof. Dr. Rob de Jeu

Second examiner: Dr. Sander Dahmen

Department of Mathematics

Faculty of Sciences

VU UNIVERSITY AMSTERDAM

Abstract

With the advent of quantum computers, isogeny-based cryptography has gained notoriety in the last decade. This thesis explores the mathematics of the post-quantum key exchange protocol CSIDH. We present the theory of elliptic curves and imaginary quadratic orders, and study the action of the class group of certain order $\mathcal{O}$ on the set of supersingular elliptic curves over the prime field $\mathbb{F}_p$ that have $\mathcal{O}$ as endomorphism ring. This class group action was shown to be free and transitive by Waterhouse [26] in a much more general setting, using techniques for general abelian varieties over finite fields. We provide some alternative proofs of these facts that avoid the use of such techniques. Furthermore, we explain how CSIDH works and discuss the requirements for its efficient implementation.

Popular Summary

The ability to communicate secretly is an essential part of our everyday life that we may sometimes take for granted. Everytime one makes a credit card transaction or sends a private email, cryptographic protocols are at play behind the scenes, making sure that no third party can access that information.

When two parties want to set up a secure channel of communication, they first have to agree on the key that they will use to encrypt and decrypt their messages. This step is not trivial at all, and it took many years to find a practical way to do this. However, there's a potential problem with the key exchange methods that are currently being employed: they are not resistant against quantum computers.

Sufficiently powerful quantum computers don't yet exist, but many believe that they will be a reality in the near future. Therefore, many cryptographers have been working on alternative methods to perform a key exchange. One of these methods is called CSIDH, and it's based on the theory of *elliptic curves* and their *isogenies*.

Generally speaking, an elliptic curve can be described as the set of points (x, y) that satisfy an equation of the form

$$y^2 = x^3 + Ax + B,$$

where A and B are some constants.

Elliptic curves have the very pleasant property of having an addition operation. Given two points in the curve, there is a way to 'add' them together and get a third point that's also in the curve. An isogeny is a function that sends a point in an elliptic curve to a point in another elliptic curve in a way that is compatible with the addition operation of each curve.

Given one of these curves, the collection of isogenies that go from the curve to itself is called the *endomorphism ring*. It turns out that there is a procedure based on this endomorphism ring that allows us to travel from an elliptic curve to another elliptic curve via an isogeny. Under certain conditions, this procedure –called the *class group action*– satisfies some very nice properties.

Thanks to these nice properties, the class group action can be used to perform a key exchange which can be efficiently computed. For this reason, it becomes the basis for the cryptographic protocol CSIDH, which is believed to be resistant against quantum attacks.

Acknowledgments

First and foremost, I would like to thank my supervisor Rob de Jeu for his time, guidance, and suggestions. I would also like to thank Sander Dahmen for agreeing to be the second reviewer of this thesis.

I want to thank my friends for their support and encouragement. I'm very lucky to have you in my life.

Finally, I would like to thank my family for their unwavering love and support. Despite the geographical distance, I always feel them close to me. Estoy eternamente agradecido.

Contents

Introduction

Suppose Alice wants to send a secret message to Bob. Before the 1970s, Alice and Bob would have to first exchange a key using a secure channel. Then Alice would encrypt her message with some type of cipher using the agreed key and send it to Bob. He would then use that same key to decrypt the message. This method quickly becomes impractical as the amount of participants increases, since establishing a secure channel to exchange a secret key is not a minor task.

A cryptographic revolution happened in the 1970s, when Whitfield Diffie and Martin Hellman introduced the Diffie-Hellman key exchange protocol [6] that allowed two parties to agree on a secret key over a public channel. The simplest version of the protocol starts with a prime number p and a generator g of the multiplicative group $\mathbb{F}_p^*$, both of which are publicly known. Then Alice and Bob do the following:

1. Alice picks a random integer $0 \leq a < p-1$, called her *private key*. Bob also picks a random integer $0 \leq b < p-1$ for his private key.

2. Alice computes g^a, known as her *public key*, and sends it over to Bob. Analogously, Bob computes g^b and sends it to Alice. This channel of communication is assumed to be public.

3. Upon receiving Bob's public key g^b, Alice computes $(g^b)^a$ using her private key. Similarly, Bob computes $(g^a)^b$. The element $g^{ab} = (g^b)^a = (g^a)^b$ is then their *shared secret key*.

The security of Diffie-Hellman relies on the fact that, given an element $h \in \mathbb{F}_p^*$, it's hard to find an integer x such that $g^x = h$. This problem is known as the *discrete logarithm problem*. Hence, the protocol's security breaks down if one is able to solve the discrete logarithm problem in a reasonable amount of time. It turns out that in the presence of a quantum computer this can be achieved using Schor's algorithm [15]. For that reason, the advent of quantum computers poses a significant risk to the currently used key exchange methods, and thus, to almost all modern secret communication. This has prompted cryptographers to search for alternative key exchange protocols that are resistant to attacks by quantum computers.

Because of their algebraic properties and practical computer implementation, a promising place to look for inspiration has been in the theory of elliptic curves. An elliptic curve over a field K is a type of projective algebraic curve that can generally be described with an equation of the form

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

where the coefficients a_i lie in K. A remarkable property of elliptic curves is that a group structure can be given to their set of points.

A fundamental object of study is the structure-preserving maps between elliptic curves. These maps are called isogenies, and they have found multiple applications in cryptography and cryptanalysis in recent years. In particular, their consideration has given rise to multiple isogeny-based cryptosystems that resemble the Diffie-Hellman key exchange. The first one was devised in 1997 by Couveignes [4] and later rediscovered by Rostovtsev and Stolbunov in 2004 [20]. Unfortunately, it has the flaw of being considerably slow.

In 2011, Jao and De Feo proposed a different protocol [9] called SIDH (Supersingular Isogeny Diffie-Hellman) which has attracted a lot of attention. It has, however, the disadvantage of requiring public keys of significant length. There are also some concerns regarding the potential leakage of information in the generation of the public keys. As an alternative, in 2018, Castryck, Lange, Martindale, Panny and Renes [1] came up with an adaptation –called CSIDH– of the Couveignes-Rostovtsev-Stolbunov key exchange that can be implemented efficiently and is believed to maintain security against quantum computers. This protocol is based on a commutative group action related to the endomorphism ring of a collection of elliptic curves.

The set of endomorphisms of an elliptic curve consists of the isogenies from the curve to itself. It has a ring structure and, under certain conditions, is an order $\mathcal{O}$ inside an imaginary quadratic number field. There is a finite commutative group associated with $\mathcal{O}$ called the class group.

In CSIDH, one considers the order $\mathcal{O} = \mathbb{Z}[\sqrt{-p}]$ with p a prime number of certain form. It turns out that the class group of $\mathcal{O}$ acts, via isogenies, on the set of elliptic curves over $\mathbb{F}_p$ that have $\mathcal{O}$ as endomorphism ring. This action has the notable property of being free and transitive, and it can be used to construct a protocol analogous to the Diffie-Hellman key exchange with the additional property of being quantum-resistant. Since this group action is commutative, the protocol was named CSIDH for Commutative Supersingular Isogeny Diffie-Hellman.

This thesis intends to be an exploration of the mathematics behind CSIDH. As such, we will mainly focus on the mathematical objects and results needed to explain, set up and implement this protocol, and not so much on the algorithmic optimality, security analysis, and possible attacks.

The first chapter is dedicated to developing the basic theory of elliptic curves. The main focus will be on isogenies, the structure-preserving maps. We will show some important properties of these maps and prove the existence of some of them under certain conditions.

In the second chapter we take a detour into algebraic number theory. We start in a more general setting by considering arbitrary number rings (subrings of number fields). We prove the Kummer-Dedekind Theorem about prime ideals in simple integral extensions of $\mathbb{Z}$. We then specialize to the case of orders inside imaginary quadratic number fields. Our study of these orders is relevant as the endomorphism ring of the elliptic curves we will consider is an order of this kind. This will be shown in Chapter 3, where supersingular elliptic curves are introduced.

Chapter 4 is entirely devoted to the group action that is the fundamental underpinning of CSIDH. As we previously mentioned, this action is free and transitive, a fact that was proven by Waterhouse [26] in a much more general setting using techniques that apply to general abelian varieties over finite fields. Since we only need the result for the very specific case that pertains to CSIDH, we will give alternative proofs that avoid the more sophisticated techniques. Hopefully, this will contribute to the accessibility of Waterhouse's result to readers that may not be familiar with the theory of abelian varieties.

Everything comes together in Chapter 5 where the CSIDH protocol is described at last. Important practical questions about the efficient computation of the class group action and the representation of keys are answered, and some possible attacks and security estimations are discussed.

1 Elliptic curves

Besides their relevant role in number theory, elliptic curves have become increasingly important in cryptography over the last few decades. Many algorithms have been developed based on their arithmetic, which can be efficiently implemented. In this chapter we will develop the essential theory of elliptic curves and the maps between them.

We assume throughout this thesis that K is a perfect field and $\bar{K}$ a fixed algebraic closure of K. We denote by $\mathrm{Gal}(\bar{K}/K)$ the Galois group of the field extension $\bar{K}/K$.

For the sections 1.1 and 1.2 we mainly follow Chapter 1 and 2 of [18].

1.1 Plane curves

In order to define elliptic curves, we start by giving the definition of a more general geometric object: plane curves. Let us begin with the affine case.

Definition 1.1. The *affine plane* is the set of pairs

$$\mathbb{A}^2 = \{P = (x, y) \mid x, y \in \bar{K}\}$$

Definition 1.2. An *affine plane curve defined over* K is a subset of $\mathbb{A}^2$ of the form

$$C = \{P \in \mathbb{A}^2 \mid F(P) = 0\},$$

where F is a polynomial in $K[x, y]$ that is irreducible in $\bar{K}[x, y]$.

To denote that the curve C is defined over K, we write C/K. In occasions we will refer to a curve simply by its defining equation $F(x, y) = 0$. Clearly, if a curve is given by $F(x, y) = 0$, then $\lambda F(x, y) = 0$, with $\lambda \in K^*$, determines the same curve.

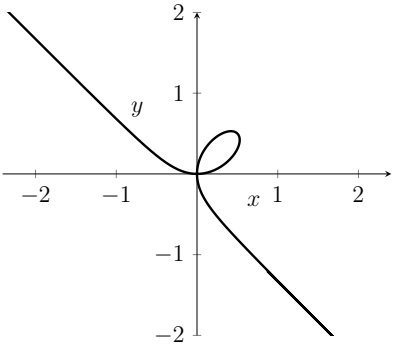
Definition 1.3. For an affine curve C/K and a field $L \supseteq K$, we define

$$C(L) := \{(x, y) \in L^2 \mid F(x, y) = 0\}.$$

We call $C(K)$ the set of *rational points* of C.

Example 1.4. Consider the curve $C/\mathbb{Q}$ given by the polynomial $F(x, y) = x^3 + y^3 - xy$. A picture of $C(\mathbb{R})$ is shown in Figure 1.1.

Figure 1.1: The real points of $x^3 + y^3 - xy = 0$



The points $(0, 0)$ and $(1/2, 1/2)$ are examples of rational points in C. In fact, $C(\mathbb{Q})$ consists of an infinite number of points.

If one attempts to draw the tangent line to the above curve at $(0, 0)$, one runs into trouble because the curve appears to self intersect at this point. We are interested in curves that look 'smooth' and where this doesn't happen. To formalize this notion, consider a curve $C : F(x, y) = 0$ and a point $P = (a, b) \in C$. The equation

$$\left(\frac{\partial F}{\partial x}(P)\right)(x - a) + \left(\frac{\partial F}{\partial y}(P)\right)(y - b) = 0$$

defines a line, called the *tangent line at P to C* (it is, in some sense, the best linear approximation to C at P), except if both $\frac{\partial F}{\partial x}(P)$ and $\frac{\partial F}{\partial y}(P)$ are equal to 0. This motivates the following definition.

Definition 1.5. Let $C : F(x, y) = 0$ be an affine plane curve and $P \in C$. Then C is *smooth at P* if

$$\frac{\partial F}{\partial x}(P) \neq 0 \quad \text{or} \quad \frac{\partial F}{\partial y}(P) \neq 0.$$

If C is smooth at every point P, we say that C *is smooth*. If C is not smooth at P, then we say that P is a *singular point* of C.

Example 1.6. Recall the curve from Example 1.4. The points (x, y) such that $\frac{\partial F}{\partial x}(x, y) = 3x^2 - y = 0$ and $\frac{\partial F}{\partial y}(x, y) = 3y^2 - x = 0$ are $(0, 0)$, $(1/3, 1/3)$, $(1/3w^2, 1/3w)$ and $(1/3w, 1/3w^2)$ with $w = e^{2\pi i/3}$. Note that only $(0, 0)$ is in the curve. Hence, the only singular point of C is $(0, 0)$.

An important set of objects is that of the polynomial functions defined at the points of $C : F(x, y) = 0$. If g and h are polynomials in $K[x, y]$ and $g - h \in (F)$, then $g(P) = h(P)$ for every $P \in C$. Therefore, we'd like to identify g with h. For this, we make the following definition.

Definition 1.7. Let C/K be an affine plane curve with defining equation $F(x, y) = 0$. The *ring of regular functions* of C is

$$K[C] = K[x, y]/(F).$$

Similarly, we define $\bar{K}[C] = \bar{K}[x, y]/(F)$.

Note that $K[C]$ and $\bar{K}[C]$ are integral domains since F is irreducible in $\bar{K}[x, y]$

Definition 1.8. Let C/K be an affine plane curve. The *function field* of C is the field of fractions of $K[C]$, denoted by $K(C)$. Similarly, $\bar{K}(C)$ is the field of fractions of $\bar{K}[C]$.

One of the advantages of working with smooth points is that for a smooth point P on a curve C and a function $f \in \bar{K}(C)$, we can talk about the *order of f at P* in an analogous way to the case of meromorphic functions in complex analysis. To do this, we first define the ring of functions defined at P.

Definition 1.9. Let P be a point in C. The *local ring of C at P* is

$$\bar{K}[C]_P := \{g/h \in \bar{K}(C) \mid g, h \in \bar{K}[C] \text{ with } h(P) \neq 0\}.$$

Note that $(g/h)(P) = g(P)/h(P)$ is well-defined if $h(P) \neq 0$. In consequence, the functions in $\bar{K}[C]_P$ are said to be *defined* at P.

The unique maximal ideal of $\bar{K}[C]_P$ is given by

$$M_P := \{f \in \bar{K}[C]_P \mid f(P) = 0\}.$$

since if f is not in M_P, then it is a unit in $\bar{K}[C]_P$. To see this, note that $f \in \bar{K}[C]_P \setminus M_P$ can be written as g/h with $g, h \in \bar{K}(C)$ and $h(P) \neq 0$. Since $f(P) \neq 0$, then $g(P) \neq 0$ so that $1/f = h/g$ is in $\bar{K}[C]_P$.

It can be shown that if P is a smooth point, then $\dim_{\bar{K}} M_P/M_P^2 = 1$, which in turn implies that $\bar{K}[C]_P$ is a *discrete valuation ring* [18, II.2 Proposition 1.1] (i.e. M_P is principal and every non-zero ideal of $\bar{K}[C]_P$ is a power of M_P).

Definition 1.10. Let C be an affine plane curve and $P \in C$ a smooth point. The *valuation on $\bar{K}[C]_P$* is given as follows: for $f \in \bar{K}[C]_P$ define

$$\mathrm{ord}_P(f) = \sup\{d \in \mathbb{Z}_{\geq 0} \mid f \in M_P^d\}.$$

We extend this to $\bar{K}(C)$ by defining $\mathrm{ord}_P(f/g)$ with $f, g \in \bar{K}[C] \subseteq \bar{K}[C]_P$ as

$$\mathrm{ord}_P(f/g) = \mathrm{ord}_P(f) - \mathrm{ord}_P(g).$$

Definition 1.11. Let $f \in \bar{K}(C)$ and consider a smooth point $P \in C$. We call $\mathrm{ord}_P(f)$ the *order of f at P*. If $\mathrm{ord}_P(f) > 0$ we say that *f has a zero at P*, and if $\mathrm{ord}_P(f) < 0$ we say that *f has a pole at P*.

Definition 1.12. Let $P \in C$ be a smooth point. Any function $t \in \bar{K}(C)$ with $\mathrm{ord}_P(t) = 1$ is called a *uniformizer for C at P* .

Remark 1.13. We have $\mathrm{ord}_P(t) = 1$ if and only if $t \in M_P \setminus M_P^2$. In that case, t is a generator for M_P. Hence, every function $f \in \bar{K}(C)$ can be written as

$$f = t^{\mathrm{ord}_P(f)} u$$

where $u \in \bar{K}[C]_P^*$.

Remark 1.14. If $t(x, y) = 0$ is a line that goes through P and is not the tangent line at P to C, then t is a uniformizer for C at P.

One can easily show that ord_P satisfies the following properties:

Proposition 1.15. For every $f, g \in \bar{K}(C)$ and $P \in C$ a smooth point, we have

- $\mathrm{ord}_P(fg) = \mathrm{ord}_P(f) + \mathrm{ord}_P(g)$.
- $\mathrm{ord}_P(f) \geq 0$ if and only if $f \in \bar{K}[C]_P$.
- $\mathrm{ord}_P(f) > 0$ if and only if $f \in \bar{K}[C]_P$ and $f(P) = 0$.

Note that if $\mathrm{ord}_P(f) < 0$, then f is not defined at P but $1/f$ is and has a zero at P.

For the description of elliptic curves, it is useful to add a 'point at infinity'. To do this, we introduce the projective plane and projective plane curves.

Definition 1.16. The *projective plane*, denoted by $\mathbb{P}^2$, is the set of all 3-tuples $(a, b, c) \in \bar{K}^3 \setminus \{(0, 0, 0)\}$ modulo the equivalence relation

$$(a_1, b_1, c_1) \sim (a_2, b_2, c_2) \iff \text{there is } \lambda \in \bar{K}^* \text{ such that } (a_1, b_1, c_1) = (\lambda a_2, \lambda b_2, \lambda c_2).$$

The equivalence class of (a, b, c) is denoted by $(a : b : c)$. The elements a, b, c are said to be projective coordinates for the point $(a : b : c)$.

Definition 1.17. For a field $L \subseteq \bar{K}$, the set of *L-rational points in* $\mathbb{P}^2$ is the set

$$\mathbb{P}^2(L) := \{(a : b : c) \in \mathbb{P}^2 \mid a, b, c \in L\}$$

We refer to $\mathbb{P}^2(K)$ simply as the *rational points* in $\mathbb{P}^2$.

In order to give the definition of a projective curve, we need to take into account that $(a : b : c) = (\lambda a : \lambda b : \lambda c)$ in $\mathbb{P}^2$ for $\lambda \in \bar{K}^*$. Note that if $F \in K[X, Y, Z]$ is a homogeneous polynomial of degree d, then

$$F(\lambda a, \lambda b, \lambda c) = \lambda^d F(a, b, c).$$

Therefore, for $P \in \mathbb{P}^2$, the statement $F(P) = 0$ is well-defined as it is independent of the representative chosen for P.

Definition 1.18. A *projective plane curve* defined over K is a subset of $\mathbb{P}^2$ of the form

$$C = \{P \in \mathbb{P}^2 \mid F(P) = 0\},$$

where $F \in K[X, Y, Z]$ is a homogeneous polynomial that is irreducible in $\bar{K}[X, Y, Z]$. For a field $K \subseteq L \subseteq \bar{K}$ we define the set of *L-rational points* of C by

$$C(L) = \{P \in \mathbb{P}^2(L) \mid F(P) = 0\}$$

We refer to $C(K)$ simply as the set of *rational points* of C.

A projective curve gives rise to an affine curve, which is often more convenient to work with. Consider the set

$$U := \{(a : b : c) \in \mathbb{P}^2 \mid c \neq 0\}.$$

If $(a : b : c)$ is in U, then we can divide each coordinates by c and get the alternative representation $(a/c : b/c : 1)$. We see that there is a well-defined bijection between U and $\mathbb{A}^2$ given by:

$$\begin{aligned} \varphi : U &\to \mathbb{A}^2 \\ (a : b : c) &\mapsto (a/c, b/c) \end{aligned}$$

where the inverse map is $(a, b) \mapsto (a : b : 1)$.

Note that if $F \in K[X, Y, Z]$ is a homogeneous polynomial of degree d and $P = (a : b : c)$ is a point in U such that $F(P) = 0$, then $F(a/c, b/c, 1) = 0$.

Let us define the polynomial $F^* \in K[x, y]$ as $F^*(x, y) = F(x, y, 1)$. Consider the projective curve C given by $F(X, Y, Z) = 0$, and the affine curve C' given by $F^*(x, y) = 0$. We have that if $(a : b : c)$ is in $C \cap U$, then $(a/c, b/c)$ is in C'. Conversely, if (a, b) is in C' then $(a : b : 1)$ is in C. Thus, φ restricts to a bijection

$$\begin{aligned} C \cap D &\to C' \\ (a : b : c) &\mapsto (a/c, b/c) \end{aligned}$$

We identify the affine curve C' with $C \cap U$ and call it an *affine patch* of C. The process of obtaining the polynomial F^* from F is called *dehomogenization with respect to Z* and is achieved by simply setting the variable $Z = 1$ and relabeling X to x and Y to y, or –alternatively– by dividing F by Z^d and defining the variables $x = X/Z$ and $y = Y/Z$.

Given the polynomial F^*, we can recover F by

$$F(X, Y, Z) = Z^d F^*(X/Z, Y/Z).$$

We say that F is the *homogenization of F^* with respect to Z of degree d.*

We could have just as well worked with the set $V = \{(a : b : c) \in \mathbb{P}^2 \mid a \neq 0\}$ or $W = \{(a : b : c) \in \mathbb{P}^2 \mid b \neq 0\}$, and dehomogenize F with respect to X (or Y) to get the polynomial corresponding to the affine curve $C \cap V$ (or $C \cap W$).

Any point $P \in C$ must lie in at least one of the affine patches $C \cap U$, $C \cap V$ or $C \cap W$. We call $C \cap U$ the *main affine patch* and the points $P \in C$ that are not in $C \cap U$ (i.e. with last coordinate zero) are called the *points at infinity of* C. It's often more convenient to work with affine coordinates, so, for a projective curve $F(X, Y, Z) = 0$ we will in occasions write the dehomogenization of F with respect to one of the variables (usually Z), instead of F.

In order to describe the function field of a projective curve, note that if $f, g \in K[X, Y, Z]$ are homogeneous polynomials of the same degree, then $f(P)/g(P)$ is well-defined for all points $P \in \mathbb{P}^2$ with $g(P) \neq 0$.

Definition 1.19. Let C/K be a projective plane curve with defining polynomial F. The *function field of* C, denoted by $K(C)$, is the field of rational functions f/g with $f, g \in K[X, Y, Z]$ such that

- f and g are homogeneous of the same degree.
- $g \notin (F)$.
- f_1/g_1 and f_2/g_2 are considered the same if $f_1 g_2 - f_2 g_1 \in (F)$.

The field $\bar{K}(C)$ is obtained by replacing K with $\bar{K}$ in the above definition.

Remark 1.20. The function field of a non-empty affine patch of C can be identified with $K(C)$. For example, if $C \cap U \neq \emptyset$, then we can identify the function field of the corresponding affine curve (in coordinates x, y) with $K(C)$ via $x = X/Z$ and $y = Y/Z$.

We can now use the definitions from the affine case to define the same notions in projective curves.

Definition 1.21. Let C be a projective curve and let $P \in C$. Choose an affine patch C' of C such that $P \in C'$. Then C is *smooth at* P if C' is smooth at P. For a function $f \in \bar{K}(C)$ we define the *order of* f *at* P, denoted by $\mathrm{ord}_P(f)$, as the order of f (seen as a function in $\bar{K}(C')$) at the point P. These can be shown to be independent of the choice of affine patch.

This is better illustrated with an example.

Example 1.22. Suppose $\mathrm{char}(K) \neq 2$ and let C/K be the smooth projective curve given by the equation

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

Note that if $Z = 0$, then $X = 0$, so the only point at infinity is $(0 : 1 : 0)$. We dehomogenize with respect to Z to get the equation of an affine curve C'/K

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

This curve is the main affine patch of C and we will often write projective curves in affine coordinates like this.

Let us consider the function $h = X/Y \in K(C)$ and the point $P = (0:0:1) \in C$. We will compute $\mathrm{ord}_P(h)$. Note that P is in $C \cap U$ and thus we can work on the main affine patch of C. Here, P corresponds to the affine point $(0,0)$ in C'. We can write h in affine coordinates recalling that $x = X/Z$ and $y = Y/Z$, hence $h = X/Y = (X/Z)/(Y/Z) = x/y$. The line $y = 0$ goes through $(0,0)$ and it's not tangent to C', therefore it is a uniformizer for C' at P. That is, $\mathrm{ord}_P(y) = 1$. Note that we can write

$$h = \frac{x}{y} = \frac{x(x^2+1)}{y(x^2+1)} = \frac{y^2}{y(1+x^2)} = \frac{y}{1+x^2}$$

The function $x^2 + 1$ is defined at P and its value when evaluated at P is not zero. It follows from Proposition 1.15 that $\mathrm{ord}_P(x^2+1) = 0$. Hence,

$$\mathrm{ord}_P(h) = \mathrm{ord}_P(y) - \mathrm{ord}_P(x^2+1) = 1.$$

Let us consider now the only point at infinity $O := (0:1:0)$ and the function $x = X/Z$. We want to compute $\mathrm{ord}_O(x)$. Since O is not in $C \cap U$ we need to consider a different affine patch. Let us work with $C \cap W$, so we dehomogenize with respect to Y. Define $u := X/Y$ and $v := Z/Y$. We get the affine curve defined by the equation

$$v = u^3 + uv^2$$

The function x written in these coordinates is $x = X/Z = (X/Y)/(Z/Y) = u/v$. The point O corresponds to $(0,0)$ in this affine patch. The line $u = 0$ goes through $(0,0)$ and is not tangent to the curve at this point. Hence, u is a uniformizer at O so $\mathrm{ord}_O(u) = 1$. Note that

$$x = \frac{u}{v} = \frac{u^3}{vu^2} = \frac{v(1-uv)}{vu^2} = \frac{1-uv}{u^2}$$

The function $1-uv$ evaluated at O is different from zero. It follows that $\mathrm{ord}_O(1-uv) = 0$. Hence,

$$\mathrm{ord}_O(x) = -\mathrm{ord}_O(u^2) = -2.$$

Additionally, notice that $v = u/x$ so $\mathrm{ord}_O(v) = 1 - (-2) = 3$. In consequence, for the function $y = Y/Z = 1/v$ we have

$$\mathrm{ord}_O(y) = -3.$$

We will now define maps between projective curves. Let C_1 and C_2 be projective curves over K. Observe that $\bar{K}(C_1)$ is a field extension of K, so that we may regard C_2 as a curve defined over $\bar{K}(C_1)$. Hence, we can consider the points $C_2(\bar{K}(C_1))$ consisting of triples $(f_0 : f_1 : f_2)$ with $f_0, f_1, f_2 \in \bar{K}(C_1)$ that satisfy the equation for C_2. We have the following definition.

Definition 1.23. Let C_1 and C_2 be projective curves defined over K. A *rational map* $\phi : C_1 \to C_2$ over $\bar{K}$ is a point $\phi \in C_2(\bar{K}(C_1))$. We may write

$$\phi = (f_0 : f_1 : f_2)$$

where $f_0, f_1, f_2 \in \bar{K}(C_1)$. The map ϕ is *defined over* K if f_1, f_2, f_3 can be taken to be in $K(C_1)$. Equivalently, if $\phi \in C_2(K(C_1))$.

Observe that if $(f_0 : f_1 : f_2) \in C_2(\bar{K}(C_1))$, then for every point $P \in C_1$ where the functions f_0, f_1, f_2 are defined and not all zero, we have

$$\phi(P) := (f_0(P) : f_1(P) : f_2(P)) \in C_2$$

Remark 1.24. Let $\sigma \in \mathrm{Gal}(\bar{K}/K)$ and let $\phi : C_1 \to C_2$ be a rational map of projective curves. We can let σ act on ϕ by applying σ to the coefficients of the rational expressions defining ϕ. Clearly, if ϕ is defined over K then $\sigma(\phi) = \phi$. The converse is also true: if $\sigma(\phi) = \phi$ for every $\sigma \in \mathrm{Gal}(\bar{K}/K)$ then ϕ is defined over K (See [18, pag. 16] and [7, Remark 5.4.14]).

Definition 1.25. A rational map $\phi = (f_0 : f_1 : f_2) : C_1 \to C_2$ is *defined at* P if there is $g \in \bar{K}(C_1)$ such that gf_0, gf_1, gf_2 are defined at P and not all zero at P. In such a case, we set

$$\phi(P) := ((gf_0)(P) : (gf_1)(P) : (gf_2)(P))$$

A rational map that is defined at every point $P \in C_1$ is called a *morphism*.

It turns out that if C_1 is smooth at every point, then a rational map $\phi : C_1 \to C_2$ is defined everywhere, so it is a morphism.

Proposition 1.26. Let C_1 and C_2 be projective curves and suppose every $P \in C_1$ is smooth. Let $\phi = (f_0 : f_1 : f_2) : C_1 \to C_2$ be a rational map. Then ϕ is a morphism.

Proof. Consider $P \in C_1$. Let $t \in \bar{K}(C)$ be a uniformizer for C at P. Let

$$n := \min_{0 \le i \le 2} \mathrm{ord}_P(f_i)$$

Then $\mathrm{ord}_P(t^{-n} f_i) \ge 0$ for every i, so $t^{-n}f_0, t^{-n}f_1, t^{-n}f_2$ are all defined at P, and there is $j \in \{0, 1, 2\}$ such that $\mathrm{ord}_P(t^{-n}f_j) = 0$ so $t^{-n}f_j$ is not zero at P. It follows that ϕ is defined at P. □

Example 1.27. Let $\mathrm{char}(K) \neq 2$ and let C_1 be the smooth curve from Example 1.22 given by the equation

$$Y^2Z = X^3 + XZ^2$$

Let C_2 be the smooth curve given by

$$Y^2Z = X^3 - 4XZ^2$$

Then $\phi : C_1 \to C_2$ with

$$\phi = \left(\frac{X^2 + Z^2}{ZX} : \frac{Y(X^2 - Z^2)}{ZX^2} : 1 \right)$$

is a morphism between these two curves.

Note that we can directly evaluate $\phi(P)$ at all points $P \in C_1$ except at $(0 : 0 : 1)$ and $(0 : 1 : 0)$. For these two points, $\phi(P)$ is a point at infinity. We must have that $\phi(0 : 1 : 0) = (0 : 1 : 0)$ and $\phi(0 : 0 : 1) = (0 : 1 : 0)$.

Using the identification $x = X/Z$ and $y = Y/Z$, we can write ϕ in the main affine coordinates as

$$\phi = \left(\frac{x^2 + 1}{x} : y\frac{x^2 - 1}{x^2} : 1 \right)$$

For convenience, we can simply write $\phi(x, y) = \left(\frac{x^2+1}{x}, y\frac{x^2-1}{x^2} \right)$ to refer to this morphism. Another alternative notation we may use is $(x, y) \mapsto \left(\frac{x^2+1}{x}, y\frac{x^2-1}{x^2} \right)$.

Remark 1.28. We will generally denote morphisms of the form $\phi = (f(x, y) : g(x, y) : 1)$ simply by $(f(x, y), g(x, y))$. Note that if $\phi = (f_1 : f_2 : f_3)$ with $f_3 \neq 0$, then ϕ can be written in this form as $(f_1/f_3, f_2/f_3)$.

Definition 1.29. Let C be a projective curve and assume without loss of generality (relabeling the variables if necessary) that it's not given by $Z = 0$. We define the *identity* morphism $id_C : C \to C$ as

$$id_C := (x : y : 1) = (x, y)$$

this map is such that $id_C(P) = P$ for every $P \in C$.

Definition 1.30. A morphism $\phi : C_1 \to C_2$ over K is an *isomorphism over* K if there exists a morphism $\psi : C_2 \to C_1$ over K such that $\psi \circ \phi = id_{C_1}$ and $\phi \circ \psi = id_{C_2}$.

Proposition 1.31. Let $\phi : C_1 \to C_2$ be a morphism of curves, then ϕ is either constant or surjective.

Proof. We refer to [14, 5.2 Theorem 1.10] for details. This proposition follows from the fact that the image of a projective curve (recall that our curves are irreducible by definition) under a morphism $\phi : C_1 \to C_2$ between projective curves is an irreducible projective variety. That is, $\phi(C_1)$ is either a single point or a curve inside C_2. But C_2 doesn't properly contain any curves, hence $\phi(C_1)$ either contains a single point or it's equal to C_2. □

Proposition 1.32. Let C/K be a smooth curve and $f \in \bar{K}(C)^*$. Then f has no poles if and only if $f \in \bar{K}^*$.

Proof. Let us denote by $\mathbb{P}^1$ the projective curve defined by $Y = 0$. Let $f \in \bar{K}(C)^*$ and define the following morphism $\phi : C \to \mathbb{P}^1$ as

$$\phi := (f : 0 : 1)$$

Note that if f is defined at P, then $\phi(P) = (f(P) : 0 : 1)$. On the other hand, if f has a pole at P then $\phi(P) = (1 : 0 : 0)$.

Suppose that f has no poles. Then ϕ is not surjective, so by Proposition 1.31 it must be that ϕ is constant. It follows that there is $a \in \bar{K}^*$ such that $(f(P) : 0 : 1) = (a : 0 : 1)$ for every $P \in C$. We conclude that $f = a \in \bar{K}^*$.

On the other hand, if f is constant and non-zero, then it has no poles.

□

Definition 1.33. Let C_1 and C_2 be projective curves defined over K and let $\phi : C_1 \to C_2$ be a non-constant morphism over K. The map

$$\phi^* : K(C_2) \to K(C_1)$$

given by

$$\phi^* f = f \circ \phi$$

is called *the pullback*.

Note that the pullback fixes K and is a ring homomorphism, hence it's also injective. Therefore, every non-constant morphism induces an injection between function fields that fixes K. The converse is also true: every injection of function fields that fixes K comes from a morphism.

Proposition 1.34. Suppose C_1 and C_2 are projective curves over K and let $\iota : K(C_2) \to K(C_1)$ be an injective ring homomorphism fixing K. Then there exists a unique non-constant morphism $\phi : C_1 \to C_2$ defined over K such that $\phi^* = \iota$.

Proof. Assume without loss of generality that C_2 is not contained in the line $Z = 0$, therefore $C_2 \cap U \neq \emptyset$ and we can work in the main affine patch with $x = X/Z$ and $y = Y/Z$ in $K(C_2)$. Suppose C_2 is defined by the polynomial $F \in K[x, y]$. Then, $F(x, y) = 0$ in $K(C_2)$ and since ι fixes K and is a ring homomorphism, we have

$$F(\iota(x), \iota(y)) = \iota(F(x, y)) = \iota(0) = 0$$

in $K(C_1)$.

It follows that $\phi = (\iota(x), \iota(y))$ is a morphism from C_1 to C_2. Note that x and y cannot both be constant, and thus, since ι is injective and fixed K, at least one of $\iota(x)$ and $\iota(y)$ is non-constant. Hence, ϕ is non-constant. In addition, since ι fixes K, we have that for $f \in K(C_2)$ it's true that $f \circ \phi = f(\iota(x), \iota(y)) = \iota(f(x, y)) = \iota(f)$. Therefore, $\phi^* = \iota$.

To show uniqueness, let $\psi := (f_1, f_2)$ be another morphism from C_1 to C_2 with $\psi^* = \iota$. Then $\iota(x) = \psi^*(x) = x \circ \psi = f_1$ and similarly $\iota(y) = f_2$. We conclude that $\psi = \phi$.

□

1.2 Elliptic Curves

Elliptic curves are a very special kind of projective curves. Their importance comes from the fact that there is a way to define a group law on their set of points. Although a more general definition can be given, the following one will be sufficient for our purposes. From now on, we assume that $\text{char}(K) \neq 2$.

Definition 1.35. Let K be a field with $\text{char}(K) \neq 2$. An *elliptic curve over* K is a smooth projective plane curve over K given by an equation of the form

$$Y^2Z = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3$$

with $a_2, a_4, a_6 \in K$.

The *base point* is $O = (0 : 1 : 0)$, the only point at infinity.

Remark 1.36. We can write the equation above in affine coordinates as

$$y^2 = x^3 + a_2x^2 + a_4x + a_6.$$

An equation of this form is called a *Weierstrass equation*.

In order for a curve given by a Weierstrass equation to be an elliptic curve, it has to be smooth. The following proposition gives a sufficient and necessary condition.

Proposition 1.37. A projective curve given by the equation $y^2 = x^3 + a_2x^2 + a_4x + a_6$ with $a_2, a_4, a_6 \in K$ is smooth if and only if

$$18a_2a_4a_6 - 4a_2^3a_6 + a_2^2a_4^2 - 4a_4^3 - 27a_6^2 \neq 0$$

Proof. It's easy to verify that $O = (0 : 1 : 0)$ is always a smooth point in this curve, and that an affine curve with equation $y^2 = f(x)$, with f a polynomial, is smooth if and only if f has no multiple roots. This is equivalent to the discriminant $\Delta(f)$ being non-zero. For $f(x) = x^3 + a_2x^2 + a_4x + a_6$, the expression above is $\Delta(f) \neq 0$ (see [8, 10.3] for the discriminant of a cubic polynomial). □

As we previously mentioned, there is a way to define a composition law between points on an elliptic curve in such a way that it gives them an abelian group structure. We do this via the following formulas. Note that all points of an elliptic curve different from O can be written in the form $(a : b : 1)$ in a unique way. Since we prefer to work with affine coordinates, we write (a, b) for $(a : b : 1)$.

The group law. Let E be an elliptic curve over K. Let $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$ be points of E. We define $P_1 + P_2$ as follows.

1. If $x_1 = x_2$ and $y_1 + y_2 = 0$, then

$$P_1 + P_2 = O$$

2. If $x_1 = x_2$ and $y_1 + y_2 \neq 0$ (i.e. if $P_1 = P_2$ and $y_1 \neq 0$), then $P_1 + P_2 = (x_3, y_3)$ with

$$x_3 = \lambda^2 - a_2 - 2x_1, \;\; y_3 = -\lambda(x_3 - x_1) - y_1, \text{ where } \lambda := \frac{3x_1^2 + 2a_2x_1 + a_4}{2y_1}$$

3. If $x_1 \neq x_2$, then $P_1 + P_2 = (x_3, y_3)$ with

$$x_3 = \lambda^2 - a_2 - x_1 - x_2, \;\; y_3 = -\lambda(x_3 - x_1) - y_1, \text{ where } \lambda := \frac{y_2 - y_1}{x_2 - x_1}$$

Finally, define $P + O = P$ and $O + P = P$ for all $P \in E$.

Remark 1.38. These formulas have a geometric interpretation: given the points P_1 and P_2 consider the line that goes through P_1 and P_2 (if $P_1 = P_2$ then take the tangent line at P_1 to E). It can be shown that this line intersects E at a third point, say Q. Reflecting Q along the x-axis gives P_3.

Proposition 1.39. Let E be an elliptic curve over K. The group law described above makes $(E, +)$ into an abelian group.

Proof. The commutativity of $+$ can be seen directly from the formulas. We also see that O is the identity element and that if $P = (x, y)$ then $-P = (x, -y)$. The associativity of $+$ can also be proven by tedious calculations with the explicit formulas and a case by case analysis. (for other approaches see [25, Section 2.4] or [18, III Proposition 3.4]). □

Remark 1.40. We see from the formulas that if $P \in E(K)$, then $-P \in E(K)$. Also, if $P_1, P_2 \in E(K)$, then $P_1 + P_2 \in E(K)$. Therefore, $E(K)$ is a subgroup of E.

The following result regarding the order of x and y at O will be useful later.

Proposition 1.41. Let E be an elliptic curve given by the equation $y^2 = x^3 + a_2x^2 + a_4x + a_6$, with O the point at infinity. Then,

$$\mathrm{ord}_O(x) = -2 \text{ and } \mathrm{ord}_O(y) = -3$$

Proof. Just as in Example 1.22, one can show that x/y is a uniformizer for E at O. We also see that $1/y = Z/Y$ so $1/y$ is defined at O and, in fact, is zero at O. We write

$$x = \frac{x/y}{1/y} = \frac{(x/y)^3}{(1/y)(x/y)^2} = \frac{1 - a_2(x/y)^2 - a_4(x/y)(1/y) - a_6(1/y)^2}{(x/y)^2}$$

Since $1 - a_2(x/y)^2 - a_4(x/y)(1/y) - a_6(1/y)^2$ is non-zero when evaluated at O, we get

$$\mathrm{ord}_O(x) = -2\,\mathrm{ord}_O(x/y) = -2,$$

and

$$\mathrm{ord}_O(y) = \mathrm{ord}_O(x/(x/y)) = -2 - 1 = -3.$$

□

Definition 1.42. Let E be an elliptic curve over K. Consider an integer $n \geq 1$. We define the set $\mathcal{L}(nO)$ as

$$\mathcal{L}(nO) = \{f \in \bar{K}(E) \mid \mathrm{ord}_P(f) \geq 0 \text{ for all } P \in E \setminus \{O\} \text{ and } \mathrm{ord}_O(f) \geq -n\} \cup \{0\}.$$

Example 1.43. From Proposition 1.41 we see that

$$x \in \mathcal{L}(2O) \setminus \mathcal{L}(1O) \quad \text{and} \quad y \in \mathcal{L}(3O) \setminus \mathcal{L}(2O)$$

For every $f, g \in \bar{K}(E)$ and every $P \in E$ we have $\mathrm{ord}_P(f+g) \geq \min\{\mathrm{ord}_P(f), \mathrm{ord}_P(g)\}$, and for every $\lambda \in \bar{K}^*$ we have $\mathrm{ord}_P(\lambda f) = \mathrm{ord}_P(f)$, so $\mathcal{L}(nO)$ is a $\bar{K}$-vector space.

The following proposition gives us the exact dimension of this space.

Proposition 1.44. Let E be an elliptic curve over K and n a positive integer, then

$$\dim_{\bar{K}} \mathcal{L}(nO) = n.$$

Proof. This is a direct consequence of the Riemann-Roch theorem (see [18, II Theorem 5.4]). □

Corollary 1.45. Let E_1 be an elliptic curve over K with equation $y^2 = x^3 + a_2x^2 + a_4x + a_6$ and E_2 an elliptic curve over K with equation $y_1^2 = x_1^3 + b_2x_1^2 + b_4x_1 + a_6$. If $\phi : E_1 \to E_2$ with $(x, y) \mapsto (x', y')$ is an isomorphism over K with $\phi(O) = O$, then there are constants $u \in K^*$ and $r, s, t \in K$ such that

$$x' = u^2x + r, \quad y' = u^3y + su^2x + t$$

Proof. By Proposition 1.41, the coordinate functions $x, y \in K(E_1)$ satisfy $\mathrm{ord}_O(x) = -2$ and $\mathrm{ord}_O(y) = -3$. Similarly, the functions $x_1, y_1 \in K(E_2)$ satisfy $\mathrm{ord}_O(x_1) = -2$ and $\mathrm{ord}_O(y_1) = -3$. Since ϕ is an isomorphism, then the pullback $\phi^* : K(E_2) \to K(E_1)$ is a ring isomorphism. It's not hard to show that a uniformizer for E_2 at O maps to a uniformizer for E_1 at O, which implies that $\mathrm{ord}_O(\phi^* f) = \mathrm{ord}_O(f)$ for all $f \in K(E_2)$. Note that $\phi^* x_1 = x'$ and $\phi^* y_1 = y'$, so that $\mathrm{ord}_O(x') = \mathrm{ord}_O(x_1) = -2$, and similarly $\mathrm{ord}_O(y') = \mathrm{ord}_O(y_1) = -3$. By the previous proposition, the $\bar{K}$-vector space $\mathcal{L}(2O) \subseteq \bar{K}(E_1)$ has dimension 2, so $\{1, x\}$ and $\{1, x'\}$ are both bases of $\mathcal{L}(2O)$. Hence, there are $u_1 \in \bar{K}^*$ and $r \in \bar{K}$ such that

$$x' = u_1x + r.$$

In fact, since x' is defined over K, we have that $u_1, r \in K$. Similarly, $\{1, x, y\}$ and $\{1, x', y'\}$ are bases for $\mathcal{L}(3O)$ so that there is $u_2 \in K^*$ and $s_2, t \in K$ such that

$$y' = u_2y + s_2x + t.$$

Since (x', y') satisfies the equation for E_2 and (x, y) satisfies the equation for E_1 where x^3 and y^2 both have coefficient 1, we must have that $u_2^2 = u_1^3$. Take $u = u_2/u_1$ and $s = s_2/u^2$. □

1.3 Isogenies

As the name suggests, one of the major players in isogeny-based cryptography are the structure-preserving maps between elliptic curves, called isogenies. For the rest of this chapter, we mostly follow [25, Section 2.9 and Chapter 12] and [21, Lectures 5, 6 and 7].

Definition 1.46. Let E_1 and E_2 be elliptic curves over K. An *isogeny* $\phi : E_1 \to E_2$ is a non-constant morphism between elliptic curves such that that $P \mapsto \phi(P)$ is a group homomorphism from $E_1(\bar{K})$ to $E_2(\bar{K})$.

Remark 1.47. It can be shown that any morphism ϕ between elliptic curves such that $\phi(O) = O$ is a group homomorphism [18, III.4 Theorem 4.8], so our definition is stronger than what is actually necessary.

Remark 1.48. From Proposition 1.31 we get that an isogeny is surjective.

In order to prove some of the properties of isogenies, it will be useful to have a standard way of writing them.

Let E_1 be an elliptic curve over K with equation $y^2 = x^3 + a_2x^2 + a_4x + a_6$ and let $\phi : E_1 \to E_2$ be an isogeny over K. Then there are rational functions $R_1(x, y)$ and $R_2(x, y)$ with coefficients in K such that

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

We can replace every appearance of y^2 in $R_1(x, y)$ by $x^3 + a_2x^2 + a_4x + a_6$, so that y has degree at most 1 in $R_1(x, y)$. Hence,

$$R_1(x, y) = \frac{p_1(x) + p_2(x)y}{p_3(x) + p_4(x)y}$$

with $p_i(x)$ polynomials.

Moreover, we can multiply the numerator and denominator of the above expression by $p_3(x) - p_4(x)y$ and replace y^2 by a polynomial in x as before, so that

$$R_1(x, y) = \frac{q_1(x) + q_2(x)y}{q_3(x)}$$

with $q_i(x)$ polynomials. Analogously, we can write

$$R_2(x, y) = \frac{s_1(x) + s_2(x)y}{s_3(x)}$$

with $s_i(x)$ polynomials.

Since ϕ is a group homomorphism, we have that $\phi(-P) = -P$. This implies that $R_1(x, -y) = R_1(x, y)$ and $R_2(x, -y) = -R_2(x, y)$. From the condition on $R_1(x, y)$ it follows that $q_2(x)$ is the zero polynomial. Similarly, from the condition on $R_2(x, y)$ it follows that $s_1(x)$ is the zero polynomial. Hence, we may write

$$\phi(x, y) = (r_1(x), r_2(x)y) \tag{1.1}$$

with $r_1(x)$ and $r_2(x)$ rational functions with coefficients in K. An isogeny written in the form (1.1) is said to be in *standard form*.

Remark 1.49. Write $r_1(x) = p(x)/q(x)$ with $p(x)$ and $q(x)$ coprime (equivalently, with no common roots in $\bar{K}$). Let $P = (x, y) \in E$. It's easy to show that $\phi(P) = O$ if and only if $q(x) = 0$.

Definition 1.50. Let E and E' be elliptic curves over K. Then $\mathrm{Hom}(E, E')$ is the set of all isogenies defined over K from E to E', together with the zero morphism $P \mapsto O$. Equivalently, $\mathrm{Hom}(E, E')$ is the set of all morphisms of curves defined over K from E to E' that are also group homomorphisms. For the set of morphisms defined over $\bar{K}$ that are also group homomorphisms we write $\mathrm{Hom}(E_{\bar{K}}, {E'}_{\bar{K}})$.

Elements of $\mathrm{Hom}(E, E')$ are points in $E'(K(E))$, thus, given two elements $\phi, \psi \in \mathrm{Hom}(E, E')$ we can add them and get another morphism $\phi + \psi$ from E to E'. It can be shown using the explicit addition formulas (see [18, Remark 3.6.1]) that for every $P \in E(\bar{K})$ we have that $(\phi + \psi)(P) = \phi(P) + \psi(P)$. Therefore, $\phi + \psi$ is also an element in $\mathrm{Hom}(E, E')$. Similarly, we have that $(-\phi)(P) = -\phi(P)$ for every $P \in E(\bar{K})$. This shows that $\mathrm{Hom}(E, E')$ is a group under addition.

1.3.1 Degree and separability

In this section we assume that all elliptic curves are defined over K, and that, unless otherwise stated, the field of definition of the isogenies is K. There is an important quantity associated to any isogeny, called the degree.

Definition 1.51. Let $\phi : E_1 \to E_2$ be an isogeny in standard form $\phi(x, y) = (r_1(x), r_2(x)y)$. Write

$$r_1(x) = p(x)/q(x)$$

with $p(x)$ and $q(x)$ coprime. We define the *degree* of ϕ as

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

The above definition is convenient from a computational perspective, but an alternative definition can also be given. It can be shown that an equivalent way of describing the degree of $\phi : E_1 \to E_2$ is as the degree of the field extension $K(E_1)/\phi^*(K(E_2))$. That is, $\deg(\phi) = [K(E_1) : \phi^*(K(E_2))]$.

Proposition 1.52. Let $\phi : E_1 \to E_2$ be an isogeny over K. Then

$$\deg(\phi) = [K(E_1) : \phi^*(K(E_2))]$$

Proof. Let $x_1, y_1 \in K(E_1)$ be the affine coordinate functions of E_1. Note that $K(E_1)/K(x_1)$ is a field extension and, since $y_1^2 = x_1^3 + a_2x_1^2 + a_4x_1 + a_6$ and $x_1^3 + a_2x_1^2 + a_4x_1 + a_6$ is not a square in $K(x_1)$, then $[K(E_1) : K(x_1)] = 2$. Let $x_2, y_2 \in K(E_2)$ be the affine coordinate functions of E_2. Similarly, $[K(E_2) : K(x_2)] = 2$. Since ϕ^* is an injection fixing K, we also have $[\phi^*K(E_2) : \phi^*K(x_2)] = 2$.

Write $\phi = (r_1(x), r_2(x)y)$ in standard form, with $r_1(x) = p(x)/q(x)$ where $p(x)$ and $q(x)$ are coprime. Note that $\phi^* x_2 = r_1(x_1)$. Consider the extension $K(x_1)/K(r_1(x_1))$. We claim that the minimal polynomial of x_1 is

$$p(T) - r_1(x_1)q(T) \in K(r_1(x_1))[T]. \tag{1.2}$$

Clearly, this polynomial has x_1 as a root. To prove that it's irreducible, look at the polynomial as an element of $K[r_1(x)][T]$. It's easy to show that it is irreducible in this ring since it's linear in $r_1(x)$, and $p(T)$ and $q(T)$ have no common factors. Then, by Gauss's Lemma, it's irreducible in $K(r_1(x))[T]$. Thus, the polynomial is indeed the minimal polynomial of x_1.

Note that $\deg(p(T) - r_1(x_1)q(T)) = \max\{\deg p(T), \deg q(T)\} = \deg(\phi)$, therefore

$$[K(x_1) : K(r_1(x))] = \deg(\phi).$$

We have

$$\begin{aligned} 2[K(E_1) : \phi^* K(E_2)] &= [K(E_1) : \phi^* K(E_2)][\phi^* K(E_2) : \phi^* K(x_2)] \\ &= [K(E_1) : K(x_1)][K(x_1) : \phi^* K(x_2)] \\ &= [K(E_1) : K(x_1)][K(x_1) : K(r_1(x_1))] \\ &= 2\deg(\phi). \end{aligned}$$

and the result follows. □

Corollary 1.53. Let $\phi : E_1 \to E_2$ and $\psi : E_2 \to E_3$ be isogenies. Then,

$$\deg(\psi \circ \phi) = \deg(\psi)\deg(\phi)$$

Proof. This is straightforward from the characterization of the degree in terms of field extensions. □

Definition 1.54. Let $\phi : E_1 \to E_2$ be an isogeny in standard form $\phi(x, y) = (r_1(x), r_2(x)y)$. We say that ϕ is *separable* if the formal derivative $r_1'(x)$ is not identically zero. Otherwise, if $r'(x) = 0$, we say that ϕ is *inseparable*.

Remark 1.55. Since an isogeny is surjective, we have that $r_1(x)$ is not constant. Hence, if $\mathrm{char}(K) = 0$, then every isogeny is separable.

Remark 1.56. A characterization of the separability of an isogeny is also possible in terms of the separability of the field extension $K(E_1)/\phi^*(K(E_2))$.

Proposition 1.57. Let $r(x) = p(x)/q(x)$ be a rational function with $p(x)$ and $q(x)$ coprime polynomials with coefficients in a field K of positive characteristic p. Then

$$r'(x) = 0 \iff p'(x) = q'(x) = 0 \iff p(x) = f(x^p) \text{ and } q(x) = g(x^p)$$

where f and g are polynomials.

Proof. We have that $r_1'(x) = 0$ if and only if $p'(x)q(x) = p(x)q'(x)$. If $p(x)$ is constant, then $p'(x) = 0$ and it follows that $q'(x) = 0$. Assume $p(x)$ is not constant, and let $\alpha \in \bar{K}$ be a root of $p(x)$. Since $p(x)$ and $q(x)$ have no roots in common, then α is a root of $p'(x)$ of at least the same multiplicity. But $\deg p'(x) < \deg p(x)$, so this is only possible if $p'(x) = 0$. Similarly, $q'(x) = 0$. The converse is clear.

To prove the second equivalence, suppose $p'(x) = 0$. Write $p(x) = \sum_n a_n x^n$. Then, $p'(x) = \sum_n n a_n x^{n-1} = 0$ if and only if $na_n = 0$ for every n. It follows that if $a_n \neq 0$ then $n = 0$ in K and thus n must be divisible by p. Thus $p(x) = \sum_m a_{pm}(x^p)^m = f(x^p)$ with $f(x) = \sum_m a_{pm}x^m$. Similarly, if $q'(x) = 0$ then $q(x) = g(x^p)$ for some polynomial g. The implication in the opposite direction is straightforward. □

Proposition 1.58. Let $\phi : E_1 \to E_2$ be an inseparable isogeny over K, where E_1 and E_2 are defined over a field K of positive characteristic p and given by the equations

$$E_1 : y^2 = x^3 + a_2x^2 + a_4x + a_6 \text{ and } E_2 : y^2 = x^3 + b_2x^2 + b_4x + b_6$$

then ϕ can be written as

$$\phi(x, y) = (r_1(x^p), r_2(x^p)y^p)$$

with $r_1(x)$ and $r_2(x)$ rational functions with coefficients in K.

Proof. Let $\phi(x, y) = (\frac{u(x)}{v(x)}, \frac{s(x)}{t(x)}y)$ be in standard form, with $u(x)$ coprime with $v(x)$ and $s(x)$ coprime with $t(x)$. By assumption, ϕ is inseparable. From the previous proposition, since $(u(x)/v(x))' = 0$, it follows that $u(x)/v(x) = r_1(x^p)$ for some rational function $r_1(x)$. Since ϕ satisfies the equation for E_2, we have that

$$\left(\frac{s}{t}y\right)^2 = \left(\frac{u}{v}\right)^3 + b_2\left(\frac{u}{v}\right)^2 + b_4\frac{u}{v} + b_6$$

It follows that

$$v^3s^2y^2 = t^2(u^3 + b_2u^2v + b_4uv^2 + b_6v^3).$$

Let $f = x^3 + a_2x^2 + a_4x + a_6$ and $w = u^3 + b_2u^2v + b_4uv^2 + b_6v^3$. Substituting $y^2 = f$ we have

$$v^3s^2f = t^2w$$

Since ϕ is inseparable, we have that $u' = v' = 0$ and it follows that $w' = 0$. Hence, $(\frac{s^2f}{t^2})' = (\frac{w}{v^3})' = 0$. Applying Proposition 1.57 again, we get that $(s^2f)' = 0$ and $t' = 0$, hence there are polynomials g and h in $K[x]$ such that $s^2(x)f(x) = g(x^p)$ and $t(x) = h(x^p)$. Since E_1 is smooth, then every root of f in $\bar{K}$ is distinct. Note that every root of $f(x)$ is a root of $g(x^p)$ and that every root of $g(x^p)$ has multiplicity divisible by p. Therefore, we may write $g(x^p) = l(x^p)f(x)^p$ with l some polynomial. Thus, $s^2(x)f(x) = l(x^p)f(x)^p$. From this, using the fact that p is odd, we see that l

must be a square. We have $s^2(x)f(x) = (m(x^p))^2 f(x)^p$ for some polynomial m. We get the following equality

$$s^2(x)y^2 = s^2(x)f(x) = (m(x^p))^2 f(x)^p = (m(x^p))^2 y^{2p}$$

in $K(E_1)$. Therefore,

$$\left(\frac{s(x)}{t(x)}y\right)^2 = \left(\frac{m(x^p)}{h(x^p)}y^p\right)^2 = (r(x^p)y^p)^2$$

with $r(x) = m(x)/h(x)$. Thus, we have that $\frac{s(x)}{t(x)}y$ is equal to either $r(x^p)y^p$ or $-r(x^p)y^p$. Take $r_2(x)$ to be $r(x)$ or $-r(x)$, respectively. Recall that in the beginning we showed $u(x)/v(x) = r_1(x^p)$ for some rational function $r_1(x)$, so this concludes the proof. □

Definition 1.59. Let K be a field of positive characteristic p and let E be an elliptic curve over K with equation $y^2 = x^3 + a_2x^2 + a_4x + a_6$. Consider an integer $n \geq 0$ and let $q = p^n$. We will denote by $E^{(q)}$ the elliptic curve with equation

$$y^2 = x^3 + a_2^q x^2 + a_4^q x + a_6^q.$$

Definition 1.60. The p-power *Frobenius morphism* $\pi : E \to E^{(p)}$ is given by

$$\pi(x, y) = (x^p, y^p)$$

Sometimes we will simply refer to π as the Frobenius.

Remark 1.61. Using the addition formulas, one can verify that the Frobenius morphism induces a group homomorphism and thus π is an isogeny. In particular, $\pi(O) = O$.

Remark 1.62. The Frobenius morphism will be of particular interest to us when $K = \mathbb{F}_p$. Note that if $x \in \mathbb{F}_p$, then $x^p = x$. Therefore, when $K = \mathbb{F}_p$, we have that $E^{(p)} = E$ and $\pi : E \to E$ is an endomorphism, an element of $\mathrm{Hom}(E, E)$.

Remark 1.63. Recall that for $n \geq 1$ and q a prime power, the Galois group $\mathrm{Gal}(\mathbb{F}_{q^n}/\mathbb{F}_q)$ is cyclic, generated by the automorphism $\sigma : x \mapsto x^q$. An element $x \in \mathbb{F}_{q^n}$ is in $\mathbb{F}_q$ if and only if x is fixed by σ. Since this is true for any n, we have that $x \in \overline{\mathbb{F}}_q$ is in $\mathbb{F}_q$ if and only if x is fixed by σ. For the particular case $q = p$, it follows that a point $P \in E$ is in $E(\mathbb{F}_p)$ if and only if $\pi(P) = P$.

Remark 1.64. The Frobenius π has degree p and is inseparable since $(x^p)' = px^{p-1} = 0$.

It turns out that we can decompose an isogeny into a separable and an inseparable part.

Proposition 1.65. Let E_1 and E_2 be elliptic curves over a field K of positive characteristic p. If $\phi : E_1 \to E_2$ is an isogeny over K, then we can write

$$\phi = \phi_{sep} \circ \pi^n$$

with π the Frobenius morphism, $n \in \mathbb{Z}_{\geq 0}$, and $\phi_{sep} : E_1^{(p^n)} \to E_2$ a separable isogeny over K. Furthermore, $\deg(\phi) = p^n \deg(\phi_{sep})$.

Proof. If ϕ is separable, then take $\phi_{sep} = \phi$ and $n = 0$. Consider ϕ inseparable. From Proposition 1.58 we have $\phi(x, y) = (r_1(x^p), r_2(x^p)y^p)$ with $r_1(x), r_2(x)$ rational functions with coefficients in K. Thus, $\phi = \phi_1 \circ \pi$ where $\phi_1 = (r_1(x), r_2(x)y)$. Note that $\deg(\phi_1)$ is $\deg(\phi)/p$. If ϕ_1 is separable then take $\phi_{sep} = \phi_1$ and $n = 1$. Otherwise, we can repeat the same argument with ϕ_1 and write $\phi = \phi_2 \circ \pi^2$ with $\deg(\phi_2) = \deg(\phi)/p^2$. We continue in this manner (this process must terminate since $\deg(\phi)$ is finite) to get $\phi = \phi_n \circ \pi^n$ with ϕ_n separable, and $\deg(\phi) = p^n \deg(\phi_n)$. □

Remark 1.66. Note that writing π^n is an abuse of notation because each π is not exactly the same morphism, as each might have different domain and codomain.

Definition 1.67. Let $\phi : E_1 \to E_2$ be an isogeny over a field K of positive characteristic p. Write $\phi = \phi_{sep} \circ \pi^n$, with ϕ_{sep} a separable isogeny. Then, the *separable degree of* ϕ is $\deg(\phi_{sep})$ and we denote it by $\deg_s(\phi)$. The *inseparable degree of* ϕ is p^n and we denoted by $\deg_i(\phi)$. An isogeny of the form π^n is said to be *purely inseparable*. We have $\deg(\phi) = \deg_s(\phi) \deg_i(\phi)$.

Definition 1.68. Let $\phi : E_1 \to E_2$ be an isogeny. We define $\ker(\phi)$ to be the kernel of the group homomorphism $E_1(\bar{K}) \to E_2(\bar{K})$ given by $P \mapsto \phi(P)$. In other words

$$\ker(\phi) = \{P \in E_1(\bar{K}) \mid \phi(P) = 0\}.$$

The size of the kernel of an isogeny is related to its degree, as the following theorem shows.

Theorem 1.69. *Let $\phi : E_1 \to E_2$ be a separable isogeny. Then*

$$\deg(\phi) = \#\ker(\phi)$$

Proof. Write $\phi = (r_1(x), r_2(x)y)$ in standard form, where $r_1(x) = p(x)/q(x)$ with $p(x)$ and $q(x)$ coprime. Since ϕ is separable, we have that $pq' - p'q$ is not the zero polynomial. Let S be the set of $x \in \bar{K}$ such that $(pq' - p'q)(x) = 0$. This set is finite. Choose $(a, b) \in E_2(\bar{K})$ such that

1. $a \neq 0$ and $b \neq 0$,
2. $\deg(p(x) - aq(x)) = \max\{\deg p(x), \deg q(x)\} = \deg(\phi)$ (equivalently, a is not equal to the ratio of the leading coefficients of p and q),
3. $a \notin r_1(S)$.

$E_2(\bar{K})$ is an infinite set, therefore such a point (a, b) must exist. Since ϕ is a group homomorphism, it suffices to determine the number of elements in $\phi^{-1}(a, b)$ since $\#\ker(\phi) = \#\phi^{-1}(a, b)$. Suppose $x_0, y_0 \in \bar{K}$ are such that $\phi(x_0, y_0) = (a, b)$. Then

$$\frac{p(x_0)}{q(x_0)} = a \text{ and } y_0 r_2(x_0) = b.$$

Since $(a,b) \neq O$ we have that $q(x_0) \neq 0$, and since $b \neq 0$ we have that $r_2(x_0) \neq 0$ and $y_0 = b/r_2(x_0)$. Therefore, y_0 is uniquely determined by x_0. Hence, we only need to count the possible values for x_0. That is, we must determine how many distinct roots $p - aq$ has (note that it has no roots in common with q). By assumption this polynomial has $\deg(\phi)$ roots in $\bar{K}$ counted with multiplicity. We show that it has no multiple roots.

Suppose $x_1 \in \bar{K}$ is a multiple root of $p - aq$. This means $p(x_1) = aq(x_1)$ and $p'(x_1) = aq'(x_1)$, which implies

$$ap(x_1)q'(x_1) = ap'(x_1)q(x_1)$$

and since $a \neq 0$, we have $(pq' - p'q)(x_1) = 0$, so $x_1 \in S$. But this is not possible since $a = r_1(x_1)$ and $a \notin r_1(S)$ by assumption. We conclude that $p - aq$ has exactly $\deg(\phi)$ distinct roots. It follows that

$$\#\ker(\phi) = \#\phi^{-1}(a,b) = \deg(\phi).$$

□

Corollary 1.70. An isogeny $\phi : E_1 \to E_2$ is inseparable if and only if $\#\ker(\phi) < \deg(\phi)$.

Proof. Write $\phi = \phi_{sep} \circ \pi^n$. Since π is bijective, we have that $\#\ker(\phi) = \#\ker(\phi_{sep})$. It follows that $\deg(\phi) = p^n \deg(\phi_{sep}) = p^n \#\ker(\phi)$. Since ϕ is inseparable if and only if $n \geq 1$, the result follows. □

Remark 1.71. Using the previous corollary it's easy to show that the composition of two isogenies $\phi \circ \psi$ is separable if and only if ϕ and ψ are both separable. Also note from Proposition 1.65 that the sum of two isogenies $\phi + \psi$ is either the zero or inseparable if ϕ and ψ are both inseparable.

We will now show that if $\phi : E_1 \to E_2$ is a separable isogeny, then $\bar{K}(E_1)/\phi^*\bar{K}(E_2)$ is a Galois extension. This will be useful later to prove the existence of certain isogenies. We start by noting that if E is an elliptic curve and $T \in E(\bar{K})$, then the map $\tau_T : E \to E$ that sends P to $P + T$ is a morphism over $\bar{K}$. We call this morphism τ_T the *translation-by-T map.*

Theorem 1.72. *Let $\phi : E_1 \to E_2$ be a separable isogeny over $\bar{K}$. The map*

$$\begin{aligned} \ker(\phi) &\to \mathrm{Aut}(\bar{K}(E_1)/\phi^*\bar{K}(E_2)) \\ T &\mapsto \tau_T^* \end{aligned}$$

is an isomorphism and $\bar{K}(E_1)/\phi^*\bar{K}(E_2)$ is a Galois extension.

Proof. We follow the proof of [18, III. Theorem 4.10(b)]. We start by showing that the map is well defined. Let $T \in \ker(\phi)$ and $f \in \bar{K}(E_2)$. Notice that $\phi \circ \tau_T = \phi$. Hence, $\tau_T^*(\phi^* f) = (\phi \circ \tau_T)^* f = \phi^* f$. Therefore, τ_T^* indeed fixes $\phi^*\bar{K}(E_2)$, and it's easy to see that it's an automorphism of $\bar{K}(E_1)$. Now, let $S \in \ker(\phi)$. We have that $\tau_{S+T} = \tau_T \circ \tau_S$ so $\tau_{S+T}^* = \tau_S^* \circ \tau_T^*$, and the map is a group homomorphism.

The number of automorphisms $\#\operatorname{Aut}(\bar{K}(E_1)/\phi^*\bar{K}(E_2))$ is bounded above by the degree of the extension $[\bar{K}(E_1) : \phi^*\bar{K}(E_2)]$ which is equal to $\deg(\phi)$. Since ϕ is separable, we have from Theorem 1.69 that $\deg(\phi) = \#\ker(\phi)$. Thus,

$$\#\operatorname{Aut}(\bar{K}(E_1)/\phi^*\bar{K}(E_2)) \leq \#\ker(\phi)$$

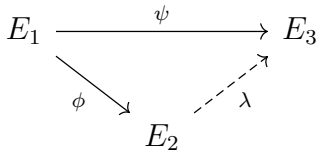
Hence, if we show that the map in injective, it will follow that it is an isomorphism. Suppose $T \in \ker(\phi)$ is such that $\tau_T^* = \tau_O^*$. Consider the coordinate function $x \in \bar{K}(E_1)$, then $\tau_O^* x = x$ which has a pole at O and no other poles. On the other hand, $\tau_T^* x$ has a pole at $-T$ and no other poles. Therefore, we must have that $T = O$ which proves the injectivity, so the map is an isomorphism.

Since $\#\ker(\phi) = [\bar{K}(E_1) : \phi^*\bar{K}(E_2)]$, we have

$$\#\operatorname{Aut}(\bar{K}(E_1)/\phi^*\bar{K}(E_2)) = [\bar{K}(E_1) : \phi^*\bar{K}(E_2)]$$

We conclude that $\bar{K}(E_1)/\phi^*\bar{K}(E_2)$ is a Galois extension. □

Corollary 1.73. Let $\phi : E_1 \to E_2$ and $\psi : E_1 \to E_3$ be isogenies defined over K. Suppose ϕ is separable. If $\ker(\phi) \subseteq \ker(\psi)$ then there is a unique isogeny $\lambda : E_2 \to E_3$ defined over K such that the following diagram commutes.



Proof. We follow the proof of [18, III. Corollary 4.11]. From the previous theorem, since ϕ is separable, we have that $\bar{K}(E_1)/\phi^*\bar{K}(E_2)$ is a Galois extension.

Let $\sigma \in \operatorname{Gal}(\bar{K}(E_1)/\phi^*\bar{K}(E_2))$. Then, by the isomorphism in the previous theorem, we have that $\sigma = \tau_T^*$ for some $T \in \ker(\phi)$. Since $\ker(\phi) \subseteq \ker(\psi)$, then $T \in \ker(\psi)$. Therefore, σ fixes $\psi^*\bar{K}(E_3)$.

It follows that $\psi^*\bar{K}(E_3)$ is inside the field fixed by every element in $\operatorname{Gal}(\bar{K}(E_1)/\phi^*\bar{K}(E_2))$ which, by Galois theory, is precisely $\phi^*\bar{K}(E_2)$. We thus have the inclusion

$$\psi^*\bar{K}(E_3) \subseteq \phi^*\bar{K}(E_2).$$

Observe that there is a unique injective homomorphism $\iota : \bar{K}(E_3) \to \bar{K}(E_2)$ such that the diagram

$$\begin{array}{ccc} \psi^*\bar{K}(E_3) & \hookrightarrow & \phi^*\bar{K}(E_2) \\ {\scriptstyle\psi^*}\uparrow\wr & & \wr\uparrow{\scriptstyle\phi^*} \\ \bar{K}(E_3) & \xhookrightarrow[\iota]{} & \bar{K}(E_2) \end{array}$$

commutes, where $\psi^*\bar{K}(E_3) \hookrightarrow \phi^*\bar{K}(E_2)$ is the inclusion map. Note that ι fixes $\bar{K}$ so we can use Proposition 1.34. This gives a unique morphism $\lambda : E_2 \to E_3$ defined over $\bar{K}$ such that $\lambda^* = \iota$. Thus, we have that

$$(\lambda \circ \phi)^* = \phi^* \circ \lambda^* = \psi^*$$

which implies $\psi = \lambda \circ \phi$.

We check that λ is a group homomorphism. Let P and Q in $\bar{K}(E_2)$ and let P' and Q' in $\bar{K}(E_1)$ be such that $\phi(P') = P$ and $\phi(Q') = Q$ (recall that ϕ is surjective). Then,

$$\begin{aligned}\lambda(P+Q) = \lambda(\phi(P') + \phi(Q')) = \lambda(\phi(P'+Q')) &= \psi(P'+Q') \\ &= \psi(P') + \psi(Q') = \lambda(P) + \lambda(Q).\end{aligned}$$

Hence, λ is an isogeny. It remains to show that it's defined over K.

To do this, we will show that it is fixed by the action of $\mathrm{Gal}(\bar{K}/K)$. Let $\sigma \in \mathrm{Gal}(\bar{K}/K)$. Then,

$$\psi = \sigma(\psi) = \sigma(\lambda) \circ \sigma(\phi) = \sigma(\lambda) \circ \phi$$

since ψ and ϕ are defined over K. Hence, $\psi = \sigma(\lambda) \circ \phi$. By the uniqueness of λ it must be the case that $\sigma(\lambda) = \lambda$. Since this is true for every element of $\mathrm{Gal}(\bar{K}/K)$, we can conclude that λ is defined over K.

□

1.4 Torsion points

Since the kernels of isogenies are finite subgroups of $E(\bar{K})$, it is very useful to understand the n-torsion subgroup of $E(\bar{K})$. To do this, we will first give a description of the multiplication-by-n map $[n] : E \to E$, which is the map that sends $P \mapsto nP$ and is, in fact, an isogeny.

We will work with an even shorter model with the intention of simplifying the computations. Let us now assume that $\mathrm{char}(K) \neq 3$. Note that if we have an elliptic curve E given by the equation

$$y^2 = x^2 + a_2x^2 + a_4x + a_6$$

then the following morphism

$$\begin{aligned}\phi : E &\to E' \\ (x, y) &\mapsto (x + a_2/3, y)\end{aligned}$$

where E' is the elliptic curve given by

$$y^2 = x^3 + Ax + B \tag{1.3}$$

with $A = a_4 - a_2^2/3$ and $B = 2a_2^3/27 - (a_2a_4)/3 + a_6$, is an isomorphism of curves. In fact, it's straightforward to verify that $\phi(P+Q) = \phi(P) + \phi(Q)$ so that ϕ is also a group isomorphism. Hence, it suffices to study the group structure of elliptic curves with equations of the form (1.3) with $A, B \in K$.

Remark 1.74. An equation of the form $y^2 = x^3 + Ax + B$ is called a *short Weierstrass equation*. From Proposition 1.37 we see that a short Weierstrass equation defines an elliptic curve if and only if $4A^3 + 27B^2 \neq 0$.

Example 1.75. As an example, let E be given by $y^2 = x^3 + Ax + B$ and consider the multiplication-by-2 map. From the addition formulas we see that

$$[2](x,y) = \left(\left(\frac{3x^2 + A}{2y} \right)^2 - 2x, -\left(\frac{3x^2 + A}{2y} \right) \left(\left(\frac{3x^2 + A}{2y} \right)^2 - 3x \right) - y \right)$$

so $[2] : E \to E$ is a morphism since it's given by rational functions. Clearly, it's also a group homomorphism, so it is an endomorphism of E.

We will give a general description of the endomorphism $[n] : E \to E$ with the help of the division polynomials. Their properties can be proved by induction and direct computations which we omit here, but give references to where the interested reader can find the details.

Definition 1.76. We define the *division polynomials* $\psi_m \in \mathbb{Z}[x, y, A, B]$ by the recursive formulas

$$\begin{aligned}
\psi_0 &= 0 \\
\psi_1 &= 1 \\
\psi_2 &= 2y \\
\psi_3 &= 3x^4 + 6Ax^2 + 12Bx - A^2 \\
\psi_4 &= 4y(x^6 + 5Ax^4 + 20Bx^3 - 5A^2x^2 - 4ABx - 8B^2 - A^3) \\
\psi_{2m+1} &= \psi_{m+2}\psi_m^3 - \psi_{m-1}\psi_{m+1}^3 \text{ for } m \geq 2 \\
\psi_{2m} &= (2y)^{-1}(\psi_m)(\psi_{m+2}\psi_{m-1}^2 - \psi_{m-2}\psi_{m+1}^2) \text{ for } m \geq 2.
\end{aligned}$$

Remark 1.77. It's easy to show that the polynomial $(\psi_m)(\psi_{m+2}\psi_{m-1}^2 - \psi_{m-2}\psi_{m+1}^2)$ is divisible by $2y$ so that ϕ_{2m} is indeed a polynomial in $\mathbb{Z}[x, y, A, B]$. See [25, Lemma 3.3].

Furthermore, we define the following polynomials:

Definition 1.78. For $m \geq 0$ define

$$\begin{aligned}
\phi_m &= x\psi_m^2 - \psi_{m+1}\psi_{m-1} \\
\omega_m &= (4y)^{-1}(\psi_{m+2}\psi_{m-1}^2 - \psi_{m-2}\psi_{m+1}^2)
\end{aligned}$$

Lemma 1.79. *For every integer $n \geq 0$ we have*

$$\begin{aligned}
\psi_n &\in \begin{cases} \mathbb{Z}[x, y^2, A, B] \text{ if } n \text{ is odd.} \\ 2y\mathbb{Z}[x, y^2, A, B] \text{ if } n \text{ is even.} \end{cases} \\
\phi_n &\in \mathbb{Z}[x, y^2, A, B] \\
\omega_n &\in \begin{cases} y\mathbb{Z}[x, y^2, A, B] \text{ if } n \text{ is odd.} \\ \mathbb{Z}[x, y^2, A, B] \text{ if } n \text{ is even.} \end{cases}
\end{aligned}$$

Proof. This follows by induction. See [25, Lemma 3.3 and Lemma 3.4] for details. □

Consider an elliptic curve E given by the equation $y^2 = x^3 + Ax + B$ with $A, B \in K$. We will consider the previous polynomials modulo the curve equation, so after replacing y^2 with $x^3 + Ax + B$ we may write $\phi_n(x)$ and $\psi_n^2(x)$, regarding both as polynomials in $\mathbb{Z}[x, A, B]$.

Lemma 1.80. *For every integer $n \geq 0$ we have*

$$\phi_n(x) = x^{n^2} + \text{ terms of lower degree in } x$$
$$\psi_n^2(x) = n^2 x^{n^2-1} + \text{terms of lower degree in } x$$

Proof. This is again shown by induction. See [25, Lemma 3.5] and [21, Lemma 6.21]. □

Proposition 1.81. Let E be an elliptic curve over K given by the equation $y^2 = x^3 + Ax + B$ and let n be a positive integer. Then the multiplication-by-n map $[n] : E \to E$ is an isogeny defined over K given by

$$[n](x, y) = \left(\frac{\phi_n(x)}{\psi_n^2(x)}, \frac{\omega_n(x, y)}{\psi_n(x, y)^3} \right)$$

Proof. This can be proved in an elementary way using induction and lengthy computations. See [21, Theorem 6.20] for an example. □

Lemma 1.82. *Let E be an elliptic curve over K given by the equation $y^2 = x^3 + Ax + B$ and let n be a positive integer. Then the polynomials $\phi_n(x)$ and $\psi_n^2(x)$ have no common roots in $\bar{K}$.*

Proof. See [25, Corollary 3.7 and Lemma 3.8]. □

Proposition 1.83. Let E be an elliptic curve over K given by $y^2 = x^3 + Ax + B$. Let $p = \text{char}(K)$ and let n be a positive integer. Then the multiplication-by-n map $[n] : E \to E$ has degree n^2. Furthermore, the map is separable if and only if p does not divide n.

Proof. Write $[n](x, y) = (r_1(x), r_2(x)y)$ in standard form. By Proposition 1.81 we have that $r_1(x) = \phi_n(x)/\psi_n^2(x)$. By Lemma 1.82 the polynomials $\phi_n(x)$ and $\psi_n^2(x)$ are coprime. From Lemma 1.80 we see that the maximum degree of these polynomials is n^2. Therefore, $\deg[n] = n^2$.

Observe from Lemma 1.80 that the leading term of the numerator of $r_1'(x)$ is

$$n^2 x^{n^2-1} n^2 x^{n^2-1} - n^2(n^2 - 1)x^{n^2-2} x^{n^2} = n^2 x^{2n^2-2}$$

which is different from zero if p does not divide n. Thus, $[n]$ is separable if p does not divide n.

Let us show that $[p]$ is not separable. Note that, by Lemma 1.79 and since p is odd, ψ_p can be written as a polynomial in x. Observe that the elements in the kernel of $[p]$ are

O and the affine points $(x_0, y_0) \in E(\bar{K})$ such that $\psi_p(x_0) = 0$. From Lemma 1.80 we see that the term $p^2 x^{p^2-1}$ of $\psi_p^2(x)$ is zero, hence the degree of $\psi_p^2(x)$ is less than $p^2 - 1$. It follows that $\psi_p(x)$ has degree less than $(p^2-1)/2$, so it has less than $(p^2-1)/2$ different roots. For each $x_0 \in \bar{K}$ there are at most two affine points in E with x-coordinate equal to x_0. Hence, the number of points $(x_0, y_0) \in E(\bar{K})$ such that $\psi_p(x_0) = 0$ is less than $p^2 - 1$, which means that the kernel of $[p]$ consists of less than p^2 points. From Corollary 1.70 we conclude that $[p]$ is inseparable. If p divides n, then we can write $[n] = [mp] = [m] \circ [p]$ with m some integer, which implies that $[n]$ is inseparable.

□

Remark 1.84. If n is a negative integer, then the multiplication-by-n map can be written as $[n] = [-1] \circ [-n]$ with $[-1] : (x, y) \mapsto (x, -y)$, a separable isogeny of degree 1. Hence, the previous proposition is actually true for all non-zero integers.

Definition 1.85. Let E be an elliptic curve and n a positive integer. We write $E[n]$ for the kernel of $[n] : E \to E$. Equivalently, it is the subgroup of n-torsion points

$$E[n] = \{P \in E(\bar{K}) \mid nP = O\}.$$

Theorem 1.86. *Let E be an elliptic curve over K and let $p = \mathrm{char}(K)$. Then, for every prime $l \neq p$ and integer $e \geq 1$ we have*

$$E[l^e] \cong \frac{\mathbb{Z}}{l^e\mathbb{Z}} \times \frac{\mathbb{Z}}{l^e\mathbb{Z}}.$$

If $p > 0$, then either

$$E[p^e] = \{O\} \text{ for all } e \geq 1 \quad \text{or} \quad E[p^e] \cong \frac{\mathbb{Z}}{p^e\mathbb{Z}} \text{ for all } e \geq 1.$$

Proof. Consider a prime $l \neq p$. Since p doesn't divide l we have that $[l] : E \to E$ is a separable isogeny of degree l^2 by Proposition 1.83. Therefore, by Theorem 1.69, we have that $\#E[l] = l^2$. But every non-zero element of $E[l]$ has order l, so it must be the case that $E[l] \cong \mathbb{Z}/l\mathbb{Z} \times \mathbb{Z}/l\mathbb{Z}$. Now, consider $e > 1$. By an analogous reasoning, $\#E[l^e] = l^{2e}$. We know that $E[l^e]$ is a finite l-subgroup, so we can write

$$E[l^e] \cong G_1 \oplus \cdots \oplus G_k$$

with $k \geq 1$ and G_i a cyclic group of order l^{e_i} for some $e_i \geq 1$. We see that

$$E[l] = l^{e_1-1}G_1 \oplus l^{e_2-1}G_2 \oplus \cdots \oplus l^{e_k-1}G_k$$

with $l^{e_i-1}G_i$ is a cyclic group of order l. So $k = 2$. Since every element of $E[l^e]$ has order at most l^e and $\#E[l^e] = l^{2e}$, we conclude that $E[l^e] \cong \mathbb{Z}/l^e\mathbb{Z} \times \mathbb{Z}/l^e\mathbb{Z}$.

Now, $[p] : E \to E$ is inseparable of degree p^2 by Proposition 1.83, so $\#E[p] < p^2$. Hence, either $\#E[p] = 0$ or $\#E[p] = p$. If $\#E[p] = 0$ then $\#E[p^e] = 0$ for every $e \geq 1$.

Let us consider the case $\#E[p] = p$. Clearly, $E[p] \cong \mathbb{Z}/p\mathbb{Z}$. Let $e > 1$. Then, by an argument analogous to the one above, one can easily show that $E[p^e]$ is cyclic since $E[p]$ is cyclic. Every element of $E[p^e]$ has order at most p^e, so it remains to show that it has an element of order exactly p^e. We show this by induction. This is true for $e = 1$. Consider $e > 1$ and let us assume the result for $e - 1$. So, by the induction hypothesis, there exists $P \in E(\bar{K})$ of order p^{e-1}. Since $[p] : E \to E$ is surjective, there is $Q \in E(\bar{K})$ such that $pQ = P$. Thus, Q has order p^e which finishes the induction. We conclude that $E[p^e]$ has a point of order p^e and thus $E[p^e] \cong \mathbb{Z}/p^e\mathbb{Z}$.

□

Corollary 1.87. Let E be an elliptic curve over $K$and let $p = \mathrm{char}(K)$. If n is a positive integer and p does not divide n, then

$$E[n] \cong \frac{\mathbb{Z}}{n\mathbb{Z}} \times \frac{\mathbb{Z}}{n\mathbb{Z}}.$$

Proof. This is an easy consequence of the previous theorem and the fact that $E[m_1 m_2] = E[m_1] \oplus E[m_2]$ if m_1 and m_2 are coprime integers. □

1.5 The dual isogeny

If there is an isogeny $\phi : E_1 \to E_2$ between two curves, we might wonder if there exists a related isogeny from E_2 to E_1. The answer is yes, and it satisfies a very useful property.

Theorem 1.88. *Let $\phi : E_1 \to E_2$ be an isogeny over K of degree m. Then there exists a unique isogeny $\hat{\phi} : E_2 \to E_1$ over K satisfying*

$$\hat{\phi} \circ \phi = [m].$$

We call it the dual isogeny to ϕ.

Proof. We start by showing uniqueness. Suppose there are two isogenies $\hat{\phi}$ and $\hat{\phi}'$ with this property. Then $(\hat{\phi}' - \hat{\phi}) \circ \phi = [m] - [m] = [0]$, where $[0]$ is the zero morphism. Since ϕ is non-constant, then $\hat{\phi}' - \hat{\phi}$ is the constant morphism sending $P \mapsto O$. Hence, $\hat{\phi}' = \hat{\phi}$.

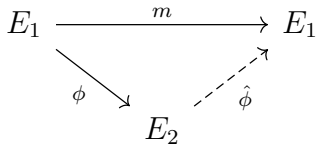
Suppose $\psi : E_2 \to E_3$ is an isogeny of degree n and suppose that we know $\hat{\phi}$ and $\hat{\psi}$ exist. Then

$$(\hat{\phi} \circ \hat{\psi}) \circ (\psi \circ \phi) = \hat{\phi} \circ [n] \circ \phi = [n] \circ \hat{\phi} \circ \phi = [n] \circ [m] = [nm],$$

which implies the existence of $\widehat{\psi \circ \phi} = \hat{\phi} \circ \hat{\psi}$.

By Proposition 1.65, the isogeny ϕ can be decomposed as $\phi = \phi_{sep} \circ \pi^n$ for some $n \geq 0$ and ϕ_{sep} a separable isogeny. Hence, it suffices to prove the existence of $\hat{\phi}$ in the case that ϕ is separable and the case $\phi = \pi$. We now consider these two cases:

- **Case 1.** ϕ is separable. From Theorem 1.69 we know that $\#\ker(\phi) = \deg(\phi) = m$ so $\ker(\phi) \subseteq E_1[m]$. It follows from Corollary 1.73 that there is an isogeny $\hat{\phi} : E_2 \to E_1$ defined over K such that the diagram



commutes. That is, $\hat{\phi} \circ \phi = [m]$.

- **Case 2**. ϕ is the Frobenius morphism. Let $p = \mathrm{char}(K) > 0$. We know from Proposition 1.83 that $[p] : E_1 \to E_1$ is not separable. Hence, $[p] = \varphi \circ \pi^n$ with $n \geq 1$ and φ a separable isogeny. Take $\hat{\pi} = \varphi \circ \pi^{n-1}$.

□

We now give some properties of the dual isogeny.

Proposition 1.89. Let $\phi : E_1 \to E_2$ be an isogeny and let $m = \deg(\phi)$. Then

1. $\deg(\hat{\phi}) = m$.
2. $\hat{\phi} \circ \phi = [m]$ in E_1 and $\phi \circ \hat{\phi} = [m]$ in E_2.
3. $\hat{\hat{\phi}} = \phi$.
4. For all $m \in \mathbb{Z}$, $\widehat{[m]} = [m]$.
5. If $\lambda : E_2 \to E_3$ is another isogeny, then $\widehat{\lambda \circ \phi} = \hat{\phi} \circ \hat{\lambda}$.
6. If $\psi : E_1 \to E_2$ is another isogeny, then $\widehat{\phi + \psi} = \hat{\phi} + \hat{\psi}$.

Proof.
(1) Recall that $\deg[m] = m^2$, so

$$\deg(\hat{\phi}) \deg(\phi) = \deg(\hat{\phi} \circ \phi) = \deg[m] = m^2$$

implies that $\deg(\hat{\phi}) = m$.
(2) $\hat{\phi} \circ \phi = [m]$ is true by definition. Note that

$$(\phi \circ \hat{\phi}) \circ \phi = \phi \circ (\hat{\phi} \circ \phi) = \phi \circ [m] = [m] \circ \phi.$$

It follows that $\phi \circ \hat{\phi} = [m]$. This also implies (3).
(4) Clearly $[m] \circ [m] = [m^2]$ and since $\deg[m] = m^2$ it follows from the uniqueness of the dual isogeny that $\widehat{[m]} = [m]$.
(5) It's straightforward to verify that $\hat{\phi} \circ \hat{\lambda}$ satisfies the defining property of the dual isogeny to $\lambda \circ \phi$.
(6) We refer to [18, III. Theorem 6.2(c)] for a proof. □

1.6 Constructing isogenies

We have shown that the kernel of an isogeny from E to another elliptic curve is a finite subgroup of $E(\bar{K})$. It is natural to ask ourselves if the converse is true: is every finite subgroup of $E(\bar{K})$ the kernel of an isogeny with domain E? The answer to this question turns out to be yes, and it will be of utmost importance in the cryptographic scheme that we will later discuss.

More precisely, in this section we will show that if G is a finite subgroup of $E(\bar{K})$, then there is an elliptic curve E' over $\bar{K}$ and an isogeny $\phi : E \to E'$ defined over $\bar{K}$ such that $\ker(\phi) = G$. For any practical application, the existence of this isogeny is not enough. We need a way to compute it. Thus, we will sketch a constructive proof that appears in [25, Theorem 12.16] with the explicit formulas due to Vélu [23].

Theorem 1.90. *Let E be an elliptic curve over K given by the equation*

$$y^2 = x^3 + a_2x^2 + a_4x + a_6.$$

Let G be a finite subgroup of $E(\bar{K})$. Then there exists an elliptic curve E' over $\bar{K}$ and and a separable isogeny $\phi : E \to E'$ defined over $\bar{K}$ such that $\ker(\phi) = G$.

Moreover, for a point $Q = (x_Q, y_Q) \in G$ different from O, define the following quantities

$$\begin{aligned} g_Q^x &= 3x_Q^2 + 2a_2x_Q + a_4 \\ g_Q^y &= -2y_Q \\ v_Q &= \begin{cases} g_Q^x & \text{if } 2Q = O \\ 2g_Q^x & \text{if } 2Q \neq O \end{cases} \\ u_Q &= (g_Q^y)^2. \end{aligned}$$

Let G_2 be the subset of points of order two of G, and let R be a subset of G such that we have the disjoint union $G = \{O\} \sqcup G_2 \sqcup R \sqcup (-R)$. That is, for any $P \in G$ that is not a 2-torsion point, put exactly one, either P or $-P$, in R. Let $S = G_2 \cup R$. Let

$$v = \sum_{Q \in S} v_Q, \quad w = \sum_{Q \in S} (u_Q + x_Q v_Q)$$

Then E' has the equation

$$Y^2 = X^3 + A_2X^2 + A_4X + A_6$$

with $A_2 = a_2$, $A_4 = a_4 - 5v$, $A_6 = a_6 - 4a_2v - 7w$.

The isogeny is given by $\phi(x, y) = (X(x, y), Y(x, y))$, where

$$X(x,y) = x + \sum_{Q \in S} \left(\frac{v_Q}{x - x_Q} + \frac{u_Q}{(x - x_Q)^2} \right) \tag{1.4}$$

$$Y(x,y) = y - \sum_{Q \in S} \left(u_Q \frac{2y}{(x - x_Q)^3} + v_Q \frac{y - y_Q}{(x - x_Q)^2} - \frac{g_Q^x g_Q^y}{(x - x_Q)^2} \right). \tag{1.5}$$

Proof. First of all, note that if $Q = (x_Q, y_Q)$, then $-Q = (x_Q, -y_Q)$. Thus, $v_Q = v_{-Q}$ and $u_Q = u_{-Q}$. This makes the value of v and w independent of the choice of R. Similarly, one can show that the functions X and Y are independent of the choice of R.

Let x, y be the affine coordinate functions on E. Let $t = x/y$ and $s = 1/y$. From Proposition 1.41 we see that t has a zero of order 1 at O and thus it's a uniformizer for E at O. We also note that s has a zero of order 3 at O. Dividing $y^2 = x^3 + a_2x^2 + a_4x + a_6$ by y^3 we find that

$$s = t^3 + a_2t^2s + a_4ts^2 + a_6s^3 \tag{1.6}$$

If we again substitute this expression for s in the equation above, we get

$$\begin{aligned} s &= t^3 + a_2t^2(t^3 + a_2t^2s + a_4ts^2 + a_6s^3) \\ &\quad + a_4t(t^3 + a_2t^2s + a_4ts^2 + a_6s^3)^2 + a_6(t^3 + a_2t^2s + a_4ts^2 + a_6s^3)^3 \\ &= t^3 + a_2t^5 + a_4t^7 + a_6t^9 + a_2^2t^4s + a_2a_4t^3s^2 + 2a_2a_4t^6s + O(t^{10}) \end{aligned}$$

Where $O(t^{10})$ denotes a function that vanishes with order at least 10 at O. Substitute (1.6) two more times to find

$$s = t^3 + a_2t^5 + (a_2^2 + a_4)t^7 + (a_2^3 + 3a_2a_4 + a_6)t^9 + O(t^{10}).$$

We can think of this as the expansion of s around O, in an analogous way to the Laurent series in complex analysis (we refer to [18, IV.1] for more details). Using the expansion for s, and since $y = 1/s$, we can compute the expansion of y around O to be

$$y = t^{-3} - a_2t^{-1} - a_4t - (a_2a_4 + a_6)t^3 + O(t^4)$$

and since $x = ty$, we have that

$$x = t^{-2} - a_2 - a_4t^2 - (a_2a_4 + a_6)t^4 + O(t^5)$$

We can substitute these expansions in the equations for X and Y in (1.4) and (1.5) to find expressions for X and Y in terms of t. See [23, page 4] for the explicit expressions. A computation then shows that

$$Y^2 = X^3 + A_2X^2 + A_4X + A_6 + O(t)$$

where the A_i are defined as in the statement of the theorem.

Looking at (1.4) and (1.5) we see that the only poles of X and Y are the points in G. Therefore, the function $Y^2 - X^3 - A_2X^2 - A_4X - A_6$ may have poles only at the points of G. Since this function is $O(t)$, it has a zero at O. We claim that it has a zero at every point of G. To prove this, we show that X and Y are invariant under translation by elements of G.

Let $P = (x_P, y_P)$ and $Q = (x_Q, y_Q)$ be points on E and write $P + Q = (x_{P+Q}, y_{P+Q})$. The following follows from the addition formulas and some lengthy computations (see [16, Lemma 3.3] for explicit calculations). If $2Q = O$, then $u_Q = 0$ and

$$x_{P+Q} - x_Q = \frac{v_Q}{x_P - x_Q},$$

Moreover, if $2Q \neq O$ then

$$x_{P+Q} - x_Q + x_{P-Q} - x_{-Q} = \frac{v_Q}{x_P - x_Q} + \frac{u_Q}{(x_P - x_Q)^2}$$

This shows that

$$X(P) = x(P) + \sum_{O \neq Q \in G} (x(P+Q) - x(Q))$$

From which is easy to see that X is invariant under translations by points of G. One can similarly show that

$$Y(P) = y(P) + \sum_{O \neq Q \in G} (y(P+Q) - y(Q))$$

so that Y is also invariant under translations by points of G.

Since X and Y are invariant under translations by G, and $Y^2 - X^3 - A_2X^2 - A_4X - A_6$ has a zero at O, we conclude that it has a zero at every point of G. But the only points it could have poles are precisely the points of G, hence, it has no poles. From Proposition 1.32 we conclude that $Y^2 - X^3 - A_2X^2 - A_4X - A_6$ is constant, and since it has a zero it must be the zero function.

This shows that $\phi(x, y) = (X, Y)$ is a rational map from E to the curve with equation $Y^2 = X^3 + A_2X^2 + A_4X + A_6$. It can be shown that the polynomial $X^3 + A_2X^2 + A_4X + A_6$ has distinct roots [25, Lemma 12.17]. It follows that this curve is nonsingular and thus an elliptic curve.

As we previously remarked, the poles of X and Y are precisely the points of G. Thus, we have $\phi(P) = O$ if and only if $P \in G$. In other words, $\ker(\phi) = G$.

Writing $X = p(x)/q(x)$ with $p(x)$ and $q(x)$ coprime, one can show [25, Exercise 12.8] that $\deg(\phi) = \max\{\deg(p), \deg(q)\} = \deg(p) = \#G$, so ϕ is a separable isogeny.

□

Note that the elliptic curve E' and the isogeny $\phi : E \to E'$ from the previous theorem need not be defined over K. In order for the isogeny and the codomain to be defined over K, we must ask G to satisfy an additional property.

If E is defined over K, then $\mathrm{Gal}(\bar{K}/K)$ acts on $E(\bar{K})$ by acting on the coordinates of a point. That is, for $\sigma \in \mathrm{Gal}(\bar{K}/K)$ we have $\sigma(O) = O$ and $\sigma(P) = (\sigma(x), \sigma(y))$ for a point $P = (x, y)$ on E. Note that the map $P \mapsto \sigma(P)$ is injective. Also observe that if E' is an elliptic curve over K and $\phi : E \to E'$ is an isogeny over K, then $\phi(\sigma(P)) = \sigma(\phi(P))$ for every $P \in E(\bar{K})$, since the coefficients of ϕ lie in K.

Definition 1.91. Let E be an elliptic curve over K and G a finite subgroup of $E(\bar{K})$. We say that G is *defined over* K if $\sigma(G) = G$ for every $\sigma \in \mathrm{Gal}(\bar{K}/K)$.

Remark 1.92. A subgroup G defined over K is not necessarily contained in $E(K)$.

Example 1.93. Let E_1 and E_2 be elliptic curves over K and let $\phi : E_1 \to E_2$ be an isogeny defined over K. Clearly, the kernel of ϕ is not necessarily contained in $E(K)$. Let $\sigma \in \mathrm{Gal}(\bar{K}/K)$. Note that if $P \in \ker(\phi)$ then

$$\phi(\sigma(P)) = \sigma(\phi(P)) = \sigma(O) = O$$

So $\sigma(P) \in \ker(\phi)$. Hence, $\sigma(\ker(\phi)) = \ker(\phi)$. We conclude that $\ker(\phi)$ is defined over K.

Corollary 1.94. If E is an elliptic curve over K and G is a finite subgroup of $E(\bar{K})$ defined over K, then there exists an elliptic curve E' over K and a separable isogeny $\phi : E \to E'$ defined over K such that $\ker(\phi) = G$.

Proof. We claim that the curve E' and the isogeny $\phi : E \to E'$ given in the statement of Theorem 1.90 are defined over K. Consider the partition $G = \{O\} \sqcup G_2 \sqcup R \sqcup (-R)$ and let $\sigma \in \mathrm{Gal}(\bar{K}/K)$. We have that $\sigma(G) = G$ and note that if P is a 2-torsion point, then $\sigma(P)$ is again a 2-torsion point in G (since the multiplication-by-2 map is defined over K), thus $\sigma(G_2) = G_2$. Also note that $\sigma(-R) = -\sigma(R)$. Therefore, we have another partition of G given by

$$G = \sigma(G) = \{O\} \sqcup G_2 \sqcup \sigma(R) \sqcup (-\sigma(R)).$$

Furthermore, note that for a point Q in G we have $\sigma(u_Q) = u_{\sigma(Q)}$ and $\sigma(v_Q) = v_{\sigma(Q)}$. As we noted in the beginning of the proof of Theorem 1.90, the values v, w and the functions X and Y are independent of the choice of R. By choosing $R' = \sigma(R)$ we can show that $\sigma(v) = v$, $\sigma(w) = w$. For example,

$$\sigma(v) = \sum_{Q \in G_2 \cup R} \sigma(v_Q) = \sum_{Q \in G_2 \cup R} v_{\sigma(Q)} = \sum_{Q' \in G_2 \cup R'} v_{Q'} = v.$$

Similarly $\sigma(w) = w$. The same kind of reasoning also shows that $\sigma(X) = X$ (where σ acts on the coefficients of X) and $\sigma(Y) = Y$, and thus $\sigma(\phi) = \phi$. Since this is true for every $\sigma \in \mathrm{Gal}(\bar{K}/K)$, we conclude that $v, w \in K$ (hence, E' is defined over K) and that ϕ is defined over K. □

The codomain E' and the isogeny $\phi : E \to E'$ are completely determined (up to isomorphism) by the group G, as the following statement shows.

Proposition 1.95. Let E_1, E_2, E_3 be elliptic curves over K. Suppose $\phi : E_1 \to E_2$ and $\varphi : E_1 \to E_3$ are separable isogenies over K such that $\ker(\phi) = \ker(\varphi)$. Then there is an isomorphism $\psi : E_2 \to E_3$ defined over K such that $\psi \circ \phi = \varphi$.

Proof. Since $\ker(\phi) \subseteq \ker(\varphi)$, we get from Corollary 1.73 that there is an isogeny $\psi : E_2 \to E_3$ defined over K such that $\psi \circ \phi = \varphi$. We claim that ψ is an isomorphism. From the other containment $\ker(\varphi) \subseteq \ker(\phi)$, applying Corollary 1.73 again, we get that there is an isogeny $\rho : E_3 \to E_2$ defined over K such that $\rho \circ \varphi = \phi$. Hence,

$$(\psi \circ \rho) \circ \varphi = \psi \circ (\rho \circ \varphi) = \psi \circ \phi = \varphi.$$

Since φ is not constant, we conclude that $\psi \circ \rho$ is the identity on E_3. Similarly, $\rho \circ \psi$ is the identity on E_2. Therefore, ψ has an inverse morphism defined over K and is therefore an isomorphism. $\square$

Definition 1.96. Let E be an elliptic curve and G a finite subgroup of $E(\bar{K})$ defined over K. Let $\phi : E \to E'$ be a separable isogeny over K with kernel G, where E' is an elliptic curve over K. We denote the curve E' (which is unique up to isomorphism) by E/G and call it E *modulo* G.

2 Orders in imaginary quadratic fields

Let E be an elliptic curve over K. In Chapter 3 we will see that the set of isogenies from E to E together with the zero morphism form a ring called the *endomorphism ring of* E. This ring is an essential ingredient in the cryptographic protocol that we will discuss, and –in some cases– it is isomorphic to an order inside a quadratic number field. To make sense of these terms, we will take a short detour into algebraic number theory.

We are mainly interested in the specific case of orders inside an imaginary quadratic number field, but we start with a higher level of generality by considering arbitrary number rings.

2.1 Number rings

For this section we mainly follow Chapter 2 and Chapter 3 of [19]. Our main goal will be to introduce the Kummer-Dedekind theorem for number rings, which will allow us to find the explicit description of prime ideals. We start by defining the concept of number field and number ring.

Definition 2.1. A *number field* $\mathcal{K}$ is a finite field extension of $\mathbb{Q}$. A *number ring* is a subring of a number field.

Example 2.2. The field $\mathbb{Q}(\sqrt{5})$ obtained by adjoining a root of 5 to $\mathbb{Q}$ is a field extension of degree 2, hence a number field. The ring $\mathbb{Z}[\sqrt{5}]$ is a subring of this field, so it's a number ring. The ring $\mathbb{Z}[\frac{1}{2}]$ is a subring of $\mathbb{Q}$, thus also a number ring.

Example 2.3. The most important example of a number ring is the *ring of integers of a number field* . This consists of the elements of a number field $\mathcal{K}$ that satisfy a monic equation $x^n + a_1x^{n-1} + \ldots + a_n = 0$ with $n \geq 1$ and $a_i \in \mathbb{Z}$. It can be shown that this is indeed a subring of $\mathcal{K}$ (see [11, Proposition 2.2]), and we denoted by $\mathcal{O}_\mathcal{K}$.

For a number ring R, there is an associated abelian group called the *class group of* R. In order to define it, we need to introduce the concept of fractional and invertible ideals.

2.1.1 Fractional ideals

Definition 2.4. Let R be a number ring with field of fractions $\mathcal{K} = Q(R)$. Then a *fractional* R*-ideal* I is a nonzero R-submodule of $\mathcal{K}$ such that $xI \subseteq R$ for some $x \in \mathcal{K}^*$. A *principal* fractional R-ideal is a fractional R-ideal of the form xR with $x \in \mathcal{K}^*$. A fractional R-ideal which is contained in R (i.e. an R-ideal) is called an *integral ideal*.

Remark 2.5. The element $x \in \mathcal{K}^*$ such that $xI \subseteq R$ can be chosen to be in R. Note that xI is an R-module contained in R and thus an R-ideal.

If I and J are fractional R-ideals then their addition and multiplication is defined in the usual way: $I+J := \{i+j \mid i \in I, j \in J\}$ and $IJ := \{\sum_{i=1}^{n} x_i y_i \mid x_i \in I, y_i \in J, n \in \mathbb{Z}_{>0}\}$. These are again fractional R-ideals. We will usually write $(x_1, x_2 \ldots, x_n)$ to denote the fractional R-ideal $x_1R + x_2R + \ldots + x_nR$ with $x_i \in \mathcal{K}^*$.

Definition 2.6. A fractional R-ideal I is *invertible* if there exists some fractional R-ideal J such that $IJ = R$. If I is invertible, we write I^{-1} for such J.

Example 2.7. Consider the number ring $R = \mathbb{Z}[\sqrt{-11}]$ with field of fractions $\mathbb{Q}(\sqrt{-11})$. Let $I = (3, \sqrt{-11} - 1)$ and consider the fractional ideal $J = (1, \frac{\sqrt{-11}+1}{3})$. We have

$$IJ = (3, \sqrt{-11} - 1, \sqrt{-11} + 1, -4)$$

which is an R-ideal containing $1 = 4 - 3$. Thus, $IJ = R$ and we conclude that I is invertible. For a non-example consider the ideal $Z = (2, \sqrt{-11} - 1)$. A simple computation shows that $Z^2 = 2Z$. If Z were invertible, then multiplying the previous equation by Z^{-1} yields $Z = (2)$. It can be shown that Z is not principal, so $Z = (2)$ is impossible. Thus, Z is not invertible.

We denote by $\mathcal{I}(R)$ the set of invertible fractional R-ideals. Note that this is a group with the operation given by multiplication of ideals. The symbol $\mathcal{P}(R)$ denotes the set of principal fractional R-ideals. Every principal ideal has an inverse which is also a principal ideal. Furthermore, the product of two principal ideals is again principal, so $\mathcal{P}(I)$ is a subgroup of $\mathcal{I}(R)$. The group that measures to which extent are invertible ideals principal is called the *class group* of R, and is defined as follows.

Definition 2.8. The *class group* of a number ring R is the quotient group

$$\mathrm{Cl}(R) = \mathcal{I}(R)/\mathcal{P}(R)$$

Remark 2.9. Some texts like [11] and [19] call this group the *Picard group* of R and reserve the name *class group* for the Picard group of the ring of integers. We instead follow the terminology in [5].

Definition 2.10. Let R be a number ring with field of fractions $\mathcal{K}$. An R-ideal I is *proper* if $R = \{x \in \mathcal{K} \mid xI \subseteq I\}$.

Remark 2.11. If I is invertible, then $xI \subseteq I$ implies $xR = xII^{-1} \subseteq II^{-1} = R$, hence $x \in R$. Therefore, every invertible R-ideal is proper.

Since a number field is a finite extension of $\mathbb{Q}$, it is an algebraic extension. Hence, every element x in a number ring satisfies an equation $a_n x^n + \ldots + a_1 x + a_0 = 0$ with $a_i \in \mathbb{Z}$. We will use this fact to prove the following theorem.

Theorem 2.12. *Let R be a number ring and $I \subseteq R$ a non-zero ideal. Then I has finite index in R.*

Proof. Let $\mathcal{K} = Q(R)$ be the field of fractions of R. Clearly, $\mathcal{K}$ is a number field. Let $t = [\mathcal{K} : \mathbb{Q}]$. Now, let $x \in I$ be a non-zero element. Since x is algebraic over $\mathbb{Q}$, it satisfies an equation $a_n x^n + \ldots + a_1 x + a_0 = 0$ with $a_i \in \mathbb{Z}$ and $a_0 \neq 0$. We have that $a_i x^i \in I$ for $i = 1, \ldots, n$, so we conclude that $a_0 \in I$. Thus, I contains a positive integer $a = |a_0|$.

The map $R/aR \to R/I$ is a surjection, so we only need to prove that R/aR is finite. Let $M \subseteq R$ be a finitely generated subgroup. Since M has no torsion elements it is a free abelian group. Let k be the rank of M. Note that $k \leq t$ as any set with more than t elements is linearly dependent over $\mathbb{Q}$. Hence, $M \cong \mathbb{Z}^k$ and M/aM has finite order a^k. The natural map $M \to R/aR$ factors through M/aM, so its image has size at most $a^k \leq a^t$. Any finitely generated subgroup of R/aR is the image of the map $M \to R/aR$ for some finitely generated subgroup $M \subseteq R$, hence all finitely generated subgroups of R/aR have order at most a^t. This implies that R/aR has finite order at most a^t. □

Corollary 2.13. A number ring R is noetherian.

Proof. Let $I \subseteq R$ be a non-zero ideal and let $x \in I$ be a non-zero element. By the previous theorem $(x) = xR$ is of finite index in R and thus $(x) \subseteq I$ is of finite index in I. We have that $I/(x)$ consists of elements $x_1 + (x), \ldots, x_n + (x)$ for some $x_1, \ldots, x_n \in I$. Thus, for $y \in I$ we have that $y - x_i \in (x)$ for some i. This shows that $y \in (x, x_1, \ldots, x_n) \subseteq I$. We conclude that $I = (x, x_1, \ldots, x_n)$, thus every ideal is finitely generated. □

Corollary 2.14. All the non-zero prime ideals of a number ring R are maximal.

Proof. Let $\mathfrak{p} \subseteq R$ be a prime ideal. Thus, $R/\mathfrak{p}$ is an integral domain. From the previous theorem we have that $R/\mathfrak{p}$ is finite. Since every finite integral domain is a field, we conclude that $\mathfrak{p}$ is a maximal ideal. □

Given that an ideal has finite index in R, we can define its *norm* as follows.

Definition 2.15. Let R be a number ring and I a non-zero integral ideal. We define the *norm of* I as the index of I in R. That is

$$N(I) = \#R/I$$

2.1.2 Localization and primary decomposition

The prime ideals play an essential role in the understanding of a number ring. The non-zero prime ideals are often called the *primes* of the number ring. Given a prime $\mathfrak{p}$, it is useful to 'forget' everything in R that has nothing to do with $\mathfrak{p}$. For this, we introduce the concept of *localization* at $\mathfrak{p}$.

Definition 2.16. Let R be a number ring with field of fractions $\mathcal{K} = Q(R)$, and let $\mathfrak{p} \subseteq R$ be a non-zero prime ideal. We define the local number ring $R_\mathfrak{p}$ as

$$R_\mathfrak{p} = \left\{ \frac{r}{s} \in \mathcal{K} \mid r \in R, s \in R \setminus \mathfrak{p} \right\}$$

Note that $R \subseteq R_\mathfrak{p}$. The fact that $\mathfrak{p}$ is a prime ideal guarantees that if s_1 and s_2 are not in $\mathfrak{p}$, then neither is $s_1 s_2$. This makes $R_\mathfrak{p}$ into a subring of $\mathcal{K}$. Furthermore, it is indeed local since it has a unique maximal ideal given by

$$\mathfrak{p}R_\mathfrak{p} = \left\{ \frac{r}{s} \in \mathcal{K} \mid r \in \mathfrak{p}, s \in R \setminus \mathfrak{p} \right\}$$

which is the complement of the unit group $R_\mathfrak{p}^* = \left\{ \frac{r}{s} \in \mathcal{K} \mid r \notin \mathfrak{p}, s \notin \mathfrak{p} \right\}$.

Definition 2.17. Let R be a number ring, $\mathfrak{p} \subseteq R$ a non-zero prime ideal, and I a fractional R-ideal. We define the *localization of I at $\mathfrak{p}$* by

$$I_\mathfrak{p} = \left\{ \frac{i}{s} \mid i \in I, s \in R \setminus \mathfrak{p} \right\}$$

It's easy to see that $I_\mathfrak{p}$ is a fractional $R_\mathfrak{p}$-ideal. Furthermore, if I is an integral R-ideal, then $I_\mathfrak{p}$ is an integral $R_\mathfrak{p}$-ideal. On the other hand, if J is an integral $R_\mathfrak{p}$-ideal, then $J \cap R$ is an integral R-ideal with localization J. Also note that if I is an integral ideal not contained in the prime $\mathfrak{p}$, then $I_\mathfrak{p} = R_\mathfrak{p}$.

Remark 2.18. The localization of $\mathfrak{p}$ at $\mathfrak{p}$ is $\mathfrak{p}R_\mathfrak{p}$, the unique maximal ideal of $R_\mathfrak{p}$.

We can recover an ideal from its localizations as follows.

Proposition 2.19. Let R be a number ring and I a fractional R-ideal. Then

$$I = \bigcap_\mathfrak{p} I_\mathfrak{p}$$

where the intersection is taken over the localizations at all primes of R.

Proof. The containment $I \subseteq \cap_\mathfrak{p} I_\mathfrak{p}$ is clear. Consider $x \in \cap_\mathfrak{p} I_\mathfrak{p}$ and let $J = \{r \in R \mid rx \in I\}$, which is an R-ideal. Note that for each prime $\mathfrak{p}$, we can write $x = a/b$ with $a \in I$ and $b \in R \setminus \mathfrak{p}$. This implies that for each prime $\mathfrak{p}$ there is an element $b \in R \setminus \mathfrak{p}$ such that $b \in J$. It follows that J is not contained in any prime (and in particular any maximal) ideal. This implies $J = R$ and thus $x = 1 \cdot x \in I$ □

Proposition 2.20. Let $R_\mathfrak{p}$ be a local number ring. Then every non-zero ideal of $R_\mathfrak{p}$ contains a power of the maximal ideal.

Proof. It easy to show that the only non-zero prime ideal of $R_\mathfrak{p}$ is the maximal ideal $\mathfrak{p}R_\mathfrak{p}$. Furthermore, one can show that since R is noetherian, so is $R_\mathfrak{p}$.

Now, suppose the statement in the Proposition is false and let M be the set of all non-zero ideals that do not contain a power of the maximal ideal. Since M is not empty and $R_\mathfrak{p}$ is noetherian, then M contains a maximal element I (otherwise we could form an infinite strictly-ascending chain of ideals). The ideal I cannot be prime since it is not the maximal ideal $\mathfrak{p}R_\mathfrak{p}$. Let $x, y \in R_\mathfrak{p} \setminus I$ be such that $xy \in I$. Note that the ideals $I + (x)$ and $I + (y)$ strictly contain I, therefore they are not in M. This means that $I + (x)$ and $I + (y)$ contain some power of the maximal ideal. But note that $I \supseteq (I + (x))(I + (y))$, so I also contains a power of the maximal ideal, a contradiction. □

There is an important characterization of invertible ideals in terms of their localizations.

Theorem 2.21. *Let R be a number ring and I a fractional R-ideal. Then I is invertible if and only if the localization $I_{\mathfrak{p}}$ at each prime $\mathfrak{p}$ is a principal fractional $R_{\mathfrak{p}}$-ideal.*

Proof. Suppose I is an invertible fractional ideal and let $\mathfrak{p}$ be a prime. Then there are $x_i \in I$ and $y_i \in I^{-1}$ such that $\sum_i x_i y_i = 1$. Each $x_i y_i$ is in R but they cannot all be in the maximal ideal of $R_{\mathfrak{p}}$. Assume without loss of generality that $x_1 y_1 \notin \mathfrak{p}R_{\mathfrak{p}}$. Hence, $x_1 y_1$ is in $R_{\mathfrak{p}}^*$. Take $x \in I$ and note that we can write

$$x = x_1 \cdot (xy_1) \cdot (x_1 y_1)^{-1}$$

with $xy_1 \in R \subseteq R_{\mathfrak{p}}$ and $(x_1 y_1)^{-1} \in R_{\mathfrak{p}}$. Thus, $x \in x_1 R_{\mathfrak{p}}$. This implies that $I_{\mathfrak{p}} \subseteq x_1 R_{\mathfrak{p}}$. Since $x_1 \in I_{\mathfrak{p}}$, we conclude that $I_{\mathfrak{p}} = x_1 R_{\mathfrak{p}}$.

For the converse, write $I = x_1 R + \ldots + x_n R$ with $x_i \in Q(R)$ (every ideal of a number ring is finitely generated). Let $\mathfrak{p}$ be a prime and let $x \in I$ be such that $I_{\mathfrak{p}} = xR_{\mathfrak{p}}$. Since every $x_i \in I_{\mathfrak{p}}$, it is possible to write $x_i = x(r_i/s)$ with $r_i \in R$ and $s \in R \setminus \mathfrak{p}$ where s is independent of i. We have that $sx^{-1}x_i = r_i \in R$ for every i. Consider the fractional ideal $J = \{y \in Q(R) \mid yI \subseteq R\}$. Then $sx^{-1} \in J$. It follows that $s = x \cdot (sx^{-1}) \in IJ$. Note that IJ is an R-ideal and is not contained in the prime $\mathfrak{p}$. Since this is true for every prime $\mathfrak{p}$, IJ is not contained in a maximal ideal. We conclude that $IJ = R$ and I is invertible. □

Related to the localizations of an ideal are the primary parts. These are defined as follows.

Definition 2.22. Given an integral ideal I of a number ring and a prime $\mathfrak{p}$, we define the $\mathfrak{p}$-*primary part of* I to be

$$I_{(\mathfrak{p})} = I_{\mathfrak{p}} \cap R.$$

If $\mathfrak{p}$ and $\mathfrak{q}$ are two distinct primes of R, then $\mathfrak{p} + \mathfrak{q} = R$ as every prime is maximal. Thus, two distinct primes are always coprime. If $0 < n \leq m$ are integers, then $\mathfrak{p}^n$ and $\mathfrak{p}^m$ are also coprime as $\mathfrak{p}^n + \mathfrak{q}^m \supseteq \mathfrak{p}^m + \mathfrak{q}^m \supseteq (\mathfrak{p} + \mathfrak{q})^{2m} = R$.

From Proposition 2.20 we get that $I_{\mathfrak{p}}$ contains a power of the maximal ideal $\mathfrak{p}R_{\mathfrak{p}}$. Since $\mathfrak{p} \subseteq \mathfrak{p}R_{\mathfrak{p}}$ we have that $\mathfrak{p}^n \subseteq (\mathfrak{p}R_{\mathfrak{p}})^n \subseteq I_{\mathfrak{p}}$ for some $n \geq 1$. Hence, $\mathfrak{p}^n \subseteq I_{(\mathfrak{p})}$. Similarly, if $\mathfrak{q}$ is another prime different from $\mathfrak{p}$, then $I_{(\mathfrak{q})}$ contains a power of $\mathfrak{q}$ and is therefore coprime with $I_{(\mathfrak{p})}$.

Proposition 2.23. If I is a non-zero integral ideal of a number ring R, then it has a *primary decomposition* given by

$$I = \prod_{\mathfrak{p} \supseteq I} I_{(\mathfrak{p})}$$

Proof. Using Proposition 2.19 we have that

$$I = R \cap I = R \cap \bigcap_{\mathfrak{p}} I_{\mathfrak{p}} = \bigcap_{\mathfrak{p}} I_{\mathfrak{p}} \cap R = \bigcap_{\mathfrak{p}} I_{(\mathfrak{p})} = \bigcap_{\mathfrak{p} \supseteq I} I_{(\mathfrak{p})}.$$

Since I is of finite index in R, it follows that there are only finitely many primes that contain I. Hence, the intersection on the right hand side is finite. By the coprimality of the various $\mathfrak{p}$-primary parts, it follows that the intersection can actually be written as a product. □

We now give another characterization of an invertible prime ideal.

Theorem 2.24. *Let $\mathfrak{p}$ be a prime of a number ring R. Then $\mathfrak{p}$ is invertible if and only if $R_{\mathfrak{p}}$ is a principal ideal domain and every non-zero $R_{\mathfrak{p}}$-ideal is a power of $\mathfrak{p}R_{\mathfrak{p}}$.*

Proof. Suppose $\mathfrak{p}$ is invertible. From Theorem 2.21 we have that the localization of $\mathfrak{p}$ at $\mathfrak{p}$ is principal. Hence, there exists $\pi \in R_{\mathfrak{p}}$ such that $\mathfrak{p}R_{\mathfrak{p}} = \pi R_{\mathfrak{p}} = (\pi)$. Note that the chain of $R_{\mathfrak{p}}$ ideals $R_{\mathfrak{p}} \supseteq (\pi) \supseteq (\pi^2) \supseteq \ldots$ is strictly descending.

Let I be a non-zero $R_{\mathfrak{p}}$ ideal. It follows from Proposition 2.20 that I contains all sufficiently large powers of (π), hence there is a largest value $m \geq 0$ such that $I \subseteq (\pi^m)$. Thus, $I \setminus (\pi^{m+1})$ is not empty. Consider any $r \in I \setminus (\pi^{m+1})$. We have that $r = a\pi^m$ where $a \notin (\pi)$, hence a is in $R_{\mathfrak{p}}^*$. We conclude that $(r) = (\pi^m) \subseteq I \subseteq (\pi^m)$, which implies $I = (\pi^m)$.

For the converse, we have that $R_{\mathfrak{p}}$ is a principal ideal domain, so the localization of $\mathfrak{p}$ at $\mathfrak{p}$ is principal. Also note that for any prime $\mathfrak{q} \neq \mathfrak{p}$ the localization of $\mathfrak{p}$ at $\mathfrak{q}$ is $R_{\mathfrak{q}}$ (since the prime $\mathfrak{p}$ is not contained in $\mathfrak{q}$) which is also principal. Hence, from Theorem 2.21 we get that $\mathfrak{p}$ is invertible. □

Remark 2.25. In other words, $\mathfrak{p}$ is invertible if and only if $R_{\mathfrak{p}}$ is a discrete valuation ring.

Remark 2.26. Using the previous theorem, it is an easy exercise to show that if $\mathfrak{p}$ is an invertible ideal, then the $\mathfrak{p}$-primary part of an integral ideal I is a power of $\mathfrak{p}$. Indeed, since $R_{\mathfrak{p}}$ is a discrete valuation ring, we have that $I_{\mathfrak{p}} = (\mathfrak{p}R_{\mathfrak{p}})^k$ for some k. Compare the localizations of $I_{(\mathfrak{p})}$ and $\mathfrak{p}^k$ at every prime to conclude that the two ideals are identical. Thus, from the primary decomposition in Proposition 2.23, it follows that if all primes $\mathfrak{p} \supseteq I$ are invertible, then I can be written as a product of prime ideals.

As the above remark suggests, determining the primes of a number ring is very useful as it allows us to factor into primes some of the invertible ideals. Note that if $\mathfrak{p}$ is a prime, then $R/\mathfrak{p}$ is a finite field of characteristic p, for some prime number p. Hence, $\mathfrak{p}$ contains a unique prime number p.

2.1.3 The Kummer-Dedekind Theorem

Given a prime number p, we call the primes $\mathfrak{p} \supseteq (p)$ the *primes above* p. We will later give an explicit description of the primes above p in a number ring of a certain form. Moreover, we'll give a simple criteria to determine if they are invertible or not. If a prime is invertible, we say it's *regular*. If it's not, we call it *singular*.

Definition 2.27. A *simple integral extension of* $\mathbb{Z}$ is a number ring of the form $\mathbb{Z}[\alpha]$, where α is the root of an irreducible monic polynomial in $\mathbb{Z}[X]$.

Example 2.28. The number ring $\mathbb{Z}[\sqrt{-5}]$ is a simple integral extension of $\mathbb{Z}$, but $\mathbb{Z}[\frac{1}{2}]$ is not.

Remark 2.29. If $f \in \mathbb{Z}[X]$ is a monic irreducible polynomial and $\alpha \in \overline{\mathbb{Q}}$ is a root of f, then $\mathbb{Z}[\alpha] \cong \mathbb{Z}[X]/(f)$. Note that $\mathbb{Z}[\alpha]$ is a free $\mathbb{Z}$-module of rank $\deg(f)$. A number ring of this type is and *order* inside its field of fractions.

Definition 2.30. An *order* $\mathcal{O}$ in a number field $\mathcal{K}$ is a subring of $\mathcal{K}$ that is a finitely generated $\mathbb{Z}$-module and contains a $\mathbb{Q}$-basis for $\mathcal{K}$. Equivalently, $\mathcal{O} \subseteq \mathcal{K}$ is an order if it is a free $\mathbb{Z}$-module of rank $[\mathcal{K} : \mathbb{Q}]$.

Remark 2.31. It can be shown that the class group of an order is finite [19, Theorem 5.4].

Theorem 2.32 (Kummer-Dedekind)**.** *Let $f \in \mathbb{Z}[X]$ be an irreducible monic polynomial, $\alpha \in \overline{\mathbb{Q}}$ a root of f, and p a prime number. Let R be the simple integral extension $\mathbb{Z}[\alpha] \cong \mathbb{Z}[X](f)$. Let $g_i \in \mathbb{Z}[X]$ be monic polynomials such that $\overline{f} = (f \mod p)$ factors as*

$$\overline{f} = \prod_{i=1}^{s} \overline{g_i}^{e_i} \in \mathbb{F}_p[X]$$

where $e_i \in \mathbb{Z}_{\geq 1}$ and $\overline{g_i} = (g_i \mod p) \in \mathbb{F}_p[X]$ are irreducible monic polynomials pairwise distinct. We have the following:

1. *The primes above p are the ideals $\mathfrak{p}_i = pR + g_i(\alpha)R$. We have $\prod_{i=1}^{s} \mathfrak{p}_i^{e_i} \subseteq pR$.*
2. *The equality $\prod_{i=1}^{s} \mathfrak{p}_i^{e_i} = pR$ holds if and only if every $\mathfrak{p}_i$ is invertible.*
3. *Let $r_i \in \mathbb{Z}[X]$ be the remainder of f after division by g_i. Then $\mathfrak{p}_i$ is singular if and only if $e_i > 1$ and p^2 divides r_i.*

Proof. (1) The primes above p correspond to the kernels of surjective homomorphisms from R to a field of characteristic p. Any such homomorphism sends an integer a to $a \bmod p$ and α to a root of $\overline{f}$. Hence, α is sent to a root of $\overline{g_i}$ for some i. We have that there is a bijective correspondence between primes above p and the kernel of the homomorphisms

$$\phi_i : R \to \mathbb{F}_p[X]/(\overline{g_i})$$

where $\mathbb{F}_p[X]/(\overline{g_i})$ is indeed a field of characteristic p. In fact, it has size $p^{\deg g_i}$.

If $t \in \mathbb{Z}[X]$, then $t(\alpha)$ is mapped to $\overline{t} \mod \overline{g_i}$. Hence, $t(\alpha)$ is mapped to zero if and only if $\overline{g_i}$ divides $\overline{t}$. One can check that $\ker(\psi_i) = pR + g_i(\alpha)R$, so the primes above p are the ideals $\mathfrak{p}_i = pR + g_i(\alpha)R$ and we have $R/\mathfrak{p}_i \cong \mathbb{F}_p[X]/(\overline{g_i})$. Note that $\prod_{i=1}^{s} g_i(\alpha)$ is in $f(\alpha) + pR = pR$ so we have the inclusion $\prod_{i=1}^{s} \mathfrak{p}_i^{e_i} \subseteq pR$.

(2) If the equality holds, then each $\mathfrak{p}_i$ is invertible as pR is a principal ideal. For the converse, let's assume that all the $\mathfrak{p}_i$ are invertible. We already have $\prod_{i=1}^{s} \mathfrak{p}_i^{e_i} \subseteq pR$. Let $I = \prod_{i=1}^{s} \mathfrak{p}_i^{e_i}$. We only need to show that I and pR have same index in R. Note that R is a free $\mathbb{Z}$-module of rank $\deg(f)$. Thus, the index of pR in R is $p^{\deg(f)}$. From the Chinese remainder theorem and the coprimality of powers of distinct primes, we have that $\#(R/I) = \prod_{i=1}^{s} \#(R/{\mathfrak{p}_\mathfrak{i}}^{e_i})$.

Note that $\#R/\mathfrak{p}_\mathfrak{i} = p^{\deg g_i}$. It can be shown that for a prime $\mathfrak{p}$ and $n \geq 0$ an integer, $R/\mathfrak{p}^n$ is isomorphic to $R_\mathfrak{p}/(\mathfrak{p}R_\mathfrak{p})^n$ (see [19, pag. 22]). Since $R_{\mathfrak{p}_\mathfrak{i}}$ is a discrete valuation ring by Theorem 2.24, we have that all the quotients $(\mathfrak{p}_\mathfrak{i}R_{\mathfrak{p}_\mathfrak{i}})^k/(\mathfrak{p}_\mathfrak{i}R_{\mathfrak{p}_\mathfrak{i}})^{k+1}$ are isomorphic. Thus,

$$\#R/\mathfrak{p}_i^{e_i} = \#R_{\mathfrak{p}_\mathfrak{i}}/(\mathfrak{p}_\mathfrak{i}R_{\mathfrak{p}_\mathfrak{i}})^{e_i} = \prod_{k=0}^{e_i-1} \#(\mathfrak{p}_\mathfrak{i}R_{\mathfrak{p}_\mathfrak{i}})^k/(\mathfrak{p}_\mathfrak{i}R_{\mathfrak{p}_\mathfrak{i}})^{k+1} = \prod_{k=0}^{e_i-1} \#R_{\mathfrak{p}_\mathfrak{i}}/\mathfrak{p}_\mathfrak{i}R_{\mathfrak{p}_\mathfrak{i}} = p^{e_i \deg g_i}.$$

We conclude that $\#R/I = p^m$ with $m = \sum_{i=1}^{s} e_i \deg g_i = \deg f$. Hence, the index of I in R is equal to the index of pR in R. It follows that $I = pR$.

(3) Notice that the remainder of f upon division by g_i is in $p\mathbb{Z}[X]$, so there are polynomials $q_i, s_i \in \mathbb{Z}[X]$ such that $f = q_i g_i + p s_i$ where s_i is either zero or $\deg(s_i) < \deg(g_i)$. Recall that $f(\alpha) = 0$, so we get a relation between $g_i(\alpha)$ and p, the two generators of $\mathfrak{p}_i$, given by

$$ps_i(\alpha) = -q_i(\alpha)g_i(\alpha).$$

Note that $\mathfrak{p}_i R_{\mathfrak{p}_i} = pR_{\mathfrak{p}_\mathfrak{i}} + g_i(\alpha)R_{\mathfrak{p}_i}$.

Suppose $\overline{g_i}$ has exponent $e_i = 1$ in the factorization of $\overline{f}$. This implies that $\overline{g_i}$ does not divide $\overline{q_i}$ and thus $q_i(\alpha)$ is not in $\mathfrak{p}_i$. Hence, $q_i(\alpha)$ is in $R^*_{\mathfrak{p}_i}$. From the relation above we have that $g_i(\alpha) = -ps_i(\alpha)q_i(\alpha)^{-1} \in pR_{\mathfrak{p}_i}$ so $\mathfrak{p}_i R_{\mathfrak{p}_i} = pR_{\mathfrak{p}_i}$. We conclude that $\mathfrak{p}_i$ is locally principal and hence invertible. Similarly, if $r_i = ps_i$ is not divisible by p^2 then p does not divide s_i so $\overline{s_i}$ is not zero. Since $\deg(s_i) < \deg(g_i)$ we have that $\overline{g_i}$ does not divide $\overline{s_i}$ so $s_i(\alpha)$ is not in $\mathfrak{p}_i$. This implies that $s_i(\alpha) \in R^*_{\mathfrak{p}_i}$ and thus $p = -q_i(\alpha)s_i(\alpha)^{-1}g_i(\alpha) \in g_i(\alpha)R_{\mathfrak{p}_i}$. It follows that $\mathfrak{p}_i R_{\mathfrak{p}_i} = g_i(\alpha)R_{\mathfrak{p}_i}$ and as before, $\mathfrak{p}_i$ is locally principal, hence invertible.

For the converse, suppose $e_i \geq 2$ and $r_i \in p^2\mathbb{Z}[X]$. This implies that both $q_i(\alpha)$ and $s_i(\alpha)$ are in $\mathfrak{p}_i$. Consider the element $x = p^{-1}q_i(\alpha)$ which is in the field of fractions of R but it's not in R. We will show that $x\mathfrak{p}_i \subseteq \mathfrak{p}_i$, which implies that $\mathfrak{p}_i$ is not proper and, hence, not invertible:

$$p^{-1}q_i(\alpha)\mathfrak{p}_i = p^{-1}q_i(\alpha)pR + p^{-1}q_i(\alpha)g_i(\alpha)R = q_i(\alpha)R + (-s_i(\alpha))R \subseteq \mathfrak{p}_i.$$

□

2.2 Orders inside imaginary quadratic fields

After considering arbitrary number rings, we will now restrict our attention to the specific case of orders (recall Definition 2.30) inside an imaginary quadratic number field. Some

of the definitions and results in this section can be stated in much more generality, but restricting to this case (which is the one we will actually need) simplifies some of the proofs. For this section we will follow Chapter 7 of [5].

A *quadratic number field* $\mathcal{K}$ satisfies $[\mathcal{K} : \mathbb{Q}] = 2$. It can be written uniquely in the form $\mathcal{K} \cong \mathbb{Q}(\sqrt{N})$ with $N \neq 0, 1$ square-free. If $N < 0$ then $\mathcal{K}$ is said to be imaginary.

Let $\mathcal{K} = \mathbb{Q}(\tau)$ be an imaginary quadratic field with τ a root of $x^2 - N$. Note that $\mathcal{K}$ has two automorphisms that fix $\mathbb{Q}$: the trivial one, and the one sending an element $x = a + b\tau$ with $a, b \in \mathbb{Q}$ to $\overline{x} = a - b\tau$. The element $\overline{x}$ is called the *conjugate of* x.

Definition 2.33. Let $\mathcal{O} = \mathbb{Z}w_1 \oplus \mathbb{Z}w_2$ be a an order inside a quadratic number field $\mathcal{K}$. Let $x \mapsto \overline{x}$ be the non-trivial automorphism of $\mathcal{K}$. We define the *discriminant of* $\mathcal{O}$ to be the number

$$D = \left(\det \begin{pmatrix} w_1 & w_2 \\ \overline{w}_1 & \overline{w}_2 \end{pmatrix} \right)^2$$

Remark 2.34. Note that this definition is independent of the $\mathbb{Z}$-basis chosen for $\mathcal{O}$. It's also easy to verify that $D \in \mathbb{Q}$ (observe that $\overline{D} = D$). We will later see that it is in fact an integer.

Definition 2.35. Let $\mathcal{K} = \mathbb{Q}(\sqrt{N})$ be a quadratic number field with N square-free. Let $x \mapsto \overline{x}$ be its non-trivial automorphism. Given $x = a + b\sqrt{N}$ with $a, b \in \mathbb{Q}$, we define the *norm of* x by

$$N(x) = x\overline{x} = a^2 - b^2 N,$$

and the *trace of* x by

$$T(x) = x + \overline{x} = 2a$$

Remark 2.36. From the definitions, it's clear that for $x, y \in \mathcal{K}$ we have $N(xy) = N(x)N(y)$ and $T(x + y) = T(x) + T(y)$.

Remark 2.37. Observe that $x \in \mathcal{K}$ satisfies the monic equation $X^2 - T(x)X + N(x)$. Therefore, if $N(x)$ and $T(x)$ are both in $\mathbb{Z}$, then $x \in \mathcal{O}_\mathcal{K}$. Conversely, if x is in $\mathcal{O}_\mathcal{K}$ then x is either an integer (in which case $T(x), N(x) \in \mathbb{Z}$) or it has minimal polynomial $X^2 - T(x)X + N(x)$. The minimal polynomial of an element in $\mathcal{O}_\mathcal{K}$ lives in $\mathbb{Z}[X]$ by Gauss Lemma, so $T(x), N(x) \in \mathbb{Z}$. Hence, $x \in \mathcal{O}_\mathcal{K}$ if and only if $N(x), T(x) \in \mathbb{Z}$.

Proposition 2.38. Let $\mathcal{K} = \mathbb{Q}(\sqrt{N})$ be a quadratic number field with N square-free. The ring of integers of $\mathcal{K}$ is given by

$$\mathcal{O}_\mathcal{K} = \begin{cases} \mathbb{Z}[\sqrt{N}] & \text{if } N \not\equiv 1 \mod 4 \\ \mathbb{Z}\left[\frac{1+\sqrt{N}}{2}\right] & \text{if } N \equiv 1 \mod 4 \end{cases}$$

Proof. This can be shown using the previous remark that $x \in \mathcal{O}_\mathcal{K}$ if and only if $N(x), T(x) \in \mathbb{Z}$. □

Note that $\mathcal{O}_\mathcal{K}$ is an order, and thus we can compute the discriminant $d_\mathcal{K}$ to be

$$d_\mathcal{K} = \begin{cases} 4N & \text{if } N \not\equiv 1 \mod 4 \\ N & \text{if } N \equiv 1 \mod 4 \end{cases}$$

Hence, if we define $w_\mathcal{K} = \frac{d_\mathcal{K} + \sqrt{d_\mathcal{K}}}{2}$, we can write

$$\mathcal{O}_\mathcal{K} = \mathbb{Z} \oplus \mathbb{Z} w_\mathcal{K}. \tag{2.1}$$

The discriminant $d_\mathcal{K}$ of the ring of integers of $\mathcal{K}$ is called *the discriminant of* $\mathcal{K}$.

Proposition 2.39. Let $\mathcal{K} = \mathbb{Q}(\sqrt{N})$ be a quadratic number field with N square-free. Then $\mathcal{O}_\mathcal{K}$ is the maximal order in $\mathcal{K}$.

Proof. Let $\mathcal{O} = \mathbb{Z}w_1 \oplus \mathbb{Z}w_2$ be an order inside $\mathcal{K}$. Consider an element $x \in \mathcal{O}$. We have that $xw_1 = aw_1 + bw_2$ and $xw_2 = cw_1 + dw_2$ for some $a, b, c, d \in \mathbb{Z}$. Hence, the matrix

$$A = \begin{pmatrix} x - a & -b \\ -c & x - d \end{pmatrix}$$

has a non trivial element $(w_1, w_2)^T$ in its kernel. We thus have $\det(A) = 0$, which gives a monic equation on x with integral coefficients. Hence, $x \in \mathcal{O}_\mathcal{K}$. We conclude that $\mathcal{O} \subseteq \mathcal{O}_\mathcal{K}$ □

The following proposition gives a description of all the orders inside a quadratic number field.

Proposition 2.40. Let $\mathcal{O}$ be an order in a quadratic number field $\mathcal{K}$. Then $\mathcal{O}$ has finite index in $\mathcal{O}_\mathcal{K}$. Moreover, let $f = [\mathcal{O}_\mathcal{K} : \mathcal{O}]$. We have

$$\mathcal{O} = \mathbb{Z} + f\mathcal{O}_\mathcal{K} = \mathbb{Z} \oplus \mathbb{Z} f w_\mathcal{K}$$

where $w_\mathcal{K}$ is as in equation (2.1).

Proof. We have that both $\mathcal{O}$ and $\mathcal{O}_\mathcal{K}$ are free abelian groups of rank 2. The quotient of two free abelian groups of the same rank is finite, hence $[\mathcal{O}_\mathcal{K} : \mathcal{O}] < \infty$. Writing $f = [\mathcal{O}_\mathcal{K} : \mathcal{O}]$ we have that $f\mathcal{O}_\mathcal{K} \subseteq \mathcal{O}$. It follows that $\mathbb{Z} + f\mathcal{O}_\mathcal{K} \subseteq \mathcal{O}$. Notice that $\mathbb{Z} + f\mathcal{O}_\mathcal{K}$ is the order $\mathbb{Z} \oplus \mathbb{Z} f w_\mathcal{K}$. This order clearly has index f in $\mathcal{O}_\mathcal{K} = \mathbb{Z} \oplus \mathbb{Z} w_\mathcal{K}$. We conclude that $\mathcal{O} = \mathbb{Z} + f\mathcal{O}_\mathcal{K}$. □

The index $f = [\mathcal{O}_\mathcal{K} : O]$ is called the *conductor of* $\mathcal{O}$. The previous proposition shows that an order $\mathcal{O}$ with conductor f can be written in the basis $1, f w_\mathcal{K}$. Using Definition 2.33 we get that the discriminant of $\mathcal{O}$ is the integer

$$D = f^2 d_\mathcal{K}. \tag{2.2}$$

2.2.1 Ideals and their norms

Let us now focus on the properties of invertible ideals of $\mathcal{O}$. We start by considering proper ideals since every invertible ideal is proper. In fact, we will show that in the context of quadratic orders a proper ideal is always invertible.

Lemma 2.41. *Let $\mathcal{K} = \mathbb{Q}(\tau)$ be a quadratic number field with τ a root of the polynomial $aX^2 + bX + c$ with a, b and c coprime integers. Then $I = \mathbb{Z} \oplus \mathbb{Z}\tau$ is a proper fractional ideal for the order $\mathcal{O} = \mathbb{Z} \oplus \mathbb{Z}a\tau$.*

Proof. We start by noting that $(a\tau)^2 + b(a\tau) + ac = 0$ so $a\tau$ satisfies a monic equation with integer coefficients, thus $\mathcal{O} = \mathbb{Z}[a\tau] = \mathbb{Z} \oplus \mathbb{Z}a\tau$ is indeed an order and I as a fractional ideal since aI is an integral $\mathcal{O}$-ideal, which is easy to verify.

Let $\beta \in \mathcal{K}$. Note that $\beta I \subseteq I$ if and only if $\beta \in I$ and $\beta\tau \in I$. This is the case if and only if there are $m, n \in \mathbb{Z}$ such that $\beta = m + n\tau$ and

$$\begin{aligned}\beta\tau &= m\tau + n\tau^2 \\ &= m\tau + \frac{n}{a}(-b\tau - c) \\ &= -\frac{cn}{a} + \left(m - \frac{bn}{a}\right)\tau \in I.\end{aligned}$$

Since $\gcd(a, b, c) = 1$, this last condition is true if and only if $a \mid n$. In consequence $\{\beta \in \mathcal{K} \mid \beta I \subseteq I\} = \mathbb{Z} + \mathbb{Z}a\tau = \mathcal{O}$. Which shows that I is a proper fractional $\mathcal{O}$-ideal. □

Proposition 2.42. Let $\mathcal{O}$ be an order inside a quadratic number field $\mathcal{K}$ and let I be a proper fractional $\mathcal{O}$-ideal. Then I is invertible.

Proof. Since $\mathcal{O}$ is a finitely generated $\mathbb{Z}$-module and I is a finitely generated $\mathcal{O}$-module (recall that a number ring is noetherian by Corollary 2.13), we have that I is a finitely generated $\mathbb{Z}$-module. Since I has finite index in $\mathcal{O}$, we must have that I is a free $\mathbb{Z}$-module of rank 2, hence we can write $I = \mathbb{Z}\alpha \oplus \mathbb{Z}\beta = \alpha(\mathbb{Z} \oplus \mathbb{Z}\tau)$, with $\tau = \beta/\alpha$.

It's easy to see that if I is proper, then $\mathbb{Z} \oplus \mathbb{Z}\tau$ is also a proper fractional ideal of the order $\mathcal{O}$. Let $aX^2 + bX + c$ with $\gcd(a, b, c) = 1$ be the minimal polynomial of τ. By Lemma 2.41 we have that $\mathbb{Z} \oplus \mathbb{Z}\tau$ is a proper ideal of the order $\mathbb{Z} \oplus \mathbb{Z}a\tau$. This implies that $\mathcal{O} = \mathbb{Z} \oplus \mathbb{Z}a\tau$. Let $x \mapsto \overline{x}$ be the non trivial automorphism of $\mathcal{K}$. Note that $\overline{\tau}$ is the other root of $aX^2 + bX + c$. Then, by Lemma 2.41, we have that $\mathbb{Z} \oplus \mathbb{Z}\overline{\tau}$ is a proper fractional ideal for the order $\mathbb{Z} \oplus \mathbb{Z}a\overline{\tau} = \mathbb{Z} \oplus \mathbb{Z}a\tau = \mathcal{O}$, hence $\overline{I} := \overline{\alpha}(\mathbb{Z} \oplus \mathbb{Z}\overline{\tau})$ is a fractional $\mathcal{O}$-ideal. Observe that

$$a I\overline{I} = a\alpha\overline{\alpha}(\mathbb{Z} \oplus \mathbb{Z}\tau)(\mathbb{Z} \oplus \mathbb{Z}\overline{\tau}) = N(\alpha)(\mathbb{Z}a + \mathbb{Z}a\tau + \mathbb{Z}a\overline{\tau} + \mathbb{Z}a\tau\overline{\tau}) = N(\alpha)\mathcal{O}.$$

The last equality follows from $a\tau + a\overline{\tau} = -b$, $a\tau\overline{\tau} = c$ and $\gcd(a, b, c) = 1$. So we have

$$I\overline{I} = \frac{N(\alpha)}{a}\mathcal{O}. \tag{2.3}$$

This shows that I is invertible. □

The proposition above states that in the context of orders in quadratic number fields, proper and invertible ideals are the same. Let us now prove some facts about the norm of an ideal.

Proposition 2.43. Let $\mathcal{O}$ be an order inside a quadratic number field. Let $\alpha \in \mathcal{O}$ with $\alpha \neq 0$. Then $N(\alpha\mathcal{O}) = |N(\alpha)|$.

Proof. From Proposition 2.40 we see that we can choose a $\mathbb{Z}$-basis for $\mathcal{O}$ that includes 1, say $\{1, w\}$ with some $w \in \mathcal{O}$. We have that $\alpha = a + bw$ and $\alpha w = c + dw$ with $a, b, c, d \in \mathbb{Z}$. Hence, the matrix

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

describes multiplication by α. Just as in the proof of Proposition 2.39, the matrix $A = \alpha\mathbb{I} - B^T$ is singular. Hence, $\det(A) = \alpha^2 - \mathrm{Tr}(B)\alpha + \det(B) = 0$. Since α is in the ring of integers of $\mathcal{K}$, it is either in $\mathbb{Z}$ or it has minimal polynomial of degree 2.

If α is in $\mathbb{Z}$ then $\alpha = a$, $b = 0$, $c = 0$ and $d = a$. Thus, $\det(B) = a^2 = N(\alpha)$. If α is not in $\mathbb{Z}$, then its minimal polynomial is $X^2 - T(\alpha)X + N(\alpha)$, which must be equal to $X^2 - \mathrm{Tr}(B)X + \det(B)$. Therefore, in either case we have that $N(\alpha) = \det(B)$. It follows that

$$\#\mathcal{O}/\alpha\mathcal{O} = \#\mathbb{Z}^2/B\mathbb{Z}^2 = |\det(B)| = |N(\alpha)|$$

□

Remark 2.44. If $\mathcal{O}$ is inside an imaginary quadratic number field, then $N(\alpha) \geq 0$ for all $\alpha \in \mathcal{O}$, so the absolute value in the above proposition may be omitted.

It turns out that when we restrict ourselves to invertible (proper) ideals, the norm has the very nice property of being multiplicative.

Proposition 2.45. Let $\mathcal{O}$ be an order inside an imaginary quadratic number field. Let I and J be proper $\mathcal{O}$-ideals. Then

$$N(IJ) = N(I)N(J).$$

Proof. We first prove the following: Let I be an $\mathcal{O}$-ideal (not necessarily proper). If $\alpha \in \mathcal{O}$ and $\alpha \neq 0$ then $N(\alpha I) = N(\alpha)N(I)$.

Note that we have the inclusions $\alpha I \subseteq \alpha\mathcal{O} \subseteq \mathcal{O}$ which give a short exact sequence

$$0 \to \alpha\mathcal{O}/\alpha I \to \mathcal{O}/\alpha I \to \mathcal{O}/\alpha\mathcal{O} \to 0.$$

It follows that $\#\mathcal{O}/\alpha\mathcal{O} = (\#\mathcal{O}/\alpha I)/(\#\alpha\mathcal{O}/\alpha I)$. Observe that multiplication by α induces an isomorphism $\mathcal{O}/I \to \alpha\mathcal{O}/\alpha I$ so that $\#\mathcal{O}/I = \#\alpha\mathcal{O}/\alpha I$. It follows that $N(\alpha\mathcal{O}) = N(\alpha I)/N(I)$. Using the previous proposition we conclude that $N(\alpha I) = N(\alpha)N(I)$.

Now, let I be a proper $\mathcal{O}$-ideal and write it in the form $I = \alpha(\mathbb{Z}\oplus\mathbb{Z}\tau)$. Let aX^2+bX+c with $\gcd(a,b,c) = 1$ be the minimal polynomial of τ. From Lemma 2.41 we get that $\mathcal{O} = \mathbb{Z}\oplus\mathbb{Z}a\tau$. The ideal $a\mathbb{Z}\oplus\mathbb{Z}a\tau = a(\mathbb{Z}\oplus\mathbb{Z}\tau)$ clearly has index a in $\mathcal{O}$, so $N(a(\mathbb{Z}\oplus\mathbb{Z}\tau)) = a$. We have $aI = \alpha a(\mathbb{Z} \oplus \mathbb{Z}\tau)$ and we get

$$a^2N(I) = N(a)N(I) = N(aI) = N(\alpha a(\mathbb{Z} \oplus \mathbb{Z}\tau)) = N(\alpha)N(a(\mathbb{Z} \oplus \mathbb{Z}\tau)) = N(\alpha)a$$

which implies that $N(I) = N(\alpha)/a$. Now, recall equation (2.3). It follows that

$$I\overline{I} = N(I)\mathcal{O}.$$

Note that the previous formula applies to any proper (invertible) $\mathcal{O}$-ideal. If J is another proper $\mathcal{O}$-ideal, then IJ is again a proper $\mathcal{O}$-ideal, so that

$$N(IJ)\mathcal{O} = (IJ)(\overline{IJ}) = I\overline{I}J\overline{J} = N(I)\mathcal{O}N(J)\mathcal{O} = N(I)N(J)\mathcal{O}.$$

this implies $N(IJ) = \mu N(I)N(J)$ with $\mu \in \mathcal{O}^*$, and since $N(IJ)$ and $N(I)N(J)$ are in $\mathbb{Z}_{>0}$ it must be the case that $\mu = 1$. □

Remark 2.46. One could also argue similarly to the proof of Theorem 2.32 (2) to conclude that if I and J can be written as the product of invertible prime ideals, then $N(IJ) = N(I)N(J)$. The previous proposition is more general as not every invertible ideal is the product of prime ideals. As an example, take the principal ideal $2\mathcal{O}$ in the order $\mathcal{O} = \mathbb{Z}[\sqrt{-11}]$.

Remark 2.47. The restriction to invertible (or proper) ideals in the previous proposition is necessary. Consider the prime ideal from Example 2.7 given by $\mathfrak{p} = (2, \sqrt{-11} - 1)$ for the order $\mathbb{Z}[\sqrt{-11}]$. We have that $\mathfrak{p}^2 = 2\mathfrak{p}$ and $N(\mathfrak{p}) = 2$. Therefore, $N(\mathfrak{p}^2) = N(2\mathfrak{p}) = N(2)N(\mathfrak{p}) = 8$ but $N(\mathfrak{p})^2 = 4$.

2.2.2 Integral quadratic forms

There is a connection between integral quadratic forms and proper ideals of an order in an imaginary quadratic number field. This connection will be convenient to derive an important property of the class group.

Definition 2.48. An *integral quadratic form in two variables* is a homogeneous polynomial in $\mathbb{Z}[X, Y]$ of degree 2. Thus, an integral quadratic form is of the form

$$f(X, Y) = aX^2 + bXY + cY^2$$

with $a, b, c \in \mathbb{Z}$.

The *discriminant* of a quadratic form is defined to be $D = b^2 - 4ac$.

Definition 2.49. We say that a quadratic form $f(X, Y)$ *represents* $m \in \mathbb{Z}$ if there are $x, y \in \mathbb{Z}$ such that $f(x, y) = m$. If x, y can be taken to be such that $\gcd(x, y) = 1$, then we say that $f(X, Y)$ *properly represents* m.

Lemma 2.50. *Let $f(X,Y) = aX^2 + bXY + cY^2$ be an integral quadratic form with $\gcd(a,b,c) = 1$, and let $M \in \mathbb{Z}$. Then $f(X,Y)$ represents an integer relatively prime to M.*

Proof. Let p be a prime number. Note that $f(1,0) = a$, $f(0,1) = c$ and $f(1,1) = a+b+c$ so at least one of these numbers is relatively prime to p, otherwise a, b and c would be divisible by p.

For each distinct prime p that divides M, let $(x_p, y_p) \in \mathbb{Z}^2$ be such that $f(x_p, y_p)$ is relatively prime to p. From the Chinese Remainder Theorem we know that there are $x, y \in \mathbb{Z}$ such that

$$\begin{aligned} x &\equiv x_p \pmod{p} \\ y &\equiv y_p \pmod{p} \end{aligned}$$

for every prime p that divides M. This implies that $f(x,y) \equiv f(x_p, y_p) \pmod{p}$ for every prime p dividing M. Since $f(x_p, y_p) \pmod{p} \neq 0$, we have that none of the primes dividing M divide $f(x,y)$. We conclude that $f(x,y)$ is relatively prime to M. □

Lemma 2.51. *Let $\mathcal{O}$ be an order of discriminant D in an imaginary quadratic number field $\mathcal{K}$. If $f(X,Y) = aX^2 + bXY + cY^2$ is a integral quadratic form with $a > 0$ and $\gcd(a,b,c) = 1$ of discriminant D, then $\mathbb{Z}a \oplus \mathbb{Z}a\tau$ with $\tau = (-b + \sqrt{D})/2a$ is a proper ideal of $\mathcal{O}$.*

Proof. Since $\mathcal{K}$ is imaginary, we have that $D < 0$. Note that $\tau = (-b + \sqrt{D})/2a$ is the unique root of $f(X,1) = aX^2 + bX + c$ in the upper half plane (i.e. with $\mathrm{im}(\tau) > 0$). From Lemma 2.41 it follows that $\mathbb{Z} \oplus \mathbb{Z}\tau$ is a proper fractional ideal for the order $\mathbb{Z}[a\tau]$. Thus, $\mathbb{Z}a \oplus \mathbb{Z}a\tau$ is a proper integral ideal of $\mathbb{Z}[a\tau]$. We claim that $\mathbb{Z}[a\tau] = \mathcal{O}$.

Let f be the conductor of $\mathcal{O}$. We have that $D = f^2 d_{\mathcal{K}}$ from (2.2). Note that

$$a\tau = \frac{-b + \sqrt{D}}{2} = \frac{-b + f\sqrt{d_{\mathcal{K}}}}{2} = -\frac{b + fd_{\mathcal{K}}}{2} + f\left(\frac{d_{\mathcal{K}} + \sqrt{d_{\mathcal{K}}}}{2}\right) = -\frac{b + fd_{\mathcal{K}}}{2} + fw_{\mathcal{K}}$$

with $w_{\mathcal{K}}$ as in (2.1). Note that $D = b^2 - 4ac$ and b have the same parity. In consequence, $fd_{\mathcal{K}}$ and b have the same parity, so $(b + fd_{\mathcal{K}})/2$ is an integer. We conclude that $\mathbb{Z}[a\tau] = \mathbb{Z} \oplus \mathbb{Z}fw_{\mathcal{K}}$ which is equal to $\mathcal{O}$ by Proposition 2.40. □

Theorem 2.52. *Let $\mathcal{O}$ be an order in an imaginary quadratic number field $\mathcal{K}$. Let I be an invertible $\mathcal{O}$-ideal and $M \in \mathbb{Z}$. Then there is some $\mathcal{O}$-ideal J in the same class of I in $\mathrm{Cl}(\mathcal{O})$ with $N(J)$ relatively prime to M.*

Proof. Let D be the discriminant of $\mathcal{O}$. Start by writing the invertible ideal I (which is a free $\mathbb{Z}$-module of rank 2) in the form $I = \alpha(\mathbb{Z} \oplus \mathbb{Z}\tau)$ with $\alpha \in \mathcal{O}$ and $\tau \in \mathcal{K}$ with $\mathrm{im}(\tau) > 0$. Then τ is the root of some polynomial $aX^2 + bX + c$ with integer coefficients and $\gcd(a,b,c) = 1$. We know that $\mathbb{Z} \oplus \mathbb{Z}\tau$ is a proper fractional $\mathcal{O}$-ideal, and from Lemma 2.41 we get that $\mathbb{Z} \oplus \mathbb{Z}\tau$ is a proper fractional $\mathbb{Z}[a\tau]$-ideal, hence $\mathcal{O} = \mathbb{Z}[a\tau]$. The discriminant of the order $\mathbb{Z}[a\tau]$ is easily computed to be $b^2 - 4ac$, so $D = b^2 - 4ac$.

Consider the quadratic form $f(X,Y) = aX^2 + bXY + cY^2$ which has discriminant D. From Lemma 2.50 we know that there are $x, y \in \mathbb{Z}$ such that $f(x,y)$ is relatively prime to M. Define $m = f(x,y)$ and let $d = \gcd(x,y)$. We have that

$$a(x/d)^2 + b(x/d)(y/d) + c(y/d)^2 = m/d^2$$

where $\gcd(x/d, y/d) = 1$. Therefore, $f(X,Y)$ properly represents $a' = m/d^2$. Let $p = x/d$ and $q = y/d$, so $f(p,q) = a'$. Since p, q are relatively prime, there are integers r, s such that $ps - qr = 1$. Define the quadratic form

$$\begin{aligned} g(X,Y) :&= f(pX + rY, qX + sY) \\ &= f(p,q)X^2 + (2apr + bps + brq + 2cqs)XY + f(r,s)Y^2 \\ &= a'X^2 + b'XY + c'Y^2, \end{aligned}$$

with $b' = 2apr + bps + brq + 2cqs$ and $c' = f(r,s)$. A direct computation shows that the discriminant of $g(X,Y)$ is again D. Let τ' be the unique root of $g(X,1)$ with $\operatorname{im}(\tau') > 0$. By Lemma 2.51, the $\mathbb{Z}$-module $J' = \mathbb{Z}a' \oplus \mathbb{Z}a'\tau'$ is a proper ideal of $\mathcal{O}$. Furthermore, from the proof of this lemma, we also have $\mathcal{O} = \mathbb{Z} \oplus \mathbb{Z}a'\tau'$. Thus, $N(J') = a'$.

Note that

$$0 = g(\tau', 1) = f(p\tau' + r, q\tau' + s) = (q\tau' + s)^2 f\left(\frac{p\tau' + r}{q\tau' + s}, 1\right)$$

so $\frac{p\tau'+r}{q\tau'+s}$ is a root of $f(X,1)$. An easy calculation shows that $\operatorname{im}(\frac{p\tau'+r}{q\tau'+s}) = |q\tau'+s|^{-1} \operatorname{im}(\tau)$. Hence, $\operatorname{im}(\frac{p\tau'+r}{q\tau'+s}) > 0$, which implies that

$$\tau = \frac{p\tau' + r}{q\tau' + s}.$$

Let $\lambda = q\tau' + s$. We have

$$\lambda(\mathbb{Z} \oplus \mathbb{Z}\tau) = \mathbb{Z}(q\tau' + s) \oplus \mathbb{Z}(p\tau' + r) = \mathbb{Z} \oplus \mathbb{Z}\tau'$$

since the matrix $\begin{pmatrix} p & q \\ r & s \end{pmatrix}$ is in $GL_2(\mathbb{Z})$. Multiplying by $\alpha a'$ the equation above we get that $\lambda a' I = \alpha J'$. It follows that I and J' are in the same class in $\mathrm{Cl}(\mathcal{O})$. Recall that $N(J') = a'$. Hence, if we define $J = dJ'$, we get

$$N(J) = N(d)N(J') = d^2 a' = m.$$

The ideal J is in the same class as I and has norm m which is relatively prime to M. This concludes the proof. □

Corollary 2.53. Let $\mathcal{O}$ be an order in an imaginary quadratic number field $\mathcal{K}$, $I \subseteq \mathcal{O}$ an invertible ideal and $M \in \mathbb{Z}$. There exists $\alpha \in I$ such that $N(\alpha I^{-1})$ is relatively prime to M.

Proof. Consider the class of the fractional ideal I^{-1}. From the previous theorem there is J an integral $\mathcal{O}$-ideal in the class of I^{-1} such that $N(J)$ is relatively prime to M. Since I^{-1} and J are in the same class, we have that there is $\alpha \in \mathcal{K}$ such that $J = \alpha I^{-1}$. Since $J \subseteq \mathcal{O}$, we conclude that $\alpha \in I$. □

Corollary 2.54. Let $\mathcal{O}$ be an order in an imaginary quadratic number field. If I is an invertible $\mathcal{O}$-ideal, then

$$N(I) = \gcd\{N(\alpha) \mid \alpha \in I\}.$$

Proof. Let $d = \gcd\{N(\alpha) \mid \alpha \in I\}$. For any $\alpha \in I$, we have that $\#\mathcal{O}/\alpha\mathcal{O} = (\#\mathcal{O}/I)(\#I/\alpha\mathcal{O})$ so $N(I) \mid N(\alpha)$. This implies that $N(I) \mid d$.

To show $d \mid N(I)$ take $\alpha \in I$ to be such that $N(\alpha I^{-1})$ is relatively prime to d, which we can do according to Corollary 2.53. Note that

$$N(\alpha) = N(\alpha\mathcal{O}) = N(\alpha I^{-1} I) = N(\alpha I^{-1})N(I),$$

We have that $d \mid N(\alpha)$ and $N(\alpha I^{-1})$ is relatively prime to d, so $d \mid N(I)$. □

There is a simple property that guarantees that an ideal is invertible. Given an order $\mathcal{O}$ with conductor f, we say that an ideal I is *prime to the conductor* f if I and $f\mathcal{O}$ are coprime ideals, that is, if $I + f\mathcal{O} = \mathcal{O}$.

Proposition 2.55. Let $\mathcal{O}$ be an order in a quadratic number field of conductor f. Let I be an $\mathcal{O}$-ideal. Then I is prime to f if and only if $N(I)$ is relatively prime to f. Furthermore, every $\mathcal{O}$-ideal prime to f is invertible.

Proof. For the first part consider the homomorphism $m_f : \mathcal{O}/I \to \mathcal{O}/I$ given by multiplication by f. Note that since $\mathcal{O}/I$ is a finite abelian group, we get from the structure theorem that m_f is an isomorphism if and only if f is relatively prime to the order of $\mathcal{O}/I$ (i.e. relatively prime to $N(I)$). Thus, we have

$$\begin{aligned} I + f\mathcal{O} = \mathcal{O} \iff m_f \text{ is surjective} &\iff m_f \text{ is an isomorphism} \\ &\iff N(I) \text{ and } f \text{ are relatively prime.} \end{aligned}$$

For the second part, suppose I is prime to f. We will show that I is proper and hence invertible by Proposition 2.42. Let $\beta \in \mathcal{K}$ with $\beta I \subseteq I$. Since I is a $\mathbb{Z}$-module of rank 2, it can be shown with a similar argument to the proof of Proposition 2.39 that $\beta \in \mathcal{O}_\mathcal{K}$. We have

$$\beta\mathcal{O} = \beta(I + f\mathcal{O}) = \beta I + \beta f\mathcal{O} \subseteq I + f\mathcal{O}_\mathcal{K} \subseteq I + \mathcal{O} \subseteq \mathcal{O}$$

since $f\mathcal{O}_\mathcal{K} \subseteq \mathcal{O}$. Therefore $\beta \in \mathcal{O}$, which shows I is proper and hence, invertible. □

2.2.3 The size of the class group

Given an order $\mathcal{O}$ in a number field $\mathcal{K}$, we denote by $h(\mathcal{O})$ the cardinality of $\mathrm{Cl}(\mathcal{O})$. The number $h(\mathcal{O}_\mathcal{K})$ is referred to as the *class number* of the number field $\mathcal{K}$.

Let us state a couple of facts about $h(\mathcal{O})$ and $h(\mathcal{O}_\mathcal{K})$. These are not necessarily required for the mathematical foundations of CSIDH, but they do provide a justification for the security of the scheme and for the choice of parameters in the implementation, as they give us an estimate of the size of the class group.

Proposition 2.56. Let $\mathcal{O}$ be an order in an imaginary quadratic number field $\mathcal{K}$. Then $h(\mathcal{O})$ is an integer multiple of $h(\mathcal{O}_\mathcal{K})$.

Proof. See [5, Theorem 7.24], which states a precise formula for $h(\mathcal{O})$ in terms of its conductor f and $h(\mathcal{O}_\mathcal{K})$. □

Given the discriminant $d_\mathcal{K}$ of an imaginary quadratic number field, we can define $h(d_\mathcal{K})$ to be the size of the class group of $\mathcal{O}_\mathcal{K}$, where $\mathcal{K}$ is the only quadratic number field with discriminant $d_\mathcal{K}$. It's natural to wonder about the behavior of $h(d_\mathcal{K})$ as $d_\mathcal{K} \to -\infty$. In 1935, Siegel [17] proved that

$$\lim_{d_\mathcal{K}\to-\infty} \frac{\log h(d_\mathcal{K})}{\log \sqrt{|d_\mathcal{K}|}} = 1.$$

This, together with Proposition 2.56, implies that if $\varepsilon > 0$ and $\mathcal{O}$ is an order inside an imaginary quadratic number field $\mathcal{K}$ with big enough discriminant (in absolute value), then

$$h(\mathcal{O}) > |d_\mathcal{K}|^{1/2-\varepsilon}.$$

In practice, this means that if we are dealing with an order $\mathcal{O}$ of big absolute discriminant, it's reasonable to assume that $h(\mathcal{O})$ is approximately as big as $\sqrt{|d_\mathcal{K}|}$.

3 Endomorphism ring of supersingular elliptic curves over $\mathbb{F}_p$

Let E, E' be elliptic curves over K. Recall that the elements of $\mathrm{Hom}(E, E')$ are the isogenies from E to E' together with the zero morphism (denoted by $[0]$ or, simply, 0), and that $\mathrm{Hom}(E, E')$ is a group under addition. In this chapter we are interested in the case where $E' = E$. Our main references for this chapter are [18, Section III.4 and III.9], Chapter 4 of [25], and [21, Lecture 13 and 14].

Definition 3.1. Let E be an elliptic curve over K. We define the *ring of endomorphisms of* E to be the additive group $\mathrm{End}(E) := \mathrm{Hom}(E, E)$, with multiplication given by composition $\circ$.

We can extend the notion of degree to the zero morphism by defining $\deg(0) := 0$. We have $\deg(\phi \circ \psi) = \deg(\phi)\deg(\psi)$ for every $\phi, \psi \in \mathrm{End}(E)$, and $\deg(\phi) = 0$ if and only if $\phi = 0$. We also define the dual $\hat{0} := 0$.

Proposition 3.2. Let E be an elliptic curve over K. The endomorphism ring $\mathrm{End}(E)$ is a ring of characteristic 0 with no zero divisors.

Proof. We first show $\mathrm{End}(E)$ is torsion-free. Let $m \in \mathbb{Z}$ and $\phi \in \mathrm{End}(E)$. Note that $m\phi = [m] \circ \phi$. Suppose that $[m] \circ \phi = 0$. Taking degrees we have

$$m^2 \deg(\phi) = \deg([m]) \deg(\phi) = \deg(0) = 0.$$

thus either $m = 0$ or $\phi = 0$. To show it has no zero divisors, let $\phi, \psi \in \mathrm{End}(E)$. Suppose that $\phi \circ \psi = 0$. Taking degrees we have

$$\deg(\phi)\deg(\psi) = \deg(0) = 0.$$

thus either $\phi = 0$ or $\psi = 0$. □

To simplify the notation, let us write $\phi\psi$ instead of $\phi \circ \psi$, and for $m \in \mathbb{Z}$ we'll simply write m instead of $[m]$. The map $\mathrm{End}(E) \to \mathrm{End}(E)$ given by $\phi \mapsto \hat{\phi}$ satisfies $\widehat{\phi\psi} = \hat{\psi}\hat{\phi}$, $\widehat{\phi + \psi} = \hat{\phi} + \hat{\psi}$ and $\hat{\hat{\phi}} = \phi$ by Proposition 1.89. Thus, it is a *involution* on $\mathrm{End}(E)$. Also note that since $\phi\hat{\phi} = \deg(\phi)$ then $\phi\hat{\phi}$ is always a non-negative integer and $\phi\hat{\phi} = 0$ if and only if $\phi = 0$. Observe that for every $\phi \in \mathrm{End}(E)$ we have

$$\deg(1 - \phi) = \widehat{(1 - \phi)}(1 - \phi) = (1 - \hat{\phi})(1 - \phi) = 1 - \phi - \hat{\phi} + \deg(\phi),$$

so that $\phi + \hat{\phi} = 1 + \deg(\phi) - \deg(1 - \phi)$ is an integer. We make the following definition.

Definition 3.3. Let $\phi \in \mathrm{End}(E)$ be an endomorphism. The *trace of* ϕ is the integer

$$\mathrm{T}(\phi) = \phi + \hat{\phi}$$

Definition 3.4. Let $\phi \in \mathrm{End}(E)$ be an endomorphism. The *characteristic polynomial of* ϕ is given by

$$X^2 - \mathrm{T}(\phi)X + \deg(\phi)$$

It's straightforward to check that ϕ is a root of this polynomial. Note that $\hat{\phi}$ has the same characteristic polynomial as ϕ.

In the case of elliptic curves over finite fields, there is a very important endomorphism.

Definition 3.5. Let $q = p^n$ be a power of a prime p and consider an elliptic curve E over $\mathbb{F}_q$. The endomorphism

$$\pi_E : E \to E, \quad (x, y) \mapsto (x^q, y^q)$$

is called the *q-power Frobenius endomorphism.* Also observe that $\pi_E = \pi^n$ where π is the p-power Frobenius morphism given in Definition 1.60. Note that π_E is purely inseparable and $\deg(\pi_E) = \deg(\pi)^n = p^n = q$.

Remark 3.6. Consider an elliptic curve $E/\mathbb{F}_q$ and an endomorphism $\phi \in \mathrm{End}(E)$. Let $n \in \mathbb{Z}$. Since ϕ is a group homomorphism, we have that $n\phi = \phi n$. Thus, integers commute with all the elements in $\mathrm{End}(E)$. The rational functions defining ϕ have coefficients in $\mathbb{F}_q$, so we also have that $\pi_E\phi = \phi\pi_E$. Therefore, the subring $\mathbb{Z}[\pi_E]$ is contained in the center of $\mathrm{End}(E)$.

3.1 Supersingular elliptic curves

Among all the elliptic curves defined over a field K of characteristic $p > 0$, there is a special and rare kind called *supersingular elliptic curves.* Recall from Theorem 1.86 that there are only two possibilities for $E[p]$, namely $\{O\}$ and $\mathbb{Z}/p\mathbb{Z}$.

Definition 3.7. Let K be a field of characteristic $p > 0$ and let E be an elliptic curve over K. We say that the curve E is *supersingular* if $E[p] = \{O\}$, and we say it's *ordinary* if $E[p] \cong \mathbb{Z}/p\mathbb{Z}$.

There are many alternative characterizations of supersingularity. In the case of finite fields, we can give one of them in terms of the trace of the q-power Frobenius endomorphism.

Theorem 3.8. *Let E be an elliptic curve over the finite field $\mathbb{F}_q$, with $q = p^n$ a prime power. Then E is supersingular if and only if* $\mathrm{T}(\pi_E) \equiv 0 \mod p$.

Proof. Suppose E is supersingular. By Theorem 1.86 we have that $E[p^e] = \{O\}$ for every $e \geq 1$. Hence, $\ker(\widehat{\pi_E}\pi_E) = E[p^n] = \{O\}$. This implies that $\ker(\widehat{\pi_E})$ is trivial, and since $\deg(\widehat{\pi_E}) = p^n$ we have that $\widehat{\pi_E}$ is inseparable. We know that π_E is also inseparable. It follows from Remark 1.71 that their sum $\mathrm{T}(\pi_E) = \pi_E + \widehat{\pi_E}$ is either zero or an inseparable endomorphism. But $\mathrm{Tr}(\pi_E) \in \mathbb{Z}$, so, by Proposition 1.83, we conclude that p divides $\mathrm{Tr}(\pi_E)$.

Conversely, if $p \mid \mathrm{T}(\pi_E)$ then $\mathrm{T}(\pi_E)$ is either zero or inseparable. This implies that $\widehat{\pi_E} = \mathrm{Tr}(\pi_E) - \pi_E$ is inseparable. Thus, $\#\ker(\widehat{\pi_E})$ is strictly less than $\deg(\widehat{\pi_E}) = p^n$. Since π_E has trivial kernel, we have that $\#E[p^n] = \#\ker(\pi_E\widehat{\pi_E}) = \#\ker(\widehat{\pi_E}) < p^n$. From Theorem 1.86 we see that the only possibility is $E[p^n] = \{O\}$, which implies $E[p] = \{O\}$.

□

The characterization of supersingular elliptic curves is even simpler over the field $\mathbb{F}_p$ with $p \geq 5$ a prime number. In order to show this, we'll need Hasse's theorem.

Lemma 3.9. *Let E be an elliptic curve over a finite field $\mathbb{F}_q$ and let $t = \mathrm{T}(\pi_E)$. Consider $r, s \in \mathbb{Z}$. Then $\deg(r\pi_E - s) = r^2q + s^2 - rst$.*

Proof. This is a straightforward computation using the fact that $\deg(r\pi_E - s) = \widehat{(r\pi_E - s)}(r\pi_E - s) = (r\widehat{\pi_E} - s)(r\pi_E - s)$, and $\widehat{\pi_E}\pi_E = q$. □

Lemma 3.10. *Let E be an elliptic curve over $\mathbb{F}_q$. Then the endomorphism $\pi_E - 1$ is separable.*

Proof. Suppose $\pi_E - 1$ is inseparable. Then $-(\pi_E - 1) = 1 - \pi_E$ is also inseparable. The sum of two inseparable endomorphisms $\pi_E + (1 - \pi_E) = 1$ is either zero or inseparable, but 1 is separable and different from zero. A contradiction. □

Lemma 3.11. *Let E be an elliptic curve over $\mathbb{F}_q$ and $t = \mathrm{T}(\pi_E)$. Then*

$$\#E(\mathbb{F}_q) = q + 1 - t$$

Proof. Note that $P \in E(\mathbb{F}_q)$ if and only if $\pi_E(P) = P$, that is if and only if $P \in \ker(\pi_E - 1)$. Since $\pi_E - 1$ is separable by Lemma 3.10, we have that

$$\#E(\mathbb{F}_p) = \#\ker(\pi_E - 1) = \deg(\pi_E - 1) = q + 1 - t,$$

where we used Lemma 3.9 with $r = s = 1$. □

Theorem 3.12 (Hasse's Theorem)**.** *Let E be an elliptic curve over $\mathbb{F}_q$. Then*

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

Proof. By Lemma 3.11 this is equivalent to showing that $|t| \leq 2\sqrt{q}$ with $t = \mathrm{T}(\pi_E)$. The degree of an endomorphism is always non-negative, thus $\deg(r\pi_E - s) \geq 0$ for all $r, s \in \mathbb{Z}$. Lemma 3.9 implies

$$q(r/s)^2 - t(r/s) + 1 \geq 0$$

for all $r, s \in \mathbb{Z}$ with $s \neq 0$. Thus $qx^2 - tx + 1 \geq 0$ for all $x \in \mathbb{Q}$, and since $\mathbb{Q}$ is dense in $\mathbb{R}$ this implies that $qx^2 - tx + 1 \geq 0$ for all $x \in \mathbb{R}$. It follows that the discriminant of the polynomial $qX^2 - tX + 1$ is negative or zero. Hence,

$$t^2 - 4q \leq 0$$

from which the desired inequality follows. □

For cryptographic purposes, one of the reasons supersingular elliptic curves over a prime field $\mathbb{F}_p$ are of interest is that they allow us to get a hold of the number of rational points that they have, as the following corollary shows.

Corollary 3.13. Let E be an elliptic curve over $\mathbb{F}_p$ with $p \geq 5$ a prime number. Then E is supersingular if and only if $\#E(\mathbb{F}_p) = p + 1$.

Proof. By Theorem 3.8, the curve E is supersingular if and only if $\mathrm{T}(\pi) \equiv 0 \mod p$. But $|\mathrm{T}(\pi)| \leq 2\sqrt{p}$ by Hasse's Theorem, and $2\sqrt{p} < p$ since $p \geq 5$. Thus, we have $p \mid \mathrm{T}(\pi)$ if and only if $\mathrm{T}(\pi) = 0$, which is equivalent to $\#E(\mathbb{F}_p) = p + 1$ by Lemma 3.11. □

3.2 The endomorphism algebra

In order to understand the endomorphism ring $\mathrm{End}(E)$ of an elliptic curve, it is useful to introduce the *endomorphism algebra*. We start by noting that $\mathrm{End}(E)$ and $\mathbb{Q}$ are both $\mathbb{Z}$-algebras.

Definition 3.14. Let E be an elliptic curve over K. The *endomorphism algebra of E* is the $\mathbb{Q}$-algebra given by

$$\mathrm{End}^0(E) = \mathrm{End}(E) \otimes_{\mathbb{Z}} \mathbb{Q}.$$

Note that every element of $\mathrm{End}^0(E)$ has an integer multiple in $\mathrm{End}(E)$. Since $\mathrm{End}(E)$ has no zero divisors, it follows that $\mathrm{End}^0(E)$ has no zero divisors. For $r \in \mathbb{Q}$ and $\phi \in \mathrm{End}(E)$ we write $r\phi$ instead of $\phi \otimes r$.

Since we will be working with supersingular elliptic curves over $\mathbb{F}_p$, we would like to have a description of their endomorphism algebra $\mathrm{End}^0(E)$. In order to do this, we first extend the involution $\phi \mapsto \hat{\phi}$ in $\mathrm{End}(E)$ to an involution in the endomorphism algebra.

Definition 3.15. Let E be an elliptic curve over K. Note that every element in $\mathrm{End}^0(E)$ can be written as $r\phi$ with $r \in \mathbb{Q}$ and $\phi \in \mathrm{End}(E)$. We define the *Rosati involution* on $\mathrm{End}^0(E)$ as

$$\widehat{r\phi} = r\hat{\phi}$$

where $\hat{\phi}$ is the dual to ϕ.

It's easy to verify that this is well defined, $\mathbb{Q}$-linear, and that for every $\phi, \psi \in \mathrm{End}^0(E)$ we have $\hat{\hat{\phi}} = \phi$ and $\widehat{\phi\psi} = \hat{\psi}\hat{\phi}$, so it is indeed an involution. Given this Rosati involution, we can extend the notions of trace and degree of endomorphisms to all elements of $\mathrm{End}^0(E)$ by defining the *reduced trace* and *reduced norm*.

Definition 3.16. Consider $\phi \in \mathrm{End}^0(E)$. The *reduced norm* and *reduced trace* of ϕ are defined as

$$N(\phi) = \phi\hat{\phi}, \qquad T(\phi) = \phi + \hat{\phi}$$

respectively.

Remark 3.17. Since the trace and the degree of endomorphisms are integers, we have that $N(\phi)$ and $T(\phi)$ are both in $\mathbb{Q}$. It's also easy to see that the norm is multiplicative and that the trace is $\mathbb{Q}$-linear. Additionally, $N(\phi) \geq 0$ for every ϕ, and $N(\phi) = 0$ if and only if $\phi = 0$.

Remark 3.18. Just as in the case of endomorphisms, an element $\phi \in \mathrm{End}^0(E)$ satisfies the characteristic polynomial $X^2 - T(\phi)X + N(\phi)$.

Remark 3.19. If $\mathrm{End}^0(E)$ is a quadratic number field, then the value of the norm and trace for elements in the endomorphism algebra in Definition 3.16 coincides with the norm and trace defined for elements in quadratic number fields in Definition 2.35, so there's no ambiguity when using the same notation. To see why this is the case, note that the Rosati involution $\phi \mapsto \hat{\phi}$ is the unique non-trivial automorphism fixing $\mathbb{Q}$, so it must coincide with conjugation.

Proposition 3.20. The endomorphism algebra $\mathrm{End}^0(E)$ is a division algebra.

Proof. Let $\phi \in \mathrm{End}^0(E)$ be a non-zero element. Since $\phi \neq 0$ we have that $N(\phi) \neq 0$. Consider $\psi = \frac{1}{N(\phi)}\hat{\phi}$. Note that $\phi\psi = \psi\phi = 1$. □

We are now ready to give a description of the endomorphism algebra of elliptic curves over a finite field when the Frobenius endomorphism is not an integer. It turns out that, in this case, the endomorphism algebra is a quadratic number field.

Theorem 3.21. *Let E be an elliptic curve defined over a finite field $\mathbb{F}_q$. Let $t = T(\pi_E)$. If $\pi_E \notin \mathbb{Z}$, then*

$$\mathrm{End}^0(E) = \mathbb{Q}(\pi_E) \cong \mathbb{Q}(\sqrt{t^2 - 4q}).$$

Proof. Since $\pi_E \notin \mathbb{Z}$ and π_E satisfies the characteristic polynomial $X^2 - tX + q$, we have that $\mathbb{Q}(\pi_E) \subseteq \mathrm{End}^0(E)$ is a quadratic number field isomorphic to $\mathbb{Q}(\sqrt{t^2 - 4q})$.

Recall that $\mathrm{End}(E)$ consists of all the endomorphisms of E defined over $\mathbb{F}_q$, so that π_E is in the center of $\mathrm{End}(E)$. It follows that $\mathbb{Q}(\pi_E)$ is contained in the center of $\mathrm{End}^0(E)$. In other words, any element of $\mathbb{Q}(\pi_E)$ commutes with every element of $\mathrm{End}^0(E)$.

Now, let us assume $\mathrm{End}^0(E) \neq \mathbb{Q}(\pi_E)$ to arrive at a contradiction. Let

$$\alpha = \pi_E - \frac{1}{2}t$$

we have that $\mathbb{Q}(\pi_E) = \mathbb{Q}(\alpha)$. Note that $T(\alpha) = T(\pi_E) - (1/2)(2t) = 0$. So α satisfies the characteristic equation $\alpha^2 + N(\alpha) = 0$. Since $\pi_E \notin \mathbb{Q}$ we have that $\alpha \neq 0$, hence $\alpha^2 = -N(\alpha) \in \mathbb{Q}^*$. Since $\mathrm{End}^0(E) \neq \mathbb{Q}(\alpha)$ we can choose $\beta' \in \mathrm{End}^0(E)$ such that $\beta' \notin \mathbb{Q}(\alpha)$. Consider the element

$$\beta = \beta' - \frac{1}{2}T(\beta') - \frac{T(\alpha\beta')}{2\alpha^2}\alpha$$

Note that $\beta' \notin \mathbb{Q}(\alpha)$ implies that $\beta \neq 0$. Using the fact that $T(\alpha) = 0$, a quick calculation shows that $T(\beta) = 0$ and $T(\alpha\beta) = 0$. Therefore, we have

$$\alpha = -\hat{\alpha}, \;\; \beta = -\hat{\beta} \text{ and } \alpha\beta = -\hat{\beta}\hat{\alpha}.$$

After substituting the first two equations in the third one, we find that

$$\alpha\beta = -\beta\alpha$$

But we know that α is in the center of $\mathrm{End}^0(E)$, so $\alpha\beta = \beta\alpha$. This, together with the previous equation, implies that $2\alpha\beta = 0$ which is a contradiction since $\alpha \neq 0$ and $\beta \neq 0$. We conclude that $\mathrm{End}^0(E) = \mathbb{Q}(\pi_E)$. □

Remark 3.22. By Hasse's Theorem $t^2 - 4q < 0$, so the previous theorem states that if $\pi_E \notin \mathbb{Z}$, then $\mathrm{End}^0(E)$ is an imaginary quadratic number field.

Remark 3.23. Theorem 3.21 doesn't tell the whole story of the endomorphism algebra since the condition $\pi_E \notin \mathbb{Z}$ is required. It can be shown that for an ordinary elliptic curve, the Frobenius is never an integer (see [21, Theorem 14.5]) so Theorem 3.21 applies. In the case of a supersingular elliptic curve, it can happen that $\pi_E \in \mathbb{Z}$. In such event, the endomorphism algebra is non-commutative (see [12, Proposition 2.60]). In fact, if E is supersingular, then $\mathrm{End}^0(E_{\bar{K}})$ is a quaternion algebra [18, V. Theorem 3.1(iii)]. Luckily, we can apply Theorem 3.21 to supersingular elliptic curves over $\mathbb{F}_p$ with $p \geq 5$ prime, which is all we really need.

We are now ready to give a description of the endomorphism ring of an elliptic curve in the case of Theorem 3.21.

Corollary 3.24. Let E be an elliptic curve defined over a finite field $\mathbb{F}_q$. If $\pi_E \notin \mathbb{Z}$, then $\mathrm{End}(E)$ is an order inside an imaginary quadratic number field.

Proof. By Theorem 3.21, the endomorphism algebra $\mathrm{End}^0(E)$ is an imaginary quadratic number field. We'll show that the subring $\mathrm{End}(E) \subseteq \mathrm{End}^0(E)$ is a free $\mathbb{Z}$-module of rank 2.

Let $\{e_1, e_2\}$ be a $\mathbb{Q}$-basis for $\mathrm{End}^0(E)$ such that $T(e_1e_2) = 0$. Multiplying by suitable integers we may assume that $e_1, e_2 \in \mathrm{End}(E)$ (for example, take $e_1 = 1$ and $e_2 = 2\alpha$ with α defined as in the proof of Theorem 3.21). Given a $\mathbb{Z}$-module $A \subseteq \mathrm{End}^0(E)$ we can define the *dual* $\mathbb{Z}$-module by

$$A^* = \{\alpha \in \mathrm{End}^0(E) \mid T(\alpha\phi) \in \mathbb{Z} \textit{ for all } \phi \in A\}$$

It's easy to see that A^* is indeed a $\mathbb{Z}$-module.

Let B be the free $\mathbb{Z}$-module spanned by $\{e_1, e_2\}$. We have that $B \subseteq \mathrm{End}(E)$. Taking duals we get $\mathrm{End}(E)^* \subseteq B^*$. Also note that $\mathrm{End}(E) \subseteq \mathrm{End}(E)^*$, thus we have

$$B \subseteq \mathrm{End}(E) \subseteq B^*.$$

Let us show that B^* is free $\mathbb{Z}$-module of rank 2. Consider $\phi \in B^*$. Since $\{e_1, e_2\}$ is a $\mathbb{Q}$-basis for $\mathrm{End}^0(E)$ we can write $\phi = a_1 e_1 + a_2 e_2$ with $a_1, a_2 \in \mathbb{Q}$. Multiplying by e_1 and taking trace on both sides, we have $T(\phi e_1) = a_1 T(e_1 e_1) + a_2 T(e_2 e_1) = a_1 T(e_1^2)$. Similarly, $T(\phi e_2) = a_2 T(e_2^2)$.

Note that $T(\phi e_1), T(\phi e_2) \in \mathbb{Z}$ since $\phi \in B^*$ and $e_1, e_2 \in B$. We have

$$\phi = T(\phi e_1)\frac{e_1}{T(e_1^2)} + T(\phi e_2)\frac{e_2}{T(e_2^2)}$$

which shows that $\{\frac{e_1}{T(e_1^2)}, \frac{e_2}{T(e_2^2)}\}$ is a $\mathbb{Z}$-basis for B^*. We conclude that B and B^* are both free $\mathbb{Z}$-modules of rank 2. Hence, $\mathrm{End}(E)$ is also a free $\mathbb{Z}$-module of rank 2 which concludes the proof. □

Corollary 3.25. Let E be a supersingular elliptic curve over $\mathbb{F}_p$ with $p \geq 5$ a prime number. Then $\mathrm{End}(E)$ is an order in the imaginary quadratic number field

$$\mathrm{End}^0(E) = \mathbb{Q}(\pi_E) \cong \mathbb{Q}(\sqrt{-p}).$$

Proof. From Corollary 3.13 we have that $T(\pi_E) = p + 1 - \#E(\mathbb{F}_p) = 0$. Also recall that $\deg(\pi_E) = p$, so π_E is a root of the polynomial $X^2 + p = 0$. In other words, $\pi_E^2 = -p$, which implies that $\pi_E \notin \mathbb{Z}$. Thus, Theorem 3.21 and Corollary 3.24 apply. □

3.3 Isogeny invariants

Let E_1 and E_2 be elliptic curves over K. We say that E_1 and E_2 are *K-isogenous* (or, simply, *isogenous*) if there exists an isogeny defined over K from E_1 to E_2. Similarly, we say E_1 and E_2 are *K-isomorphic* (or, simply, *isomorphic*) if there exists an isomorphism defined over K between them, and we denote it by $E_1 \cong E_2$.

Note that being isogenous is an equivalence relation. In particular, it is symmetric since, if there is an isogeny from E_1 to E_2, then the dual isogeny is an isogeny from E_2 to E_1.

There are some quantities and objects associated to an elliptic curve that remain the same among all curves in an isogeny class. One of them is the endomorphism algebra. If E_1 and E_2 are isogenous, then their endomorphism algebras are isomorphic, as the following proposition shows.

Proposition 3.26. Let E_1 and E_2 be elliptic curves over K and suppose there is an isogeny $\varphi : E_1 \to E_2$ defined over K. Then the map

$$\iota : \mathrm{End}^0(E_1) \to \mathrm{End}^0(E_2), \quad \phi \mapsto \frac{1}{\deg(\varphi)}\varphi\phi\hat{\varphi}$$

is an isomorphism of $\mathbb{Q}$-algebras.

Proof. Showing that this is a ring homomorphism that fixes $\mathbb{Q}$ is straightforward using the fact that $\varphi\hat{\varphi} = \deg(\varphi)$. It's easy to verify that the inverse map is given by $\mathrm{End}^0(E_2) \to \mathrm{End}^0(E_1)$ with $\psi \mapsto \frac{1}{\deg(\varphi)}\hat{\varphi}\psi\varphi$. □

Remark 3.27. Although two isogenous elliptic curves have the same endomorphism algebra, their endomorphism rings need not be isomorphic as we will see in the next chapter.

The endomorphism algebra is not the only isogeny invariant. The number of rational points, the trace of Frobenius, and the property of being supersingular are all isogeny invariant.

Proposition 3.28. Let E_1 and E_2 be elliptic curves over the finite field $\mathbb{F}_q$. If E_1 and E_2 are isogenous, then $\#E_1(\mathbb{F}_q) = \#E_2(\mathbb{F}_q)$.

Proof. If E_1 and E_2 are isogenous, then there is an isogeny $\phi : E_1 \to E_2$ defined over $\mathbb{F}_q$. Recall from the proof of Lemma 3.11 that $\#E_1(\mathbb{F}_q) = \deg(1 - \pi_{E_1})$ and $\#E_2(\mathbb{F}_q) = \deg(1 - \pi_{E_2})$. Note that since the coefficients of the rational functions defining ϕ are in $\mathbb{F}_q$, we have that $\phi \circ \pi_{E_1} = \pi_{E_2} \circ \phi$. This implies that

$$\phi \circ (1 - \pi_{E_1}) = (1 - \pi_{E_2}) \circ \phi.$$

Taking degrees and dividing by $\deg(\phi)$, we find that $\deg(1 - \pi_{E_1}) = \deg(1 - \pi_{E_2})$ which proves the result. □

Corollary 3.29. Let E_1 and E_2 be elliptic curves over the finite field $\mathbb{F}_q$. If E_1 and E_2 are isogenous, then $T(\pi_{E_1}) = T(\pi_{E_2})$.

Proof. From Lemma 3.11 we have that $T(\pi_{E_1}) = q + 1 - \#E_1(\mathbb{F}_q)$ and $T(\pi_{E_2}) = q + 1 - \#E_2(\mathbb{F}_q)$. By the previous proposition E_1 and E_2 have the same number of rational points, so the claim follows. □

Corollary 3.30. Let E_1 and E_2 be elliptic curves over the finite field $\mathbb{F}_q$. Suppose E_1 and E_2 are isogenous. If E_1 is supersingular (ordinary) then E_2 is supersingular (ordinary).

Proof. By the previous Corollary, E_1 and E_2 have the same trace of Frobenius. Now use Theorem 3.8. □

It turns out that the converse of Corollary 3.29 (and thus of Proposition 3.28) is also true, as was shown by Tate in 1966. We state, without proof, the following result.

Theorem 3.31 (Tate)**.** *Let E_1 and E_2 be elliptic curves over the finite field $\mathbb{F}_q$. Then E_1 and E_2 are isogenous if and only if $T(\pi_{E_1}) = T(\pi_{E_2})$.*

Proof. See [22, Theorem 1 (c)]. □

4 The class group action

This chapter will be dedicated to the action of the class group of a certain quadratic order on a set of supersingular elliptic curves. The protocol of CSIDH is based on this group action, and relies on the fact that the class group is an abelian group. Before we can state the main theorem, we need some definitions. Recall that if E/K is an elliptic curve, then the kernel of a non-zero endomorphism is a finite subgroup of $E(\bar{K})$ defined over K (see Example 1.93).

Definition 4.1. Let E be an elliptic curve over a field K with endomorphism ring $\mathcal{O}$ an order inside an imaginary quadratic number field. Let $I \subseteq \mathcal{O}$ be a non-zero $\mathcal{O}$-ideal. We define the finite subgroup of $E(\bar{K})$, called *the kernel of* I, by

$$E[I] := \bigcap_{\alpha \in \mathcal{I}} \ker(\alpha)$$

It's easy to see that the intersection of subgroups defined over K is again a subgroup defined over K. Hence, $E[I]$ is a finite subgroup defined over K. From Corollary 1.94 we have that there is a separable isogeny $\varphi_I : E \to E/E[I]$ with kernel $E[I]$ and with $E/E[I]$ an elliptic curve over K. Recall from Proposition 1.95 that the curve $E/E[I]$ is unique up to isomorphism.

CSIDH works with supersingular elliptic curves over $\mathbb{F}_p$. The main reason for this is that this allows us to control the number of rational points in the elliptic curves by choosing a suitable prime p (recall Corollary 3.13, which states that over the prime field $\mathbb{F}_p$, with $p \geq 5$, a curve is supersingular if and only if $\#E(\mathbb{F}_p) = p + 1$). This will be useful to compute the class group action of certain ideal classes in an efficient way. Therefore, it's convenient to let $p \geq 5$ be a prime and to fix the field $K = \mathbb{F}_p$ from now on.

Definition 4.2. Let $\mathcal{O}$ be an order inside an imaginary quadratic number field, and consider an element $\pi \in \mathcal{O}$. We write $Ell(\mathcal{O}, \pi)$ for the set of elliptic curves E over K (up to K-isomorphism) that have endomorphism ring $\mathrm{End}(E) = \mathcal{O}$, such that the element π corresponds to the Frobenius endomorphism of E.

Remark 4.3. In accordance with the previous definition, we consider two curves identical if they are isomorphic over K. Note that two curves can be isomorphic over $\bar{K}$ but not over K.

Example 4.4. Consider the order $\mathcal{O} = \mathbb{Z}[\pi]$ with $\pi^2 + p = 0$. If E is a curve in $Ell(\mathcal{O}, \pi)$, then its Frobenius endomorphism satisfies the characteristic polynomial $X^2 + p = 0$. This means that the trace of Frobenius is zero, which implies that E is supersingular. Thus, $Ell(\mathcal{O}, \pi)$ consists of supersingular elliptic curves that have endomorphism ring $\mathbb{Z}[\pi]$.

The previous example will be of central importance for the rest of the chapter as the curves considered in CSIDH will be the supersingular elliptic curves over $K = \mathbb{F}_p$ that have endomorphism ring exactly $\mathcal{O} = \mathbb{Z}[\pi]$ with $\pi^2 + p = 0$. Note that the endomorphism ring of a supersingular elliptic curve over K includes $\mathbb{Z}[\pi]$ but may be bigger in the case that $\mathbb{Z}[\pi]$ is not the maximal order of the endomorphism algebra $\mathbb{Q}(\pi)$ (which happens, by Proposition 2.38, precisely when $p \equiv 3 \mod 4$).

In order for a cryptographic protocol to be of practical significance, it must be computationally feasible. One difficulty that arises is the representation of an elliptic curve. Typically, one needs to store at least two coefficients to specify an elliptic curve (for example, by writing it in short Weierstrass form). We will see in the next chapter that by demanding $p \equiv 3 \mod 8$, the elements of $Ell(\mathcal{O}, \pi)$ may be specified with a single coefficient in $\mathbb{F}_p$. Thus, the CSIDH protocol utilizes a prime p such that $p \equiv 3 \mod 8$, which in particular implies that $\mathcal{O} = \mathbb{Z}[\pi]$ with $\pi^2 + p = 0$ is not the maximal order of $\mathbb{Q}(\pi)$. Let us make some remarks about this order $\mathcal{O}$.

Remark 4.5. The ring $\mathcal{O}$ is an order in the imaginary quadratic number field $\mathcal{K} = \mathbb{Q}(\pi) \cong \mathbb{Q}(\sqrt{-p})$. Since $p \equiv 3 \mod 4$ we have by Proposition 2.38 that the maximal order is $\mathcal{O}_\mathcal{K} = \mathbb{Z}[\frac{1+\pi}{2}]$, so that $\mathcal{O}$ has conductor $f = 2$.

Remark 4.6. Since $\mathcal{O}$ has conductor 2, then every prime ideal above a prime different from 2 is invertible by Proposition 2.55. To find the primes above 2 we use the Kummer-Dedekind Theorem (Theorem 2.32).

We see that $X^2 + p \equiv X^2 - 1 \equiv (X-1)^2 \mod 2$. Hence, the only prime above 2 is $\mathfrak{p} = (2, \pi - 1)$. The remainder of $X^2 + p$ after division by $X - 1$ is $p + 1$, which is divisible by 2^2 since $p \equiv 3 \mod 4$. Hence, by part 3 of Theorem 2.32, we conclude that $\mathfrak{p}$ is singular. Another way to see this is to note that $\mathfrak{p}^2 = 2\mathfrak{p}$.

The goal in this chapter will be to prove the following theorem, which is a particular case of a more general result proved by Waterhouse [26, Theorem 4.5]. For some of the proofs of the partial results, Waterhouse uses techniques that apply generally to abelian varieties. Since we will only need the result for elliptic curves defined over $\mathbb{F}_p$ with endomorphism ring $\mathbb{Z}[\pi]$, we will give alternative proofs that avoid the more sophisticated techniques.

Theorem 4.7 (The Class Group Action)**.** *Let p be a prime such that $p \equiv 3 \mod 8$ and let $K = \mathbb{F}_p$. Consider the order $\mathcal{O} = \mathbb{Z}[\pi]$ with $\pi^2 + p = 0$, which lies inside an imaginary quadratic field. Then the ideal class group $\mathrm{Cl}(\mathcal{O})$ acts freely and transitively on the set $Ell(\mathcal{O}, \pi)$ via the map*

$$\begin{aligned} \mathrm{Cl}(\mathcal{O}) \times Ell(\mathcal{O}, \pi) &\to Ell(\mathcal{O}, \pi) \\ ([I], E) &\mapsto E/E[I] \end{aligned}$$

where I is chosen to be an integral representative of its class.

There are multiple things about the map given in the statement above that we need to verify: that its codomain is indeed $Ell(\mathcal{O}, \pi)$, that it is well-defined, that it induces a group action, and that this action is free and transitive. We start by showing that it is well-defined and that it behaves compatibly with multiplication of invertible ideals.

4.1 Well-definedness

From this point forward, let p be a prime number such that $p \equiv 3 \mod 8$. Let $\mathcal{O} = \mathbb{Z}[\pi]$ with $\pi^2 + p = 0$, and recall that we have fixed the base field $K = \mathbb{F}_p$.

Lemma 4.8. *Let $E \in Ell(\mathcal{O}, \pi)$. Let I be an $\mathcal{O}$-ideal and $0 \neq \rho \in \mathcal{O}$. Then $E/E[I\rho] \cong E/E[I]$.*

Proof. This proof is from [12]. Consider the separable isogeny $\varphi_I : E \to E/E[I]$ with kernel $E[I]$. Consider the composition $\varphi_I \circ \rho : E \to E/E[I]$ which has kernel $\rho^{-1}(E[I])$. Note that

$$\rho^{-1}(E[I]) = \rho^{-1}\left(\bigcap_{\alpha \in I} \ker(\alpha)\right) = \bigcap_{\alpha \in I} \rho^{-1}(\ker(\alpha)) = \bigcap_{\alpha \in I} \ker(\alpha\rho) = E[I\rho]$$

Thus, the isogeny $\varphi_I \circ \rho : E \to E/E[I]$ has kernel $E[I\rho]$. By the uniqueness of the quotient implied by Proposition 1.95, we have

$$E/E[I\rho] \cong E/E[I].$$

□

The previous lemma implies that if two integral ideals $I, J \subseteq \mathcal{O}$ belong to the same class in $\mathrm{Cl}(\mathcal{O})$, then $E/E[I] \cong E/E[J]$. Since we consider curves up to K-isomorphism, then the two ideals produce the same element. This shows that the map in Theorem 4.7, if it exists, is well-defined. To prove the existence of this map, we still need to show that $E/E[I] \in Ell(\mathcal{O}, \pi)$.

Let E be a curve with endomorphism ring $\mathcal{O}$. It's not obvious that if $I \subseteq \mathcal{O}$ is an invertible ideal, then $E/E[I]$ also has endomorphism ring $\mathcal{O}$. Before we prove this, we need some other results. Let us simplify notation and write $E_I := E/E[I]$. Since supersingularity is isogeny invariant by Corollary 3.30, E_I is again supersingular so its Frobenius satisfies the same characteristic polynomial. Since $\mathrm{End}(E_I)$ contains $\mathbb{Z}[\pi]$, we may regard $\mathrm{End}(E)$ as a subset of $\mathrm{End}(E_I)$. More precisely, let $\varphi_I : E \to E_I$ be the separable isogeny with kernel $E[I]$ and consider the isomorphism of $\mathbb{Q}$-algebras given in Proposition 3.26.

$$\iota : \mathrm{End}^0(E) \to \mathrm{End}^0(E_I), \quad \phi \mapsto \frac{1}{\deg(\varphi_I)} \varphi_I \phi \hat{\varphi}_I$$

Recall that if $\varphi : E \to E_I$ is an isogeny over K, then $\pi_{E_I} \circ \varphi = \varphi \circ \pi_E$. Also, for any $n \in \mathbb{Z}$ we have that $[n] \circ \varphi = \varphi \circ [n]$. Hence, for any element $\psi = [n] + [m]\pi_E$ in $\mathrm{End}(E)$ with $n, m \in \mathbb{Z}$ it's easy to verify that $\iota(\psi) = [n] + [m]\pi_{E_I} \in \mathrm{End}(E_I)$. Thus, the map induces an embedding of $\mathrm{End}(E)$ into $\mathrm{End}(E_I)$.

We aim to prove that the action described in Theorem 4.7 is indeed a group action. Let $I, J \subseteq \mathcal{O}$ be ideals and $E \in Ell(\mathcal{O}, \pi)$. We do not claim yet that E_I has endomorphism ring $\mathcal{O}$, but, by the previous observation, we may regard J as a subset of $\mathrm{End}(E_I)$. Let us define $E_I[J] := \bigcap_{\beta \in J} \ker(\beta) \subseteq E_I(\bar{K})$ and $(E_I)_J := E_I/E_I[J]$. Let $\tilde{\varphi}_J : E_I \to (E_I)_J$ be the separable isogeny with kernel $E_I[J]$.

Lemma 4.9. *Let $E_{IJ} = E/E[IJ]$ and $\varphi_{IJ} : E \to E_{IJ}$ the separable isogeny with kernel $E[IJ]$. Then,*

$$(E_I)_J \cong E_{IJ}$$

and

$$\varphi_{IJ} = \tilde{\varphi}_J \circ \varphi_I$$

(the equality is up to composition with a K-isomorphism).

Proof. It suffices to show that the composition $\tilde{\varphi}_J \circ \varphi_I : E \to (E_I)_J$ has kernel $E[IJ]$.

We noted before that if $\varphi : E \to E_I$ is an isogeny defined over K, then $\pi_{E_I} \circ \varphi = \varphi \circ \pi_E$. Hence, if $\beta \in \mathcal{O} = \mathbb{Z}[\pi]$, then the following diagram

$$\begin{array}{ccc} E & \xrightarrow{\varphi_I} & E_I \\ {\scriptstyle\beta}\downarrow & & \downarrow{\scriptstyle\beta} \\ E & \xrightarrow{\varphi_I} & E_I \end{array}$$

commutes. Therefore, we have that

$$\begin{aligned} P \in \ker(\tilde{\varphi}_J \circ \varphi_I) &\iff \beta\varphi_I(P) = O \text{ for every } \beta \in J \\ &\iff \varphi_I(\beta P) = O \text{ for every } \beta \in J \quad \text{(by commutativity of the diagram)} \\ &\iff \beta P \in E[I] \text{ for every } \beta \in J \\ &\iff \alpha\beta P = O \text{ for every } \alpha \in I \text{ and every } \beta \in J. \\ &\iff P \in E[IJ]. \end{aligned}$$

By Proposition 1.95, we conclude that there is an isomorphism ψ (defined over K) such that

$$\begin{array}{ccc} E & \xrightarrow{\varphi_I} & E_I \\ {\scriptstyle\varphi_{IJ}}\downarrow & & \downarrow{\scriptstyle\tilde{\varphi}_J} \\ E_{IJ} & \xrightarrow{\psi} & (E_I)_J \end{array}$$

commutes. This proves the statement.

□

4.2 Kernel ideals

In order to prove that E_I is in $Ell(\mathcal{O}, \pi)$ and that the group action is free, we need to introduce the concept of a *kernel ideal*. Recall that if $I \subseteq \mathcal{O}$ is an ideal and $E \in Ell(\mathcal{O}, \pi)$, then $E[I]$ is a finite subgroup of $E(\bar{K})$ defined over K.

Is it possible to recover I from $E[I]$? Not necessarily. For example, note that the two ideals $\pi\mathcal{O}$ and $\mathcal{O}$ produce the same subgroup $E[\pi\mathcal{O}] = E[\mathcal{O}] = \{O\}$. Given the finite subgroup $E[I]$ one might consider all the endomorphisms of E that include $E[I]$ in their kernel. The kernel ideals are precisely the ones that can be reconstructed in this way.

Definition 4.10. Let $E \in Ell(\mathcal{O}, \pi)$. We say that an ideal $I \subseteq \mathcal{O}$ is a *kernel ideal* if

$$I = \{\alpha \in \mathcal{O} \mid \alpha(E[I]) = O\}$$

Our goal for this section will be to prove that every class in $\mathrm{Cl}(\mathcal{O})$ has a representative that is a kernel ideal. For this, we need some lemmas.

Note that $X^2 + p \equiv X^2 \mod p$, so we find using the Kummer-Dedekind Theorem (Theorem 2.32) that the ideal $(\pi) = \pi\mathcal{O}$ is an invertible prime ideal of norm p, that it's the only prime ideal above p and that $(p) = (\pi)(\pi)$. Let $I \subseteq \mathcal{O}$ be a non-zero ideal. By Remark 2.26 we have that the (π)-primary part of I is a power of (π). Hence, from Proposition 2.23 we get that I can be written as $I = (\pi)^n\mathfrak{a}$ for some n and $\mathfrak{a} \subseteq \mathcal{O}$ an ideal such that $(\pi) \not\supseteq \mathfrak{a}$.

Lemma 4.11. *Let $\mathfrak{a} \subseteq \mathcal{O}$ be an ideal such that $(\pi) \not\supseteq \mathfrak{a}$. Then $N(\mathfrak{a})$ is relatively prime to p.*

Proof. Since (π) is a maximal ideal and $\mathfrak{a} \not\subseteq (\pi)$ then $\mathfrak{a} + (\pi) = \mathcal{O}$, so $\mathfrak{a}$ and (π) are coprime ideals. It follows that $\mathfrak{a}$ and $(p) = (\pi)^2$ are coprime ideals so $p\mathcal{O} + \mathfrak{a} = \mathcal{O}$. As in the proof of Proposition 2.55, we have that the group homomorphism $\mathcal{O}/\mathfrak{a} \to \mathcal{O}/\mathfrak{a}$ given by multiplication by p is surjective, and, hence, a group isomorphism. This implies that the order of the group $\#\mathcal{O}/\mathfrak{a} = N(\mathfrak{a})$ is relatively prime to p. □

Lemma 4.12. *Let E be an elliptic curve over K with endomorphism algebra $\mathbb{Q}(\pi)$ such that the Frobenius satisfies $\pi^2 + p = 0$ (thus, $\mathcal{O} \subseteq \mathrm{End}(E)$). Let $I \subseteq \mathcal{O}$ be an invertible ideal and $\varphi_I : E \to E_I$ the separable isogeny with kernel $E[I] = \cap_{\alpha \in I} \ker(\alpha)$. Write $I = (\pi)^n\mathfrak{a}$ with an ideal such that $(\pi) \not\supseteq \mathfrak{a}$. Then $\deg \varphi_I$ divides $N(\mathfrak{a})$.*

Proof. Note that $\#E[I] = \#E[\mathfrak{a}]$ since the Frobenius has trivial kernel. Observe that $E[\mathfrak{a}]$ is a subgroup of $\ker(\alpha)$ for every $\alpha \in \mathfrak{a}$, thus $\#E[\mathfrak{a}]$ divides $\#\ker(\alpha)$ for every $0 \neq \alpha \in \mathfrak{a}$.

On the other hand, for every $0 \neq \alpha \in \mathfrak{a}$, we have that $\#\ker(\alpha) = \deg_s(\alpha)$ divides $\deg(\alpha)$ and, by Remark 3.19, $\deg(\alpha) = N(\alpha)$. Thus, $\#E[\mathfrak{a}]$ divides $N(\alpha)$ for every $\alpha \in \mathfrak{a}$ (note that $N(0) = 0$). By Corollary 2.54, since $\mathfrak{a}$ is invertible, we have that $N(\mathfrak{a}) = \gcd\{N(\alpha) \mid \alpha \in \mathfrak{a}\}$ which implies that $\#E[\mathfrak{a}]$ divides $N(\mathfrak{a})$. Since φ_I is separable, we have that $\deg(\varphi_I) = \#E[I]$. It follows that $\deg(\varphi_I)$ divides $N(\mathfrak{a})$. □

Lemma 4.13. *Let $E \in Ell(\mathcal{O}, \pi)$ and consider $I \subseteq \mathcal{O}$ an invertible ideal. Let $\varphi_I : E \to E$ be the separable isogeny with kernel $E[I]$. Write $I = (\pi)^n\mathfrak{a}$ with $(\pi) \not\supseteq \mathfrak{a}$. Then $N(\mathfrak{a}) = \deg(\varphi_I)$ and thus*

$$N(I) = p^n \deg(\varphi_I).$$

Proof. From Proposition 2.45 we have:

$$N(I) = N(\pi^n)N(\mathfrak{a}) = p^n N(\mathfrak{a}).$$

By the previous lemma we have that $\deg(\varphi_I)$ divides $N(\mathfrak{a})$.

To show that $N(\mathfrak{a})$ divides $\deg(\varphi_I)$ we will argue similarly to the proof of [24, Proposition 42.2.16.a], which works with a maximal order in a quaternion algebra instead of our non-maximal order $\mathcal{O}$ in a quadratic field. We will rely on the invertibility of $\mathfrak{a}$.

We start by using Corollary 2.53 and choosing $\alpha \in \mathfrak{a}$ such that the integral $\mathcal{O}$-ideal $\alpha\mathfrak{a}^{-1}$ has norm relatively prime to $N(\mathfrak{a})$. Let $\mathfrak{a}' = \alpha\mathfrak{a}^{-1}$. Thus, $N(\mathfrak{a})$ and $N(\mathfrak{a}')$ are relatively prime.

From Lemma 4.9 we have that there is an isomorphism ψ such that the diagram

$$\begin{array}{ccc} E & \xrightarrow{\varphi_{\mathfrak{a}}} & E_{\mathfrak{a}} \\ {\scriptstyle \varphi_{\mathfrak{a}\mathfrak{a}'}}\downarrow & & \downarrow{\scriptstyle \tilde{\varphi}_{\mathfrak{a}'}} \\ E_{\mathfrak{a}\mathfrak{a}'} & \xrightarrow{\psi} & (E_{\mathfrak{a}})_{\mathfrak{a}'} \end{array}$$

commutes. It follows that $\deg(\varphi_{\mathfrak{a}\mathfrak{a}'}) = \deg(\tilde{\varphi}_{\mathfrak{a}'}\varphi_{\mathfrak{a}})$. From Lemma 4.12 we also have that $\deg(\tilde{\varphi}_{\mathfrak{a}'})$ divides $N(\mathfrak{a}')$.

By Proposition 1.65, we may write $\alpha = (\pi)^m \alpha_{sep}$ for some positive integer m and a separable endomorphism α_{sep}. Thus, $\alpha_{sep} : E \to E$ is a separable isogeny with kernel $\ker(\alpha)$. Note that $E[\mathfrak{a}\mathfrak{a}'] = E[\alpha\mathcal{O}] = \ker(\alpha)$, so $\varphi_{\mathfrak{a}\mathfrak{a}'}$ is a separable isogeny with kernel $\ker(\alpha)$. Hence, we may take $\varphi_{\mathfrak{a}\mathfrak{a}'} = \alpha_{sep}$ and $E_{\mathfrak{a}\mathfrak{a}'} = E$. We have that $\alpha = (\pi)^m \varphi_{\mathfrak{a}\mathfrak{a}'}$. Thus,

$$N(\mathfrak{a})N(\mathfrak{a}') = N(\alpha\mathcal{O}) = N(\alpha) = \deg(\alpha) = p^m \deg(\varphi_{\mathfrak{a}\mathfrak{a}'}) = p^m \deg(\tilde{\varphi}_{\mathfrak{a}'}) \deg(\varphi_{\mathfrak{a}})$$

Recall that $\deg(\tilde{\varphi}_{\mathfrak{a}'})$ divides $N(\mathfrak{a}')$ and that $N(\mathfrak{a})$ and $N(\mathfrak{a}')$ are relatively prime. Thus, $\deg(\tilde{\varphi}_{\mathfrak{a}'})$ is relatively prime to $N(\mathfrak{a})$.

By Lemma 4.11 we have that $N(\mathfrak{a})$ is relatively prime to p. Hence, $N(\mathfrak{a})$ is relatively prime to $p^m \deg(\tilde{\varphi}_{\mathfrak{a}'})$. From the equation above we conclude that $N(\mathfrak{a})$ divides $\deg(\varphi_{\mathfrak{a}})$. But since $E[I] = E[\mathfrak{a}]$, we have that $\deg(\varphi_{\mathfrak{a}}) = \deg(\varphi_I)$. This concludes the proof. □

We are now ready to show that every ideal class in $\mathrm{Cl}(\mathcal{O})$ includes a kernel ideal.

Theorem 4.14. *Let $E \in Ell(\mathcal{O}, \pi)$ and consider $I \subseteq \mathcal{O}$ an invertible ideal. Then there is a kernel ideal $J \subseteq \mathcal{O}$ such that $[I] = [J]$ in $\mathrm{Cl}(\mathcal{O})$.*

Proof. By Theorem 2.52 there is an invertible $\mathcal{O}$-ideal in the same class as I such that its norm is relatively prime to 2. Hence, let us assume without loss of generality that $N(I)$ is relatively prime to 2. Define the $\mathcal{O}$-ideal:

$$J := \{\alpha \in \mathcal{O} \mid \alpha(E[I]) = O\}$$

Clearly $I \subseteq J$. Therefore, $E[I] \supseteq E[J]$. On the other hand, if $\alpha \in J$ then $E[I] \subseteq \ker(\alpha)$. Thus,

$$E[I] \subseteq \bigcap_{\alpha \in J} \ker(\alpha) = E[J].$$

We conclude that equality holds $E[I] = E[J]$. This implies that J is a kernel ideal. Also, if φ_I and φ_J are the separable isogenies with kernel $E[I]$ and $E[J]$, respectively, then we have that $\varphi_I = \varphi_J$.

We write $I = (\pi)^n \mathfrak{a}$ with $(\pi) \not\supseteq \mathfrak{a}$. Since I is invertible, we have from Lemma 4.13 that $N(I) = p^n \deg(\varphi_I)$. Also note that

$$[\mathcal{O} : I] = [\mathcal{O} : J][J : I]. \tag{4.1}$$

In particular, $N(J)$ divides $N(I)$, so $N(J)$ is also relatively prime to 2. Recall that $\mathcal{O}$ has conductor 2. By Proposition 2.55 we have that J is invertible. We may write $J = (\pi)^t \mathfrak{b}$ with $\mathfrak{b} \subseteq \mathcal{O}$ an ideal such that $(\pi) \not\supseteq \mathfrak{b}$. Note that $E[\mathfrak{b}] = E[J]$, so

$$J = (\pi)^t \mathfrak{b} \subseteq \mathfrak{b} \subseteq \{\alpha \in \mathcal{O} \mid \alpha(E[\mathfrak{b}]) = O\} = \{\alpha \in \mathcal{O} \mid \alpha(E[J]) = O\} = J \tag{4.2}$$

Hence, $J = \mathfrak{b}$ and $t = 0$. By Lemma 4.13 we have that $N(J) = \deg(\varphi_J)$. Using (4.1) we see that $N(I) = N(J)[J : I]$ and since $\deg(\varphi_I) = \deg(\varphi_J)$ and $N(I) = p^n \deg(\varphi_I)$ we conclude that $[J : I] = p^n$. This implies that $p^n J \subseteq I$. In other words,

$$(\pi)^{2n} J \subseteq I \subseteq J.$$

Taking localizations at each each prime ideal, we see that for every prime $\mathfrak{p}$ different from (π), the $\mathfrak{p}$-primary part of I and J are equal. The ideal J has trivial (π)-primary part since $(\pi) \not\supseteq J$, while the (π)-primary part of I is a power of (π) by Remark 2.26. Thus, looking at the primary decomposition of I we have that

$$I = (\pi)^n J$$

so $[I] = [J]$. □

Proposition 4.15. Let $E \in Ell(\mathcal{O}, \pi)$ and consider a kernel ideal $I \subseteq \mathcal{O}$. Let $\rho \in \mathcal{O}$ be a separable non-zero endomorphism. Then ρI is also a kernel ideal.

Proof. We follow Waterhouse [26, Lemma 3.10]. Clearly $\rho I \subseteq \{\alpha \in \mathcal{O} \mid \alpha(E[\rho I]) = O\}$. To prove the opposite containment, consider $\alpha \in \mathcal{O}$ such that $\alpha(E[\rho I]) = O$. Since $\rho I \subseteq \rho\mathcal{O}$ we have that $E[\rho I] \supseteq E[\rho\mathcal{O}] = \ker(\rho)$. Hence, $\ker(\rho) \subseteq \ker(\alpha)$. Since ρ is separable, we have by Corollary 1.73 that $\alpha = \beta\rho$ for some $\beta \in \mathcal{O}$. We claim that $\beta(E[I]) = O$.

Note that $E[\rho I] = E[I\rho] = \rho^{-1}(E[I])$. Let $Q \in E[I]$. Since ρ is surjective, there is $P \in \rho^{-1}(E[I]) = E[\rho I]$ such that $\rho(P) = Q$. Hence, $\beta Q = \beta\rho P = \alpha P = 0$. It follows that $\beta Q = 0$ for every $Q \in E[I]$. In other words, $\beta(E[I]) = O$. Since I is a kernel ideal, this implies that $\beta \in I$. Therefore, $\alpha \in \rho I$ which finishes the proof. □

We are able to give a characterization of the invertible kernel ideals.

Corollary 4.16. Let $E \in Ell(\mathcal{O}, \pi)$ and consider an invertible ideal $I \subseteq \mathcal{O}$. Then I is a kernel ideal if and only if $(\pi) \not\supseteq I$.

Proof. One direction is easy. Assume I is a kernel ideal and write it in the form $I = (\pi)^t \mathfrak{a}$ with $(\pi) \not\supseteq \mathfrak{a}$. Just as in (4.2), we conclude that $t = 0$ so $(\pi) \not\supseteq I$.

For the other implication, assume $(\pi) \not\supseteq I$. By Lemma 4.11, the norm $N(I)$ is relatively prime to p. From Theorem 4.14 we have that there is a kernel ideal J in the same class as I. Hence, there are non-zero $\alpha, \beta \in \mathcal{O}$ such that

$$\alpha I = \beta J. \tag{4.3}$$

Since J is a kernel ideal, it's not contained in (π). Again by Lemma 4.11, we have that $N(J)$ is relatively prime to p.

Let us write $\alpha = \pi^m \alpha'$ and $\beta = \pi^n \beta'$ where n, m are some non-negative integers, and α', β' are separable endomorphisms. Note that $N(\alpha') = \deg(\alpha') = \#\ker(\alpha')$ which is relatively prime to p since there are no point of order p in E. Similarly, $N(\beta')$ is relatively prime to p. We have

$$p^m N(\alpha') N(I) = N(\alpha I) = N(\beta J) = p^n N(\beta') N(J).$$

We conclude that $m = n$. Canceling the power of Frobenius on both sides of (4.3), we get

$$\alpha' I = \beta' J$$

Since β' is a separable endomorphism and J is a kernel ideal, we get from Proposition 4.15 that $\beta' J$ is a kernel ideal. It follows that $\alpha' I$ is a kernel ideal.

To prove that I is a kernel ideal as well, suppose that $\sigma \in \mathcal{O}$ is such that $\sigma(P) = O$ for every $P \in E[I]$. Consider any $Q \in E[\alpha' I]$ and note that $\alpha'(Q) \in E[I]$. Hence, $\sigma(\alpha'(Q)) = (\sigma\alpha')(Q) = O$. This implies that $\sigma\alpha' \in \{\gamma \in \mathcal{O} \mid \gamma(E[\alpha' I]) = O\} = \alpha' I$ since $\alpha' I$ is a kernel ideal. It follows that $\sigma \in I$, which proves that I is a kernel ideal. □

4.3 The endomorphism ring of E_I

We have seen that if $E \in Ell(\mathcal{O}, I)$ and $I \subseteq \mathcal{O}$ is an invertible ideal, then $E_I = E/E[I]$ also has endomorphism algebra $\mathbb{Q}(\pi)$ since E and E_I are K-isogenous. The ring $\mathrm{End}(E_I)$ is an order in $\mathbb{Q}(\pi)$ containing $\mathcal{O}$, but it's not immediately clear that $\mathrm{End}(E_I) = \mathcal{O}$. In fact, this is dependent on I being invertible. In this section we will show that $E_I \in Ell(\mathcal{O}, \pi)$ by exhibiting an isomorphism $\mathrm{End}(E) \to \mathrm{End}(E_I)$ that sends the Frobenius to the Frobenius.

Proposition 4.17. Let $E \in Ell(\mathcal{O}, \pi)$ and let I be an invertible $\mathcal{O}$-ideal. Then the curve $E_I = E/E[I]$ is in $Ell(\mathcal{O}, \pi)$.

Proof. First of all, recall that if two ideals I and J are in the same class, then $E_I = E_J$ by Lemma 4.8 (the equality being up to K-isomorphism). By Theorem 4.14, there is a kernel ideal in the class of I. Hence, replacing I with this kernel ideal in the same class, we may assume without loss of generality that I is a kernel ideal.

Let $\varphi_I : E \to E_I$ be the separable isogeny with kernel $E[I]$. Let $n = \deg(\varphi_I) = \#E[I]$. Note that n is relatively prime with p. If this wasn't the case, then, by Cauchy's theorem, the finite group $E[I]$ would have an element of order p. But this is not possible since E is supersingular so $E[p] = \{O\}$.

Recall the isomorphism of $\mathbb{Q}$-algebras given in Proposition 3.26 by.

$$\iota : \mathrm{End}^0(E) \to \mathrm{End}^0(E_I), \quad \phi \mapsto \frac{1}{n}\varphi_I \phi \hat{\varphi}_I$$

We claim that ι restricts to a ring isomorphism $\iota : \mathrm{End}(E) \to \mathrm{End}(E_I)$ that sends the Frobenius to the Frobenius. We previously remarked that $\iota(\mathrm{End}(E)) \subseteq \mathrm{End}(E_I)$ and that $\iota(\pi_E) = \pi_{E_I}$.

Hence, it remains to be shown that $\iota^{-1}(\mathrm{End}(E_I)) \subseteq \mathrm{End}(E)$. Consider $\alpha \in \iota^{-1}(\mathrm{End}(E_I)) \subseteq \mathrm{End}^0(E)$, then $\iota(\alpha) = \beta$ for some $\beta \in \mathrm{End}(E_I)$. We have that $\beta = \frac{1}{n}\varphi_I \alpha \hat{\varphi}_I$. Thus,

$$\alpha = \frac{1}{n}\hat{\varphi}_I \beta \varphi_I$$

Let $\tau \in I$. We claim that $\tau\alpha \in \mathcal{O}$. Note that

$$\tau\alpha = \frac{1}{n}\tau\hat{\varphi}_I \beta \varphi_I$$

Clearly $\tau\hat{\varphi}_I\beta\varphi_I \in \mathrm{End}(E)$. Note that $\tau\hat{\varphi}_I : E_I \to E$ has kernel ${\hat{\varphi}_I}^{-1}(\ker(\tau))$. Also observe that, since $\tau \in I$, we have that $E[I] \subseteq \ker(\tau)$. Therefore, $\ker(\varphi_I) \subseteq \ker(\tau)$. This implies that

$$E_I[n] = \ker([n]_{E_I}) = \ker(\varphi_I\hat{\varphi}_I) = {\hat{\varphi}_I}^{-1}(\ker(\varphi_I)) \subseteq {\hat{\varphi}_I}^{-1}(\ker(\tau)) = \ker(\tau\hat{\varphi}_I)$$

It follows that the kernel of the multiplication-by-n map $[n]_{E_I} : E_I \to E_I$ is contained in the kernel of $\tau\hat{\varphi}_I : E_I \to E$. Since n is relatively prime to p, we have that $[n]_{E_I} : E_I \to E_I$ is separable. By Corollary 1.73 there is an isogeny $\rho : E_I \to E$ such that $\tau\hat{\varphi}_I = \rho[n]_{E_I} = [n]_E\rho$. We conclude that

$$\tau\alpha = \frac{1}{n}\tau\hat{\varphi}_I\beta\varphi_I = \frac{1}{n}n\rho\beta\varphi_I = \rho\beta\varphi_I \in \mathrm{End}(E),$$

so $\tau\alpha \in \mathrm{End}(E) = \mathcal{O}$. In fact, $\tau\alpha \in I$. To see this, consider $P \in E[I]$. Since $\varphi_I(P) = O$ we see from the expression above that $(\tau\alpha)(P) = O$ so $(\tau\alpha)(E[I]) = \{O\}$. This implies that

$$\tau\alpha \in \{x \in \mathcal{O} \mid x(E[I]) = O\} = I$$

since I is a kernel ideal. This is true for every $\tau \in I$, so $I\alpha \subseteq I$.

Recall that every invertible ideal is proper by Remark 2.11. This implies that $\alpha \in \mathcal{O}$, which is what we wanted to show.

□

Remark 4.18. As we have previously emphasized, the assumption that I is invertible is crucial. Recall that the only prime above 2 is $\mathfrak{p} = (2, \pi - 1)$ and it is singular. We will later see that if $E \in Ell(\mathcal{O}, \pi)$ then $E/E[\mathfrak{p}]$ has endomorphism ring $\mathbb{Z}[\frac{1+\pi}{2}]$ and not $\mathcal{O}$.

The previous proposition together with Lemma 4.8 shows that the map in Theorem 4.7 exists. From Lemma 4.9 we get that this is a group action of $\mathrm{Cl}(\mathcal{O})$ on $Ell(\mathcal{O}, \pi)$.

4.4 Proof that the class group action is free

We are now ready to prove that the group action described in Theorem 4.7 is free.

Proposition 4.19. Consider $E \in Ell(\mathcal{O}, \pi)$ and let $I, J \subseteq \mathcal{O}$ be ideals. Suppose

$$E/E[I] \cong E/E[J].$$

Then there exist a separable endomorphism $\rho \in \mathrm{End}(E)$ and a non-zero integer M relatively prime to p such that

$$E[I\rho] = E[JM].$$

Proof. This proof is based on Waterhouse's proof in [26, Proposition 3.6]. We have that there is an elliptic curve E' over K and separable isogenies $\varphi_I, \varphi_J : E \to E'$ such that $\ker(\varphi_I) = E[I]$ and $\ker(\varphi_J) = E[J]$. Take $N = \#E[I]$. Since E is supersingular, the order of any finite subgroup is relatively prime to p, thus N is relatively prime to p. Note that $N^{-1}(E[J]) \supseteq N^{-1}(0) \supseteq E[I] = \ker \varphi_I$. We see that $\ker(N\varphi_J) = \ker(\varphi_J N) = N^{-1}(E[J])$, hence $\ker(N\varphi_J) \supseteq \ker(\varphi_I)$.

From Corollary 1.73 it follows that there is an isogeny $\sigma : E' \to E'$ such that $\sigma\varphi_I = N\varphi_J$. Since both $\deg(\varphi_J) = \#E[J]$ and N are relatively prime to p, we have that $\deg(\sigma)$ is also relatively prime to p. We conclude that σ is separable. Now, consider the separable endomorphism $\rho := \widehat{\varphi}_I \sigma \varphi_I$. Note that $\varphi_I \rho = \deg(\varphi_I)\sigma\varphi_I = N\sigma\varphi_I$. Therefore,

$$\varphi_I \rho = N\sigma\varphi_I = N^2\varphi_J = M\varphi_J.$$

where $M = N^2$. Thus, we have that $\ker(\varphi_I\rho) = \ker(M\varphi_J)$. But $E[I\rho] = \ker(\varphi_I\rho)$ and $E[JM] = \ker(M\varphi_J)$, which proves the result.

□

Corollary 4.20. The ideal class group action described in Theorem 4.7 is free.

Proof. It suffices to show that if $[I], [J] \in \mathrm{Cl}(\mathcal{O})$ are such such that $E/E[I] \cong E/E[J]$, then $[I] = [J]$. By Theorem 4.14 we can always take the representatives $I, J \subseteq \mathcal{O}$ to be kernel ideals.

From Proposition 4.19 we know that there is an integer M relatively prime to p and a separable endomorphism $\rho \in \mathrm{End}(E)$ such that $E[I\rho] = E[JM]$. Note that since M is relatively prime to p, the multiplication-by-M map is also separable. Hence, by Proposition 4.15 we get that $I\rho$ and JM are kernel ideals. We conclude that

$$I\rho = \{\alpha \in \mathcal{O} \mid \alpha(E[I\rho]) = O\} = \{\alpha \in \mathcal{O} \mid \alpha(E[JM]) = O\} = JM$$

This shows that

$$[I] = [I\rho] = [JM] = [J].$$

□

4.5 Proof that the class group action is transitive

In this section we show that the class group action is transitive. That is, we will show that if E and E' are in $Ell(\mathcal{O}, \pi)$, then there is an invertible ideal $I \subseteq \mathcal{O}$ such that $E' \cong E/E[I]$.

Note that since E and E' have the same trace of Frobenius, then, by Tate's Isogeny Theorem (Theorem 3.31) there is an isogeny $\phi : E \to E'$ over K. We may write $\phi = \phi_{sep} \circ \pi^n$ with ϕ_{sep} separable by Proposition 1.65, and since π is an endomorphism of E, we have that ϕ_{sep} is an isogeny from E to E'. So, without loss of generality, we may take the isogeny ϕ to be separable.

We will prove our desired result by considering the different possible structures of the finite subgroup $\ker(\phi)$. The fact that $\ker(\phi)$ is defined over K is essential.

Remark 4.21. Consider a finite subgroup $G \subseteq E(\bar{K})$. Recall that G is defined over K if $\tau(G) = G$ for every $\tau \in \mathrm{Gal}(\bar{K}/K)$. Also recall that if $n \geq 1$ then $\mathrm{Gal}(\mathbb{F}_{p^n}/\mathbb{F}_p)$ is cyclic, generated by the automorphism $\sigma : x \to x^p$. Since G is finite, there is $n \geq 1$ such that $G \subseteq E(\mathbb{F}_{p^n})$. An element $\tau \in \mathrm{Gal}(\bar{K}/K)$ restricted to $\mathbb{F}_{p^n}$ is equal to σ^m for some $m \geq 0$. Hence, if $P \in G$ we have that $\tau(P) = \sigma^m(P) = \pi^m(P)$. Consequently, G is defined over K if and only if $\pi(G) = G$.

We start by making some observations about finite subgroups of $E(\bar{K})$.

Proposition 4.22. Let E be an elliptic curve over K and l a prime number. Suppose G is a subgroup of $E(\bar{K})$ with order a power of l. Then

$$G \cong \mathbb{Z}/l^n\mathbb{Z} \times \mathbb{Z}/l^m\mathbb{Z}$$

with $n, m \geq 0$.

Proof. This is clearly true if G is trivial. Suppose G is not trivial and thus, by the structure theorem of finite abelian groups, we have that

$$G \cong \mathbb{Z}/l^{n_1}\mathbb{Z} \times \mathbb{Z}/l^{n_2}\mathbb{Z} \times \ldots \times \mathbb{Z}/l^{n_k}\mathbb{Z}$$

with $0 < n_1 \leq \ldots \leq n_k$ and $k \geq 1$. Then G has a subgroup H isomorphic to $(\mathbb{Z}/l^{n_1}\mathbb{Z})^k$. Every element of H has order dividing l^{n_1} so $H \subseteq E[l^{n_1}]$. But, there are exactly $(l^{n_1})^2$ elements in $E[l^{n_1}]$ so k is at most 2. □

For now, we will work with primes $l \neq 2, p$. The case $l = 2$ will be handled later since it requires a slightly different treatment. The case $l = p$ never shows up since our curves will be supersingular so they have no subgroups of order p. We will show that if G is a subgroup of order relatively prime with 2 defined over K, then there is an invertible ideal $I \subseteq \mathcal{O}$ such that $E[I] = G$. We do this in incremental steps, starting from the simplest possible structure for G.

Lemma 4.23. *Consider $E \in Ell(\mathcal{O}, \pi)$. Let l be an odd prime and G a subgroup of $E(\bar{K})$ of order l defined over K. Then l splits in $\mathcal{O}$ and there is an invertible prime ideal $\mathfrak{p}$ above l such that $E[\mathfrak{p}] = G$.*

Proof. Since there are no subgroups of order p in a supersingular curve, l is automatically different from p. We have $G \cong \mathbb{Z}/l\mathbb{Z}$. Let P be a generator of G. Since G is defined over K we have that $\pi(P) \in G$ and thus there is $\lambda \in \{1, \ldots, l-1\}$ such that $\pi(P) = \lambda P$. Applying π again we get

$$[-p](P) = \pi^2(P) = \pi(\pi(P)) = \pi(\lambda P) = \lambda\pi(P) = \lambda^2 P.$$

Thus, $-p \equiv \lambda^2 \mod l$. Since l is not 2 nor p, then $\lambda \not\equiv -\lambda \mod l$. Therefore, the polynomial $X^2 + p$ splits modulo l as

$$X^2 + p \equiv (X - \lambda)(X + \lambda) \mod l.$$

By the Kummer-Dedekind Theorem (Theorem 2.32) we have that l splits in $\mathcal{O}$ into prime ideals as

$$(l) = (l, \pi - \lambda)(l, \pi + \lambda).$$

Let $\mathfrak{p} = (l, \pi - \lambda)$, which has norm l since the polynomial $X - \lambda$ has degree 1 (see proof of Theorem 2.32, part 1).

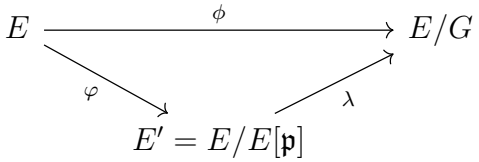
Let $Q \in G$. Note that $Q \in E[l]$. Also, $Q = nP$ for some integer n, so $\pi(Q) = n\pi(P) = n\lambda P = \lambda Q$. This means that Q is in $\ker(\pi - \lambda)$. It follows that $G \subseteq E[\mathfrak{p}]$. But $\#E[\mathfrak{p}] = N(\mathfrak{p}) = l$ by Lemma 4.13, therefore we must have $E[\mathfrak{p}] = G$. □

Lemma 4.24. *Consider $E \in Ell(\mathcal{O}, \pi)$. Let l be an odd prime and G a cyclic subgroup of $E(\bar{K})$ of order l^n defined over K. Then there is an invertible ideal $I \subseteq \mathcal{O}$ such that $E[I] = G$.*

Proof. We have $G \cong \mathbb{Z}/l^n\mathbb{Z}$. We will prove the result by induction. It is clearly true for $n = 0$ by taking $I = \mathcal{O}$. Now, let $n \geq 1$ and suppose the result is true for $n - 1$.

Since G is defined over K there is a separable isogeny $\phi : E \to E/G$ over K with kernel G. Note that $l^{n-1}G$ is a cyclic group of order l and that $\pi(l^{n-1}G) = l^{n-1}\pi(G) = l^{n-1}G$ since G is defined over K. Hence, $l^{n-1}G$ is defined over K. From the previous lemma we know that there is an invertible prime ideal $\mathfrak{p}$ such that $E[\mathfrak{p}] = l^{n-1}G$.

Consider the separable isogeny $\varphi : E \to E/E[\mathfrak{p}]$ over K with kernel $l^{n-1}G$. Since the kernel of φ is contained in G, we have by Corollary 1.73 that there is an isogeny $\lambda : E/E[\mathfrak{p}] \to E/G$ over K such that



commutes. Note that since $\mathfrak{p}$ is invertible, we have that $E' := E/E[\mathfrak{p}]$ has endomorphism ring $\mathcal{O}$ by Proposition 4.17. Also note that $\ker(\lambda)$ is a subgroup defined over K and $\ker(\lambda) = \ker(\phi)/\ker(\varphi) \cong \mathbb{Z}/l^{n-1}\mathbb{Z}$ so that $\ker(\lambda)$ is a cyclic subgroup of $E'(\bar{K})$ defined over K of order l^{n-1}. By the induction hypothesis we know that there is an invertible ideal $J \subseteq \mathcal{O}$ such that $E'[J] = \ker(\lambda)$.

Let $I = \mathfrak{p}J$ which is an invertible ideal. From Lemma 4.9 we get that

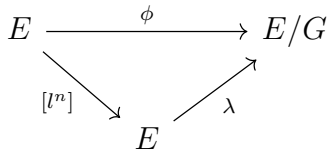
$$E[I] = E[\mathfrak{p}J] = \ker(\lambda \circ \varphi) = \ker(\phi) = G.$$

This shows the result for n and concludes the proof. □

We are now ready to drop the assumption that G is cyclic.

Lemma 4.25. *Consider $E \in Ell(\mathcal{O}, \pi)$. Let l be an odd prime and G a subgroup of $E(\bar{K})$ of order a power of l defined over K. Then there is an invertible ideal $I \subseteq \mathcal{O}$ such that $E[I] = G$.*

Proof. Consider the separable isogeny $\phi : E \to E/G$ over K with kernel G. Let n be the biggest integer such that $E[l^n] \subseteq G$. Then, we can factor $\phi = \lambda \circ [l^n]$ with $\lambda : E \to E/G$ a separable isogeny defined over K.



Note that $\deg(\lambda)$ is a power of l so that $\ker(\lambda)$ has order a power of l. We claim that $\ker(\lambda)$ is cyclic. Suppose is not, then from Proposition 4.22 we have that $\ker(\lambda) \cong \mathbb{Z}/l^n\mathbb{Z} \times \mathbb{Z}/l^m\mathbb{Z}$ with $m, n \geq 1$. Then $\ker(\lambda)$ has a subgroup isomorphic to $\mathbb{Z}/l\mathbb{Z} \times \mathbb{Z}/l\mathbb{Z}$, which is precisely $E[l]$. Thus, $E[l^{n+1}] \subseteq G$ which contradicts the fact that n is the biggest integer such that $E[l^n] \subseteq G$.

Therefore, $\ker(\lambda)$ is indeed a cyclic group of order a power of l defined over K. From the previous lemma we get that there is an invertible ideal $J \subseteq \mathcal{O}$ such that $E[J] = \ker(\lambda)$. Let $I = l^n J$ which is an invertible ideal. From Lemma 4.9 we get

$$E[I] = E[l^n J] = \ker(\lambda \circ [l^n]) = \ker(\phi) = G.$$

□

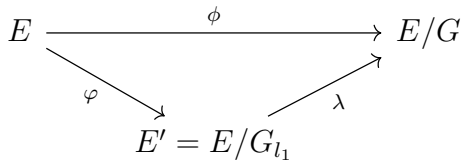
We can now state the general proposition.

Proposition 4.26. Consider $E \in Ell(\mathcal{O}, \pi)$. Let G be a subgroup of $E(\bar{K})$ of order relatively prime to 2 defined over K. Then there is an invertible ideal $I \subseteq \mathcal{O}$ such that $E[I] = G$.

Proof. We will show the result by induction on the number of distinct primes, say n, dividing $\#G$. The case $n = 0$ is trivially true by taking $I = \mathcal{O}$. The case $n = 1$ is true by Lemma 4.25. Let $n > 1$ and assume the result for $n-1$. Let $l_1, \ldots, l_n$ be the distinct primes dividing $\#G$. We can write G as a direct sum of its Sylow l-subgroups.

$$G = G_{l_1} \oplus \cdots \oplus G_{l_n}$$

Consider the separable isogeny $\phi : E \to E/G$ over K with kernel G. Observe that the l_1-subgroup G_{l_1} is defined over K: if $P \in G_{l_1}$ has order l_1^k then $l_1^k(\pi(P)) = \pi(l_1^k P) = O$ so the order of $\pi(P)$ is also a power of l_1. Hence, $\pi(P) \in G_{l_1}$. Therefore, there is a separable isogeny $\varphi : E \to E/G_{l_1}$ defined over K with kernel G_{l_1} and a separable isogeny $\lambda : E/G_{l_1} \to E/G$ over K such that



commutes. Since G_{l_1} is a subgroup defined over K of order a power of l_1 we have that, by Lemma 4.25, there is an invertible ideal $J \subseteq \mathcal{O}$ such that $E[J] = G_{l_1}$. Since J is invertible we also have that $E' := E/E[J] = E/G_{l_1}$ has endomorphism ring $\mathcal{O}$. Note that $\#\ker(\lambda) = \#G/\#G_{l_1}$ so that exactly $n-1$ distinct primes divide $\#\ker(\lambda)$. By the induction hypothesis we get that there is an invertible ideal $J' \subseteq \mathcal{O}$ such that $E'[J'] = \ker(\lambda)$. Let $I = JJ'$, which is an invertible ideal. From Lemma 4.9 we get

$$E[I] = E[JJ'] = \ker(\lambda \circ \varphi) = \ker(\phi) = G.$$

□

Corollary 4.27. Let E and E' be elliptic curves over K and let E have endomorphism ring $\mathcal{O}$. Suppose $\phi : E \to E'$ is a separable isogeny over K with degree coprime with 2. Then there is an invertible ideal $I \subseteq \mathcal{O}$ such that

$$E' \cong E/E[I].$$

Furthermore, E' has endomorphism ring $\mathcal{O}$.

Proof. The curve E' is isomorphic to $E/\ker(\phi)$ by Proposition 1.95. We have that $\#\ker(\phi) = \deg(\phi)$ is relatively prime to 2, so we apply Proposition 4.26 to conclude that there is an invertible ideal $I \subseteq \mathcal{O}$ such that $E[I] = \ker(\phi)$. By Proposition 4.17 the endomorphism ring of $E/E[I]$ is again $\mathcal{O}$. □

We will now work on the case $l = 2$. Here a difficulty arises since the only prime above 2 given by $\mathfrak{p} = (2, \pi - 1)$ is not invertible. This is related to the fact that an isogeny of degree 2 changes the endomorphism ring of the codomain, as we will later see.

Lemma 4.28. *Let $E \in Ell(\mathcal{O}, \pi)$. Then $E(\mathbb{F}_p)[2]$ (the rational points in $E[2]$) consists of precisely two points.*

Proof. Since $E(\mathbb{F}_p)[2]$ is a subgroup of $E[2]$, it consists of 1, 2 or 4 points. Since $\#E(\mathbb{F}_p) = p + 1$ is divisible by 2, $E(\mathbb{F}_p)$ must have an element of order 2. Hence, $\#E(\mathbb{F}_p)[2]$ is either 2 or 4. Suppose it's 4, then

$$\ker[2] = E[2] = E(\mathbb{F}_p)[2] \subseteq E(\mathbb{F}_p) = \ker(\pi - 1).$$

Therefore we can write $\pi - 1 = \phi \circ [2]$ with $\phi : E \to E$ an endomorphism defined over K. But then $\phi = \frac{\pi-1}{2}$ is an endomorphism of E that is not in $\mathcal{O} = \mathbb{Z}[\pi]$, a contradiction. □

Lemma 4.29. *Let E and E' be elliptic curves over K and let E have endomorphism ring $\mathcal{O}$. Suppose $\phi : E \to E'$ is an isogeny over K of degree 2. Then* $\mathrm{End}(E') \cong \mathbb{Z}[\frac{\pi-1}{2}]$.

Proof. From the previous Lemma, $E[2]$ has only two rational points, say $\{O, P\}$. Let Q be a non-rational point in $E[2]$. Then,

$$E[2] = \{O, P, Q, Q + P\}$$

Note that $\ker(\phi)$ is a subgroup of order 2 defined over K. It's easy to see that the two points in this subgroup are fixed by the Frobenius, so $\ker(\phi)$ consist of only rational points. Therefore, $\ker(\phi) = \{O, P\}$.

We will show that $E'[2]$ is all rational. Since $\#E(\mathbb{F}_p) = p + 1$ and $p \equiv -1 \mod 4$, we have that $4 \mid p + 1$. Since E doesn't have full rational 2-torsion, there must exist a rational point $R \in E(\mathbb{F}_p)$ of order 4. Clearly, since ϕ is defined over K, we have that $\phi(R)$ is a rational point in E'. Note that $2R = P$, so $2\phi(R) = \phi(P) = O$. It's also clear that $\phi(R) \neq O$. It follows that $\phi(R)$ is a rational point of order 2.

Now, consider the non-rational point Q in $E[2]$. Observe that $\pi(Q)$ is a point of order 2 different from Q, and also different from O and P since π is injective. Hence, $\pi(Q) = Q + P$. Clearly, $\phi(Q)$ is a point of order 2, and note that $\pi(\phi(Q)) = \phi(\pi(Q)) = \phi(P + Q) = \phi(Q)$, so $\phi(Q)$ is rational.

We claim that $\phi(Q)$ and $\phi(R)$ are distinct. For the sake of contradiction, assume $\phi(Q) = \phi(R)$. This implies that $Q - R \in \ker(\phi)$. But $\ker(\phi)$ consists of only rational points, so Q would be the sum of two rational points and, hence, rational. A contradiction. We conclude that

$$E'[2] = \{O, \phi(R), \phi(Q), \phi(R) + \phi(Q)\}$$

is all rational.

Consequently, $E'[2] \subseteq \ker(\pi - 1) = E'(\mathbb{F}_p)$ and we can write $\pi - 1 = \varphi \circ [2]$ with φ an endomorphism over K. This implies that $\varphi = \frac{\pi-1}{2} \in \mathrm{End}(E')$. Therefore $\mathbb{Z}\left[\frac{\pi-1}{2}\right] \subseteq \mathrm{End}(E')$. But $\mathbb{Z}\left[\frac{\pi-1}{2}\right]$ is the maximal order of $\mathrm{End}^o(E') = \mathbb{Q}(\pi)$, so we must have $\mathrm{End}(E') = \mathbb{Z}\left[\frac{\pi-1}{2}\right]$. □

Remark 4.30. Let E be an elliptic curve with endomorphism ring $\mathcal{O}$. The previous lemma shows that if G is a subgroup of order 2 defined over K, then it's impossible to find an invertible ideal $I \subseteq \mathcal{O}$ such that $E[I] = G$. This is because for any invertible ideal $I \subseteq \mathcal{O}$ the curve $E/E[I]$ has endomorphism ring $\mathcal{O}$ while E/G has endomorphism ring larger than $\mathcal{O}$.

As an example, consider $\mathfrak{p} = (2, \pi - 1)$ again. Note that $E[\mathfrak{p}] = E(\mathbb{F}_p)[2]$ which consists of two elements. Hence, $E[\mathfrak{p}]$ is a subgroup defined over K of order 2. By the previous lemma $E/E[\mathfrak{p}]$ has endomorphism ring $\mathbb{Z}[\frac{\pi-1}{2}]$, and thus $\mathfrak{p}$ is not invertible, which we already knew.

It turns out that if G is a cyclic subgroup of order 4 defined over K, then it is possible to find such an invertible ideal I with $E[I] = G$.

Lemma 4.31. *Consider $E \in Ell(\mathcal{O}, \pi)$. Let G be a cyclic subgroup of $E(\bar{K})$ of order 4 defined over K. Then there is an invertible ideal $I \subseteq \mathcal{O}$ such that $E[I] = G$.*

Proof. The proof is similar to Lemma 4.23. Let $P \in G$ be a generator. Since G is defined over K we have that $\pi(P) = \lambda P$ for some integer λ. This implies that $-pP = \pi(\pi(P)) = \lambda^2 P$ so that $-p \equiv \lambda^2 \mod 4$. Recall that $-p \equiv 1 \mod 4$, so $\lambda \equiv 1 \mod 4$ or $\lambda \equiv -1 \mod 4$.

Consider the following product of ideals

$$(4, \pi - 1)(4, \pi + 1) = (16, 4\pi + 4, 4\pi - 4, \pi^2 - 1) = (16, 4\pi + 4, 4\pi - 4, -p - 1) \subseteq (4).$$

Recall that $p \equiv 3 \mod 8$. Therefore, 4 is the highest power of 2 that divides $-p - 1$. Thus $\gcd(16, -p-1) = 4$. There are $x, y \in \mathbb{Z}$ such that $16x+(-p-1)y = 4$ which implies that $4 \in (16, 4\pi+4, 4\pi-4, -p-1)$. We conclude that $(4) = (16, 4\pi+4, 4\pi-4, -p-1)$. Thus,

$$(4, \pi - 1)(4, \pi + 1) = (4)$$

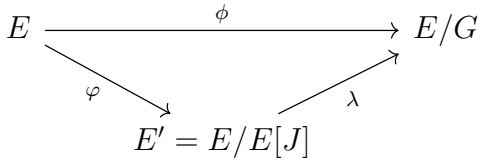
and $\mathfrak{a} := (4, \pi - 1)$ and $\mathfrak{b} := (4, \pi + 1)$ are invertible ideals of norm 4. Clearly $4G = 0$ and either $\pi(P) = P$ for every $P \in G$ or $\pi(P) = -P$ for every $P \in G$, since $\lambda \equiv 1, -1 \mod 4$. Thus, G is contained in either $E[\mathfrak{a}]$ or $E[\mathfrak{b}]$. But $\#E[\mathfrak{a}] = N(\mathfrak{a}) = 4$ and $\#E[\mathfrak{b}] = N(\mathfrak{b}) = 4$ so, in fact, G must be equal to either $E[\mathfrak{a}]$ or $E[\mathfrak{b}]$. □

Lemma 4.32. *Let E and E' be elliptic curves over K, both with endomorphism ring $\mathcal{O}$. Suppose that there is an isogeny $\phi : E \to E'$ over K of degree a power of 2 such that $\ker(\phi)$ is cyclic. Then there is an invertible ideal $I \subseteq \mathcal{O}$ such that $E[I] = \ker(\phi)$.*

Proof. We will prove it by induction on the size of $\ker(\phi)$. The result is clearly true for $\#\ker(\phi) = 1$. Now suppose $\ker\phi$ is not trivial and let $\ker(\phi) = n$. We assume the result for kernel sizes less than n and we prove it for n.

Let $G = \ker(\phi)$. We have that $G \cong \mathbb{Z}/2^k\mathbb{Z}$ is a subgroup defined over K. Since G is not trivial, we have $k \geq 1$. By Lemma 4.29 we have that $k \neq 1$ since E' has endomorphism ring $\mathcal{O}$, so $k \geq 2$. Consider the subgroup $2^{k-2}G$ which is cyclic of order 4 and is defined over K. From Lemma 4.31 we have that there is an invertible ideal $J \subseteq \mathcal{O}$ such that $E[J] = 2^{k-2}G$. Consider the separable isogeny $\varphi : E \to E/E[J]$ over K with kernel $2^{k-2}G$.

What follows is almost identical to the proof of Lemma 4.24: since the kernel of φ is contained in G, there is a separable isogeny $\lambda : E/E[J] \to E/G$ over K such that



commutes. Note that since J is invertible, we have that $E' := E/E[J]$ has endomorphism ring $\mathcal{O}$. Also, $\ker(\lambda)$ is cyclic of size a power of 2 and strictly smaller than $\ker(\phi)$. Hence, we apply the induction hypothesis to the isogeny $\lambda : E' \to E$ to get an invertible ideal $J' \subseteq \mathcal{O}$ such that $E'[J'] = \ker(\lambda)$. We define the invertible ideal $I := JJ'$. From Lemma 4.9 we get that $E[I] = \ker(\phi)$. □

Proposition 4.33. Let E and E' be elliptic curves over K, both with endomorphism ring $\mathcal{O}$. Suppose $\phi : E \to E'$ is an isogeny over K with degree a power of 2. Then there is an invertible ideal $I \subseteq \mathcal{O}$ such that $E[I] = \ker(\phi)$, and thus

$$E' \cong E/E[I].$$

Proof. The proof is identical to Lemma 4.25. Let n be the biggest integer such that $E[2^n] \subseteq \ker(\phi)$ and write $\phi = \lambda \circ [2^n]$. Then $\lambda : E \to E'$ is an isogeny of degree a power of 2 and cyclic kernel. Apply Lemma 4.32 to find $J \subseteq \mathcal{O}$ invertible ideal with $E[J] = \ker(\lambda)$. Let $I = 2^n J$, then $E[I] = \ker(\phi)$ and $E' \cong E/E[I]$. □

The previous proposition is very similar to Corollary 4.27 but has an essential difference. In Corollary 4.27 only the domain curve E is assumed to have endomorphism ring $\mathcal{O}$, and the fact that the codomain E' also has endomorphism ring $\mathcal{O}$ is a consequence of the assumptions. On the other hand, in Proposition 4.33 we assume that both the domain E and codomain E' have endomorphism ring $\mathcal{O}$. If we don't make this assumption, the statement is not true as we have seen in Remark 4.30.

We combine these results to prove the following theorem.

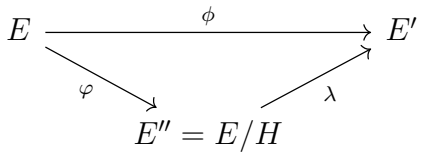
Theorem 4.34. *Let E and E' be elliptic curves over K, both with endomorphism ring $\mathcal{O}$. Suppose $\phi : E \to E'$ is a separable isogeny over K. Then there is an invertible ideal $I \subseteq \mathcal{O}$ such that $E[I] = \ker(\phi)$, and thus*

$$E' \cong E/E[I].$$

Proof. Let $G = \ker(\phi)$. Let G_l denote the Sylow l-subgroup and write

$$G = H \oplus G_2$$

with $H = G_{l_1} \oplus \cdots \oplus G_{l_n}$ where $l_1, \ldots, l_n$ are distinct primes (all different from 2). Since each G_{l_i} is defined over K we have that H is defined over K. Hence, there is a separable isogeny $\varphi : E \to E/H$ defined over K with kernel H and a separable isogeny $\lambda : E/H \to E'$ over K so that the diagram



commutes

Furthermore, by Proposition 4.26 we have that there is an invertible ideal $J \subseteq \mathcal{O}$ such that $E[J] = H$. Let $E'' := E/H$. Since J is invertible we have that E'' has endomorphism ring $\mathcal{O}$. Note that $\#\ker(\lambda) = \#G/\#H = \#G_2$ so λ is an isogeny of degree a power of 2 between two curves with endomorphism ring $\mathcal{O}$. By Proposition 4.33 we conclude that there is an invertible ideal $J' \subseteq \mathcal{O}$ such that $E''[J'] = \ker(\lambda)$. We define $I = JJ'$. It follows that $E[I] = \ker(\lambda \circ \varphi) = \ker(\phi)$. Thus,

$$E' \cong E/E[I]$$

□

Corollary 4.35. The class group action on $Ell(\mathcal{O}, \pi)$ described in Theorem 4.7 is transitive. In other words, if E and E' are elliptic curves with endomorphism ring $\mathcal{O}$, then there is an invertible ideal $I \subseteq \mathcal{O}$ such that $E' \cong E/E[I]$.

Proof. We sketched this argument at the start. Since E and E' have the same trace of Frobenius, there is an isogeny $\phi : E \to E'$ defined over K by Theorem 3.31. Without loss of generality we can assume this isogeny to be separable (just consider the separable part of ϕ, since π is an endomorphism of E). From the previous theorem we get that there is an invertible ideal $I \subseteq \mathcal{O}$ such that

$$E' \cong E/E[I]$$

which concludes the proof. □

In the next chapter we will use this class group action to construct the key exchange protocol CSIDH.

5 CSIDH

We have now developed the mathematical theory that we needed, and hence, we are ready to explain how CSIDH works. We will mainly follow the original paper [1].

CSIDH is based on the action of the abelian group $\mathrm{Cl}(\mathcal{O})$ on the set of supersingular elliptic curves over $\mathbb{F}_p$ with endomorphism ring $\mathcal{O} = \mathbb{Z}[\pi]$ described in Theorem 4.7. Before going into the details of the implementation, let us first give a sketch of how we could use this class group action to achieve a key exchange. It will become apparent that the commutativity of the group $\mathrm{Cl}(\mathcal{O})$ is crucial.

Just as in the previous chapter, we consider a prime $p \geq 5$ such that $p \equiv 3 \mod 8$. We let $\mathcal{O} = \mathbb{Z}[\pi]$ with $\pi^2 + p = 0$, and we fix the field $K = \mathbb{F}_p$. Recall that the class $[I] \in \mathrm{Cl}(\mathcal{O})$ acts on the elliptic curve $E \in Ell(\mathcal{O}, \pi)$ resulting in the curve $E/E[I]$. In order to make this group action more explicit, we will write $[I]E = E/E[I]$.

Let us fix a curve $E_0 \in Ell(\mathcal{O}, \pi)$ which is known to everyone. If Alice and Bob want to exchange a key over a public channel, they do the following:

1. Alice chooses a random element $[\mathfrak{a}] \in \mathrm{Cl}(\mathcal{O})$. This is her *private key*. Similarly, Bob chooses a random element $[\mathfrak{b}] \in \mathrm{Cl}(\mathcal{O})$ for his private key.

2. Alice computes $[\mathfrak{a}]E_0$ and sends it over to Bob (over a possibly public channel). This is Alice's *public key*. Bob proceeds analogously, he computes $[\mathfrak{b}]E_0$ and sends it over to Alice.

3. Upon receiving $[\mathfrak{b}]E_0$, Alice uses her private key to compute $[\mathfrak{a}][\mathfrak{b}]E_0$. Bob, after receiving $[\mathfrak{a}]E_0$, computes $[\mathfrak{b}][\mathfrak{a}]E_0$ using his private key. Since $[\mathfrak{a}][\mathfrak{b}] = [\mathfrak{b}][\mathfrak{a}]$ in $\mathrm{Cl}(\mathcal{O})$, Alice and Bob are now in possession of the same element of $Ell(\mathcal{O}, \pi)$. This is their *shared secret key*.

Note that there's apparently nothing special about the group action that we are using. One might think of using any other action of a finite commutative group G on a set X to construct an analogous key exchange protocol. The elements of G would be the private keys, and X would contain the possible public and shared keys. The issue lies, of course, in the efficiency and the security of such protocol.

In order to implement it efficiently, we need the action of $g \in G$ on $x \in X$ to be fast to compute. To guarantee that the space of shared keys is as big as possible (and hence making a brute-force search attack slower) we would like for this action to be transitive, so that any element $x \in X$ could be reached by the action of G on a fixed element $x_0 \in X$. On the other hand, just as with the discrete logarithm problem, given an element $x_0 \in X$ and an element $x \in X$, it should be hard to find $g \in G$ such that

$gx_0 = x$. To make this search harder, we would like for this group action to be free so that such g is unique.

The group action of $\mathrm{Cl}(\mathcal{O})$ on $Ell(\mathcal{O}, \pi)$ turns out to be, with some caveats, a good choice of G and X, respectively. We know that this group action is free and transitive, but there are still multiple questions that we need to answer in order to sustain this claim. How can we efficiently compute the class group action? How can we represent the elements of $\mathrm{Cl}(\mathcal{O})$ and $Ell(\mathcal{O}, \pi)$? How can we implement this protocol? What makes this protocol secure?

5.1 Computing the class group action

The main reason for choosing to work with supersingular elliptic curves over $\mathbb{F}_p$ is that the trace of Frobenius is zero. This allows us to control the number of rational points $p + 1$ and the characteristic polynomial $X^2 + p$ of the Frobenius by choosing a suitable prime p. This will enable us to make $\mathcal{O}$ have prime ideals for which their action is very easy and fast to compute.

Since we have freedom over the choice of p, we will make the key assumption that $p + 1$ is divisible by many small primes. More specifically, we will suppose p is a large prime of the form $p = 4\ell_1 \dots \ell_n - 1$ with ℓ_i small distinct odd primes. Note that the the characteristic polynomial of π satisfies $X^2 + p \equiv (X - 1)(X + 1) \mod \ell_i$. Hence, by the Kummer-Dedekind Theorem, we have that in $\mathcal{O}$ the ideal (ℓ_i) splits

$$(\ell_i) = (\ell_i, \pi - 1)(\ell_i, \pi + 1) = \mathfrak{l}_i \overline{\mathfrak{l}_i}$$

with $\mathfrak{l}_i = (\ell_i, \pi - 1)$ and $\overline{\mathfrak{l}_i} = (\ell_i, \pi + 1)$. Note that $\mathfrak{l}_i$ and $\overline{\mathfrak{l}_i}$ are the prime ideals above ℓ_i and they both have norm ℓ_i. Also observe that in $\mathrm{Cl}(\mathcal{O})$ we have that $[\mathfrak{l}_i]^{-1} = [\overline{\mathfrak{l}_i}]$. We claim that the action of both $[\mathfrak{l}_i]$ and $[\overline{\mathfrak{l}_i}]$ on a curve $E \in Ell(\mathcal{O}, \pi)$ is fast to compute.

Let $E \in Ell(\mathcal{O}, \pi)$. Note that $E[\mathfrak{l}_i]$ consists of the ℓ_i-torsion points P such that $\pi(P) = P$. That is, $E[\mathfrak{l}_i]$ is the group of rational ℓ_i-torsion points. By Lemma 4.13, we have $\#E[\mathfrak{l}_i] = N(\mathfrak{l}_i) = \ell_i$, so that $E[\mathfrak{l}_i]$ is a cyclic group of order ℓ_i. Hence, any rational point P of order ℓ_i is a generator of $E[\mathfrak{l}_i]$. Finding such $P \in E(K)$ can be done with the following algorithm.

Algorithm 1 Finding a rational point of order ℓ_i

Input: An elliptic curve $E \in Ell(\mathcal{O}, \pi)$ and one of the primes ℓ_i
Output: A point $P \in E(K)$ of order ℓ_i.

1. Randomly pick a point $Q \in E(K)$
2. Set $P \leftarrow [(p + 1)/\ell_i]Q$
3. **If** $P \neq O$ then **return** P

Note that the previous algorithm fails to find P only in the case that ℓ_i does not divide

the order of Q. Since $\#E(K) = p+1$ we have that $E(K) \cong E(K)[4] \times \mathbb{Z}/\ell_1\mathbb{Z} \times \cdots \times \mathbb{Z}/\ell_n\mathbb{Z}$ so that the probability of ℓ_i dividing the order of Q is $(\ell_i - 1)/\ell_1$.

This means that the algorithm has a success probability of $(\ell_i - 1)/\ell_1$, so we expect to find such P in less than two tries, on average. We also note that scalar multiplication in elliptic curves is fast in current computer algebra systems (using double-and-add type algorithms [18, XI.1]), so that Algorithm 1 can be implemented with very low cost. After finding $P = (x, y)$ of order ℓ_i with $x, y \in \mathbb{F}_p$, we can compute the subgroup $E[\mathfrak{l}_i] = \langle P \rangle$ by simply adding P repeatedly.

Now we can use Vélu's formulas in Theorem 1.90 to compute the curve $[\mathfrak{l}_i]E = E/E[\mathfrak{l}_i]$. Looking at these formulas, we see that the number of arithmetic operations in $\mathbb{F}_p$ that one needs to perform in order to compute the codomain of an isogeny with kernel $G \subseteq E(\mathbb{F}_p)$ is linear in $\#G$. Hence, using Vélu's formulas for large groups is unfeasible. But, in this case, since the primes ℓ_i are small (for example, for the implementation in [1] the largest one is 587) and the arithmetic operations all happen in $\mathbb{F}_p$, this computation is also performed quickly. Therefore, we can efficiently compute the action of $[\mathfrak{l}_i]$.

Now we will focus on computing the action of $[\overline{\mathfrak{l}_i}]$. The story is very similar, but with an important distinction. The group $E[\overline{\mathfrak{l}_i}]$ consists of the ℓ_i-torsion points P such that $\pi(P) = -P$. Again, by Lemma 4.13, we have $\#E[\overline{\mathfrak{l}_i}] = N(\overline{\mathfrak{l}_i}) = \ell_i$, so that $E[\overline{\mathfrak{l}_i}]$ is a cyclic group of order ℓ_i. Hence, any point P such that $\pi(P) = -P$ of order ℓ_i is a generator of $E[\mathfrak{l}_i]$.
Note that if $P = (x, y) \in E(\bar{K})$ is a point with $x \in \mathbb{F}_p$ and $y \in \mathbb{F}_{p^2} \setminus \mathbb{F}_p$ then $\pi(P)$ is a point with the same x-coordinate as P, so either $\pi(P) = P$ or $\pi(P) = -P$, but since $y \notin \mathbb{F}_p$ we must have $\pi(P) = -P$. Hence, we can adapt Algorithm 1 to find a point of order ℓ_i with $\pi(P) = -P$ as follows.

Algorithm 2 Finding a point P of order ℓ_i such that $\pi(P) = -P$

Input: An elliptic curve $E \in Ell(\mathcal{O}, \pi)$ and one of the primes ℓ_i
Output: A point $P \in E(\bar{K})$ of order ℓ_i with $\pi(P) = -P$.

1. Randomly pick a point $Q = (x, y) \in E(\bar{K})$ with $x \in \mathbb{F}_p$ and $y \in \mathbb{F}_{p^2} \setminus \mathbb{F}_p$
2. Set $P \leftarrow [(p+1)/\ell_i]Q$
3. **If** $P \neq O$ then **return** P

Note that $\deg(\pi + 1) = p + 1$ by Lemma 3.9, and the endomorphism $\pi + 1$ is separable, so the subgroup of points P such that $\pi(P) = -P$ also has order $p + 1$. The success probability of the above algorithm is again overwhelming.

Just as before, after we've found the desired P, we can compute the subgroup $E[\overline{\mathfrak{l}_i}] = \langle P \rangle$ and use Vélu's formulas to find $[\overline{\mathfrak{l}_i}]E$. Note that this time $E[\overline{\mathfrak{l}_i}] \subseteq E(\mathbb{F}_{p^2})$, so the required arithmetic operations in the formulas must be performed in $\mathbb{F}_{p^2}$, which makes the computations slower than in $\mathbb{F}_p$ (for example, see [21, Theorem 4.3] for complexity of multiplication in finite fields), but still feasible. In fact, we will later state

an adaptation of Vélu's formulas that only involves arithmetic over $\mathbb{F}_p$. We conclude that the action of $[\mathfrak{l}_i]^{-1} = [\overline{\mathfrak{l}_i}]$ can also be computed efficiently.

By repeatedly applying the action of these prime ideals, we can efficiently compute the action of an ideal class represented by an element of the form

$$I = \mathfrak{l}_1^{e_1}\mathfrak{l}_2^{e_2}\dots\mathfrak{l}_n^{e_n} \tag{5.1}$$

with $e_i \in \{-m, \dots, m\}$, where $m \in \mathbb{Z}_{>0}$ is not too big.

Remark 5.1. There is a modification to this method that allows for an even faster computation of the action of an ideal of this form. The details can be found in [1, Algorithm 2].

As we have previously stated, in order for the key exchange protocol to be practical, we need to be able to compute the class group action efficiently. If we sample a random element $[\mathfrak{a}] \in \mathrm{Cl}(\mathcal{O})$ it might not be feasible to compute the action of $[\mathfrak{a}]$ unless the representative of this class is given in the form (5.1).

Hence, in CSIDH, instead of sampling any random elements $[\mathfrak{a}], [\mathfrak{b}] \in \mathrm{Cl}(\mathcal{O})$ for Alice's and Bob's private keys, we restrict ourselves to classes with representatives of the form (5.1). To do this, Alice and Bob each pick a n-tuple $(e_1, \dots, e_n)$ of integers, each sampled randomly from the set $\{-m, \dots, m\}$, which represents the class $[\mathfrak{l}_1^{e_1}\mathfrak{l}_2^{e_2}\dots\mathfrak{l}_n^{e_n}]$.

Definition 5.2. Given a class $[\mathfrak{a}] \in \mathrm{Cl}(\mathcal{O})$, we say that the ideal $I = \mathfrak{l}_1^{e_1}\mathfrak{l}_2^{e_2}\dots\mathfrak{l}_n^{e_n}$ with $e_i \in \{-m, \dots, m\}$ is a *short representation of* $[\mathfrak{a}]$ if $[\mathfrak{l}_1^{e_1}\mathfrak{l}_2^{e_2}\dots\mathfrak{l}_n^{e_n}] = [\mathfrak{a}]$.

Ideally, we want every element of $\mathrm{Cl}(\mathcal{O})$ to be a possible private key. Hence, we would like for every ideal class to have at least one short representation. On the other hand, if an attacker wants to conduct a brute-force search of the private keys by sampling random n-tuples $(e_1, \dots, e_n)$, we want their chances of success to be as low as possible. Thus, the number of short representations of each ideal class should be small. The following heuristic argument gives an approximation on the number of short representations per ideal class. We first make a remark.

Remark 5.3. In cryptography, the size of the objects that depend on the parameters is more naturally measured in bits. In other words, given an object of size N, we care about the number of bits needed to write down N, that is, $\log_2 N$. When estimating the size of some object, we are content with just having a decent approximation of its number of bits since a difference of a couple of bits is not relevant in the asymptotic security analysis.

Heuristic. We start by assuming that $\mathrm{Cl}(\mathcal{O})$ is 'almost cyclic', meaning that $\mathrm{Cl}(\mathcal{O}) \cong H_1 \times H_2$ with H_1 a big cyclic group of order N, where N is not much smaller than $h(\mathcal{O}) = \#\,\mathrm{Cl}(\mathcal{O})$. By a heuristic of Cohen and Lenstra [2, §9], this is likely to be the case. By projecting $\mathrm{Cl}(\mathcal{O})$ to H_1 we can suppose that there is a surjective group homomorphism

$$\rho : \mathrm{Cl}(\mathcal{O}) \twoheadrightarrow (\mathbb{Z}/N, +).$$

Let $\alpha_i = \rho([\mathfrak{l}_i])$ for every $i = 1, \ldots, n$. We make the additional assumption that at least one of the classes $[\mathfrak{l}_i]$ has order N, so that we can take $\alpha_1 = 1$ without loss of generality. Hence, for $[\mathfrak{a}] \in \mathrm{Cl}(\mathcal{O})$ any short representation $[\mathfrak{l}_1^{e_1}\mathfrak{l}_2^{e_2}\ldots\mathfrak{l}_n^{e_n}] = [\mathfrak{a}]$ gives the following linear congruence

$$e_1 + e_2\alpha_2 + \ldots + e_n\alpha_n \equiv \rho([\mathfrak{a}]) \mod N$$

Hence, the number of solutions to this congruence with $e_i \in \{-m, \ldots, m\}$ gives an upper bound to the number of short representations of $[\mathfrak{a}]$.

If we now assume that for a random choice of $(e_1, \ldots, e_n)$ the probability of $e_1 + e_2\alpha_2 + \ldots + e_n\alpha_n$ being equal to $\rho([\mathfrak{a}])$ in $\mathbb{Z}/N\mathbb{Z}$ is approximately $1/N$, then we can expect to have $(2m+1)^n/N$ solutions to the congruence with $e_i \in \{-m, \ldots, m\}$ (a more careful analysis using the Gaussian heuristic for lattices can be found in [1, Section 7.1]).

The ratio $(2m+1)^n/N$ is not much bigger than $(2m+1)^n/h(\mathcal{O})$. So, by choosing m to be the smallest integer such that $(2m+1)^n \geq h(\mathcal{O})$, it's reasonable to expect that the number of classes with short representation is approximately $h(\mathcal{O})$, and that two short representations lie in the same class only very occasionally.

5.2 Representing the elements of $Ell(\mathcal{O}, \pi)$

Recall that in $Ell(\mathcal{O}, \pi)$ two elliptic curves are considered identical if they are K-isomorphic, so that the elements of $Ell(\mathcal{O}, \pi)$ are actually classes of K-isomorphic curves. In order to implement CSIDH, we would like a short way to represent these classes in a computer.

Recall that an elliptic curve over K has a representation in short Weierstrass form $y^2 = x^3 + Ax + B$ with $A, B \in K$. Hence, we could think of representing the class of this elliptic curve with the pair of coefficients $(A, B) \in K^2$. Since two curves with different short Weierstrass equation can be K-isomorphic, these pair of coefficients might not be unique. The problem of unique representation is solved by considering not the short Weierstrass form, but the *Montgomery form* of the curves.

Definition 5.4. An elliptic curve E/K is said to be in *Montgomery form* if it is given by the equation $y^2 = x^3 + Ax^2 + x$, with $A \in K$. We call A the *Montgomery coefficient*.

Remark 5.5. By Proposition 1.37, the projective curve given by the equation $y^2 = x^3 + Ax^2 + x$ is an elliptic curve if and only if $A^2 \neq 4$.

The assumption that $p \equiv 3 \mod 8$ will come into play to guarantee that each element of $Ell(\mathcal{O}, \pi)$ can be represented in Montgomery form in a unique way, and, conversely, that each supersingular elliptic curve in Montgomery form represents an element of $Ell(\mathcal{O}, \pi)$.

For $A \in K$ with $A^2 \neq 4$, let E_A denote the elliptic curve given in Montgomery form by the equation $y^2 = x^2 + Ax^2 + x$. In order to prove certain facts about supersingular elliptic curves of this form, we will need a related curve called the *quadratic twist*.

Definition 5.6. Given the elliptic curve E_A over K, we define its *quadratic twist* to be the curve E_{-A}.

Remark 5.7. Let $\sqrt{-1}$ be an element in $\mathbb{F}_{p^2}$ that squares to -1. Note that $\phi : E_A \to E_{-A}$ given by $\phi(x, y) = (-x, \sqrt{-1}y)$ is an isomorphism defined over $\bar{K}$. It is not defined over K, however, since $\sqrt{-1} \notin \mathbb{F}_p$ as $p \equiv -1 \mod 4$.

Lemma 5.8. *If E_A is a supersingular elliptic curve, then its quadratic twist E_{-A} is also supersingular.*

Proof. One way to see this is by noting that a supersingular curve E/K remains supersingular over any finite extension of K since the condition $E[p] = \{O\}$ is relative only to the algebraic closure $\bar{K}$. Since E_A and E_{-A} are isogenous over $\mathbb{F}_{p^2}$, then the supersingularity of E_A implies that of E_{-A} by Corollary 3.30. We give an alternative proof, which will be useful later, by counting the number of rational points in E_{-A}. Consider the following two equations

$$y^2 = x^3 + Ax^2 + x \tag{5.2}$$
$$y^2 = -(x^3 + Ax^2 + x) \tag{5.3}$$

The first equation is simply the affine equation for E_A. Note that the number of pairs $(x, y) \in K^2$ that satisfy the equation for E_{-A} is exactly the same as the number of pairs satisfying the second equation: if $(x, y) \in E_{-A}(K)$ then $(-x, y)$ satisfies the second equation and vice versa . For each $x \in K$ we consider the solutions to both (5.2) and (5.3). We recall that -1 is not a square in K. Consider $x \in K$. We have three cases:

- **Case 1:** If $x^3 + Ax^2 + x = 0$, then $x = 0$ as the polynomial has distinct roots. Thus, $(0, 0)$ is the only solution to (5.2) and (5.3).
- **Case 2:** If $x^3 + Ax^2 + x$ is a non-zero square in K, then $(x, \pm\sqrt{x^2 + Ax^2 + x})$ are two solutions of (5.2), while such x gives rise to no solutions of (5.3) since $-(x^3 + Ax^2 + x)$ is not a square.
- **Case 3:** If $x^3 + Ax^2 + x$ is not a square in K, then $(x, \pm\sqrt{-(x^2 + Ax^2 + x)})$ are two solutions of (5.3), while such x gives rise to no solutions of (5.2)

Let S_1 be the set of solutions of (5.2) and S_2 the set of solutions of (5.3). By the previous analysis we have that each $x \in K$ contributes 2 to the sum $\#S_1 + \#S_2$. Hence, $\#S_1 + \#S_2 = 2p$. E_A is supersingular, so $\#E_A(K) = p + 1$. Hence, $\#S_1 = p$. It follows that $\#S_2 = p$ and so, including the point at infinity in our count, we have that $\#E_{-A}(K) = p + 1$. □

Lemma 5.9. *Let $A \in K$ and suppose that E_A is a supersingular elliptic curve. Then E has endomorphism ring $\mathcal{O} = \mathbb{Z}[\pi]$.*

Proof. Since E_A is supersingular, it has endomorphism algebra $\mathbb{Q}(\pi)$ with $\pi^2 + p = 0$. The endomorphism ring is an order inside this algebra that includes $\mathbb{Z}[\pi]$, which has conductor $f = 2$. Hence, there only two possibilities: either $\mathrm{End}(E_A) = \mathbb{Z}[\pi]$ or $\mathrm{End}(E_A) = \mathbb{Z}[\frac{\pi-1}{2}]$.

For the sake of contradiction, suppose that $\mathrm{End}(E_A) = \mathbb{Z}[\frac{\pi-1}{2}]$. Consider the endomorphism $\phi = \frac{\pi-1}{2}$. We have that $2\phi = \pi - 1$. Therefore, $E_A[2] \subseteq \ker(\pi - 1) = E_A(K)$. That is, E_A has full rational 2-torsion. To arrive at a contradiction, we analyze the size of $E(K)[4]$ following [3, Table 1].

The curve E_A is given by the equation $y^2 = x^2 + Ax^2 + x$. Since all the 2-torsion points are rational, we have that the polynomial $x^3 + Ax^2 + x = x(x^2 + Ax + 1)$ splits completely over K. Hence, $A^2 - 4$ is a square in K and $\alpha = \frac{-A+\sqrt{A^2-4}}{2} \in K$ is a root of $x^2 + Ax + 1$. Consequently, $E_A[2]$ consist of $\mathcal{O}$, $(0,0)$, $(\alpha, 0)$ and $(1/\alpha, 0)$. Now, since $A^2 - 4$ is a square, we have to possibilities:

- **Case 1:** $A - 2$ and $A + 2$ are both squares. In this case, it's easy to check using the addition formulas that $P = (1, \sqrt{A+2})$ is a rational point of order 4 and that $2P = (0,0)$. Hence, $\langle P \rangle \oplus \langle (\alpha, 0) \rangle$ is a subgroup of $E_2(K)$ of order 8. We conclude that 8 divides $\#E_A(K)$. But this is a contradiction since $\#E_A(K) = p + 1$ and $p \equiv 3 \mod 8$.

- **Case 1:** Neither $A - 2$ nor $A + 2$ are squares. In that case $-A - 2$ and $-A + 2$ are both squares. We can consider the quadratic twist E_{-A} which, by Lemma 5.8, is again supersingular and use the same reasoning as in Case 1 to conclude that 8 divides $\#E_{-A}(K)$ reaching again a contradiction since $\#E_{-A}(K) = p + 1$.

We conclude that $\mathrm{End}(E_A) = \mathcal{O}$. □

Lemma 5.10. *Suppose E has endomorphism ring $\mathrm{End}(E) = \mathcal{O}$. Then there exists $A \in K$ such that $E \cong E_A$.*

Proof. We give a different proof than the one in [1]. Since E has endomorphism ring $\mathcal{O}$, it is supersingular. Note that 4 divides $\#E(K) = p+1$ so that $E(K)$ has a subgroup of order 4. By Lemma 4.28, E doesn't have full 2-torsion, so we can conclude that $E(K)$ has an element of order 4, say $P = (x_0, y_0)$ with $x_0, y_0 \in K$.

There's only one point $(a, 0)$ in $E(K)$ with order 2. We may assume (after applying the K-isomorphism $(x, y) \mapsto (x - a, y)$) that this point is $(0, 0)$. Hence, E has equation

$$y^2 = x(x^2 + a_2x + a_4) \text{ with } a_2, a_4 \in K$$

Note that $a_4 \neq 0$ since the polynomial $x(x^2 + a_2x + a_4)$ cannot have multiple roots. The point $2(x_0, y_0)$ is a rational point of order 2, so the only possibility is $2(x_0, y_0) = (0, 0)$. Using the addition formulas for the x-coordinate, we see that this implies

$$\left(\frac{3x_0^2 + 2a_2x_0 + a_4}{2y_0}\right)^2 - a_2 - 2x_0 = \frac{x_0^4 - 2a_2x_0^2 + a_4^2}{4y_0^2} = 0$$

Hence, $x_0^4 - 2a_2x_0^2 + a_4^2 = (x_0^2 - a_4)^2 = 0$. It follows that $x_0^2 = a_4$, so that a_4 has a square root in K. Note that either x_0 or $-x_0$ is a square in K, denote by $\sqrt{a_4}$ such element. Since $\sqrt{a_4}$ is a square, there is $u \in K$ such that $u^2 = \sqrt{a_4}$. Denote u by $\sqrt[4]{a_4}$. We have

that $(\sqrt[4]{a_4})^2 = \sqrt{a_4}$ and $(\sqrt{a_4})^2 = a_4$. Now, consider the curve E_A with $A = a_2/\sqrt{a_4}$, which has equation

$$y^2 = x^3 + \frac{a_2}{\sqrt{a_4}}x^2 + x.$$

Its easy to check that

$$\phi(x, y) = \left(\frac{1}{\sqrt{a_4}}x, \frac{1}{(\sqrt[4]{a_4})^3}y\right)$$

is an isomorphism $\phi : E \to E_A$. □

Lemma 5.11. *Let $A \in K$ such that E_A is a supersingular elliptic curve. Suppose that $B \in K$ is such that $E_A \cong E_B$. Then $A = B$.*

Proof. We first note that by Lemma 5.9, the curve E_A has endomorphism ring $\mathcal{O}$.

Let $\phi : E_B \to E_A$ with $(x, y) \to (x', y')$ be an isomorphism over K. By Corollary 1.45 we have that there is $u \in K^*$ and $r, s, t \in K$ such that

$$x' = u^2x + r, \quad y' = u^3y + su^2x + t$$

We have that (x', y') satisfies the equation of E_A, while (x, y) satisfies the equation for E_B. Hence, we have

$$\begin{aligned}(u^3y + su^2x + t)^2 &= (u^2x + r)^3 + A(u^2x + r)^2 + u^2x + r \\ y^2 &= x^3 + Bx^2 + x\end{aligned}$$

Subtracting u^6 times the second equation from the first equation, we get

$$\begin{aligned}&2su^5xy + s^2u^4x^2 + 2u^3ty + 2su^2tx + t^2 \\ &\quad = (3u^4r + Au^4 - u^6B)x^2 + (3u^2r^2 + 2Au^2r + u^2 - u^6)x + (r^3 + Ar^2 + r) \quad (5.4)\end{aligned}$$

Note that $\{1, x, y, x^2, xy\}$ are linearly independent elements of $\mathcal{L}(5O)$. Hence, by comparing the coefficient of (5.4) we get information about u, r, s, t. Looking at the coefficients of xy and y we can see that $s = t = 0$. Now, looking at the coefficients of 1, we see that $r^3 + Ar^2 + r = t^2 = 0$. Since E_A has endomorphism ring $\mathcal{O}$, we have from Lemma 4.28 that E_A only has one rational point of order 2, namely $(0, 0)$. Hence, the polynomial $X^2 + AX^2 + X$ only has one root in K, which is 0. It follows that $r = 0$.

From the coefficients of x we have $3u^2r^2+2Au^2r+u^2-u^6 = 2su^2t$. Since $s = t = r = 0$, we conclude that $u^4 = 1$. Since -1 is not a square in K, it follows that $u = \pm 1$ and $u^2 = 1$. Finally, looking at the coefficients of x^2 we have that $s^2u^4 = (3u^4r+Au^4-u^6B)$. Since $s = r = 0$ and $u^2 = 1$, we conclude that $A = B$.

□

Theorem 5.12. *Let E/K be a supersingular elliptic curve. Then* $\mathrm{End}(E) = \mathbb{Z}[\pi]$ *if and only if there exists $A \in K$ such that $E \cong E_A$. If such A exists, it is unique.*

Proof. This follows from Lemmas 5.9, 5.10 and 5.11. □

The previous theorem solves the problem of representation. Any element in $Ell(\mathcal{O}, \pi)$ can be represented uniquely by a single coefficient $A \in K$. Conversely, any $A \in K$ such that E_A is a supersingular elliptic curve represents a unique element in $Ell(\mathcal{O}, \pi)$. We have

$$Ell(\mathcal{O}, \pi) = \{[E_A] \mid A \in K \text{ such that } E_A \text{ is a supersingular elliptic curve } \}$$

where $[E_A]$ denotes the K-isomorphism class of E_A.

5.3 Implementation of CSIDH

We will now make an adaptation to the protocol that we sketched in the beginning of the chapter, which allows for a fast computation of the class group action and a convenient representation of the private, public and shared keys.

Before we proceed, we must settle on a fixed curve in $Ell(\mathcal{O}, \pi)$ to initiate the protocol.

Proposition 5.13. The elliptic curve E_0 over K given by the equation $y^2 = x^3 + x$ has endomorphism ring $\mathcal{O}$.

Proof. By Lemma 5.9 it suffices to show that E_0 is supersingular. Note that the proof of Lemma 5.8 shows that if S_1 is the set of pairs in K^2 that satisfy the affine equation for E_A, and S_2 is the set of of pairs in K^2 that satisfy the affine equation for E_{-A}, then $\#S_1 + \#S_2 = 2p$. The quadratic twist of E_0 is again E_0, so $\#S_1 = \#S_2$. Hence, $\#S_1 = p$, which means that, including the point at infinity, $\#E_0(K) = p + 1$. This shows that E_0 is supersingular. □

Recall Remark 5.3. Note that by the results mentioned in Section 2.2.3, it's reasonable to expect $\mathrm{Cl}(\mathcal{O})$ to be approximately as big as $\sqrt{p}$, since $-p$ is the discriminant of the imaginary quadratic number field $\mathbb{Q}(\pi)$. Hence, we will use $\log_2 \sqrt{p}$ as an estimator for the size of $h(\mathcal{O})$ in bits. This is good enough for our purposes. Hence, by the heuristic in Section 5.2, we should choose $m \in \mathbb{Z}_{>0}$ such that $n \log_2(2m + 1) \approx \log_2 \sqrt{p}$.

The CSIDH protocol

The parameters of the protocol are a large prime $p = 4\ell_1 \dots \ell_n - 1$ where the ℓ_i are distinct odd small primes, the elliptic curve E_0 over K given by the equation $y^2 = x^3 + x$ (by the previous proposition $E_0 \in Ell(\mathcal{O}, \pi)$), and an integer m such that $n \log_2(2m + 1) \approx \log_2 \sqrt{p}$.

1. Alice's private key is an n-tuple $(e_1, \dots, e_n)$ where the integers e_i are sampled randomly from $\{-m, \dots, m\}$, this represents the ideal class $[\mathfrak{a}] = [\mathfrak{l}_1^{e_1} \mathfrak{l}_2^{e_2} \dots \mathfrak{l}_n^{e_n}]$. Bob's public key is also an n-tuple of integers randomly sampled from $\{-m, \dots, m\}$ that represents a class ideal $[\mathfrak{b}]$.

2. Alice efficiently computes $[\mathfrak{a}]E_0$ using the method described in Section 5.1 to compute the action of the class of an ideal with short representation. She can then transform the curve into Montgomery form E_A, with some $A \in K$, using the transformation given in Lemma 5.10. The Montgomery coefficient $A \in K$ is her public key. In an analogous manner, Bob computes $[\mathfrak{b}]E_0$ and finds the unique coefficient $B \in K$ corresponding to its Montgomery form. This Montgomery coefficient is his public key.

3. Alice sends A to Bob and Bob sends B to Alice. Alice computes $[\mathfrak{a}]E_B$ and puts it in Montgomery form. Bob computes $[\mathfrak{b}]E_A$ and puts it in Montgomery form. Since these two curves are isomorphic over K, their Montgomery form is equal by Lemma 5.11. Hence, the corresponding Montgomery coefficient $S \in K$ is their shared secret key.

Remark 5.14. Given a curve $E_A \in Ell(\mathcal{O}, \pi)$, there is in fact a more straightforward method for computing the action of an ideal class $[\mathfrak{l}_i^{\pm 1}]$ on E_A that automatically yields the Montgomery coefficient corresponding to $[\mathfrak{l}_i^{\pm 1}]E_A$. Instead of using the classical Vélu's formulas, one uses the Vélu-type formulas in [13, Proposition 1] that take a curve in Montgomery form E_A and a point P of prime order ℓ as input, and output the codomain of an isogeny with kernel $\langle P \rangle$ in Montgomery form. We state these formulas below.

Let E_A be an elliptic curve over K in Montgomery form. Consider $P \in E_A(\bar{K})$ a point of prime order $\ell \geq 3$. For $1 \leq i < \ell$ let x_i be the x-coordinate of iP. Define

$$\tau = \prod_{i=1}^{\ell-1} x_i, \quad \sigma = \sum_{i=1}^{\ell-1} \left(x_i - \frac{1}{x_i} \right), \quad f(x) = x \prod_{i=1}^{\ell-1} \frac{x x_i - 1}{x - x_i}$$

then

$$\phi : E_A \to E_B, \text{ with } \phi(x, y) = (f(x), c_0 y f'(x))$$

where $B = \tau(A - 3\sigma)$ and $c_0^2 = \tau$, is an isogeny with kernel $\langle P \rangle$.

The curve E_B and the isogeny ϕ are defined over K in the cases that interest us. Indeed, suppose that $\pi(P) = P$ or $\pi(P) = -P$. Then, as we have previously noted, $x_i \in \mathbb{F}_p$ for every $1 \leq i < \ell$. Hence, $\tau, \sigma \in \mathbb{F}_p$, so E_B is defined over K. Also note that $f(x)$ is defined over K. Furthermore, observe that $(\ell - i)P = -iP$ which has the same x-coordinate as iP, therefore $x_i = x_{\ell - i}$ for every $1 \leq i \leq \ell - 1$. We have that $\tau = (\prod_{i=1}^{(\ell-1)/2} x_i)^2$ so $c_0 \in \mathbb{F}_p$. Hence, the isogeny ϕ is defined over K.

Therefore, these formulas may be used by Alice and Bob to compute the action of their private keys without later needing to transform the curve into Montgomery form. They have the additional advantage of only involving arithmetic over $\mathbb{F}_p$.

To show the class group action in practice, we have implemented a toy example in `sage` using these Vélu-type formulas and the 'small' prime $p = 14841476269619 = 4\ell_1 \dots \ell_{11} - 1$, where $\ell_1, \dots, \ell_{11}$ are the smallest 11 odd primes. A Jupyter notebook

is available for download here [1]. We note that significant speed-ups are possible (see Remark 5.1), so this serves only an illustrative purpose.

5.4 Security

We make some remarks about the security of CSIDH. The most obvious way for an attacker to break the protocol would be to simply guess the secret key S, which corresponds to guessing the correct curve in $Ell(\mathcal{O}, \pi)$. Since the class group action is free and transitive, we have that

$$\#Ell(\mathcal{O}, \pi) = \#h(\mathcal{O})$$

Hence, an attacker would have to search a space of roughly $\log_2 \sqrt{p}$ bits. It follows that a brute-force approach has a time complexity of $O(\sqrt{p}) = O(2^{\frac{1}{2}\log_2 p})$, which is exponential in the number of bits of p. This makes the attack unfeasible for large p.

On the other hand, note that if an attacker is able to recover Alice's private key $[\mathfrak{a}]$ from her public key $E_A = [\mathfrak{a}]E_0$, they would be able to compute the shared key by simply computing the action of $[\mathfrak{a}]$ on Bob's public key E_B. Hence, the security of CSIDH relies on the following problem, which is analogous to the discrete logarithm problem in the Diffie Hellman key exchange.

Problem. Given two elliptic curves E and E' over K with endomorphism ring $\mathcal{O}$, find an invertible $\mathcal{O}$-ideal $\mathfrak{a}$ such that $[\mathfrak{a}]E = E'$. It must be possible to compute the action of $[\mathfrak{a}]$ on a curve efficiently.

There is currently no known algorithm (classical or quantum) that can solve this problem in a reasonable amount of time.

A naive first approach is to conduct a *brute-force search*. In the case of CSIDH this corresponds to searching through all possible n-tuples $(e_1, \ldots, e_n)$ with $e_i \in \{-m, \ldots, m\}$, each representing the ideal $[\mathfrak{a}] = [\mathfrak{l}_1^{e_1}\mathfrak{l}_2^{e_2} \ldots \mathfrak{l}_n^{e_n}]$, until one is found such that $[\mathfrak{a}]E = E'$. By our choice of m, this search space is again of size $\log \sqrt{p}$ bits approximately, so this method would require around $\sqrt{p}$ evaluations of the class group action. Hence, it has a time complexity of $O(\sqrt{p})$.

A more sophisticated method would be the *meet-in-the-middle attack*. Consider $\nu \in \{1, \ldots, n\}$ and note that $[\mathfrak{l}_1^{e_1}\mathfrak{l}_2^{e_2} \ldots \mathfrak{l}_n^{e_n}]E = E'$ if and only if

$$[\mathfrak{l}_1^{e_1} \ldots \mathfrak{l}_\nu^{e_\nu}]E = [\mathfrak{l}_{\nu+1}^{-e_{\nu+1}} \ldots \mathfrak{l}_n^{-e_n}]E'$$

Now, instead of searching through all $(2m+1)^n$ tuples $(e_1, \ldots, e_n)$, we split our search space into two. First, for all ν-tuples $(e_1, \ldots, e_\nu)$ with $e_i \in \{-m, \ldots, m\}$ we compute and store the values of $[\mathfrak{l}_1^{e_1} \ldots \mathfrak{l}_\nu^{e_\nu}]E$ alongside the corresponding ν-tuple. Similarly, we compute and store the values of $[\mathfrak{l}_{\nu+1}^{-e_{\nu+1}} \ldots \mathfrak{l}_n^{-e_n}]E'$ for all the $(n-\nu)$-tuples $(e_{\nu+1}, \ldots, e_n)$

[1] `https://drive.google.com/file/d/1V29g6-0yPJmFe_4CFkYBVDhhuPNqluIt/view?usp=sharing`

with $e_j \in \{-m, \ldots, m\}$. We then search for a match between these two lists of stored values. Looking at the corresponding tuples we find the desired n-tuple $(e_1, \ldots, e_n)$.

If we choose $\nu \approx n/2$, then the space of ν-tuples and $(n-\nu)$-tuples is approximately of $(1/2)\log_2 \sqrt{p} = \log_2 \sqrt[4]{p}$ bits. Hence, we require around $\sqrt[4]{p} + \sqrt[4]{p}$ evaluations of the class group action. Therefore, this algorithm has time complexity of $O(\sqrt[4]{p})$, but with a major caveat: we need a lot of storage. There are, however, modifications that allow for better performance [10, Section 3.4].

Explaining the quantum algorithms proposed to attack CSIDH is beyond the scope of this thesis, but we note that the best known quantum attacks are based on algorithms designed to solve the *abelian hidden-shift problem*, and they have estimated complexity of $2^{O(\sqrt{\log_2 p})}$, which is subexponential in the number of bits of p (see [1, Table 1] for complexity estimations of possible quantum attacks). This is tolerable. For example, there are classical subexponential integer-factorization algorithms [7, Chapter 15] that can attack the famous encryption protocol RSA, but this has not stopped its widespread use.

We have answered all of the questions that we stated in the beginning of this chapter. In the quest for a quantum-resistant key exchange protocol, CSIDH –which implements the commutative group action in Theorem 4.7– offers us a possible solution.

Bibliography

[1] Wouter Castryck, Tanja Lange, Chloe Martindale, Lorenz Panny, and Joost Renes. CSIDH: an efficient post-quantum commutative group action. In *International Conference on the Theory and Application of Cryptology and Information Security*, pages 395–427. Springer, 2018.

[2] Henri Cohen and Hendrik W. Lenstra. Heuristics on class groups of number fields. In *Number Theory Noordwijkerhout 1983*, pages 33–62. Springer, 1984.

[3] Craig Costello and Benjamin Smith. Montgomery curves and their arithmetic. *Journal of Cryptographic Engineering*, 8(3):227–240, 2018.

[4] Jean-Marc Couveignes. Hard homogeneous spaces. *Cryptology ePrint Archive*, 2006.

[5] David A. Cox. *Primes of the form $x^2+ ny^2$: Fermat, class field theory, and complex multiplication*, volume 34. John Wiley & Sons, 2011.

[6] Whitfield Diffie and Hellman Martin E. New directions in cryptography. *IEEE Transactions on Information Theory*, 22(6), 1976.

[7] Steven D. Galbraith. *Mathematics of public key cryptography*. Cambridge University Press, 2012.

[8] Ronald S. Irving. *Integers, polynomials, and rings: a course in algebra*. Springer, 2004.

[9] David Jao and Luca De Feo. Towards quantum-resistant cryptosystems from supersingular elliptic curve isogenies. In *International Workshop on Post-Quantum Cryptography*, pages 19–34. Springer, 2011.

[10] Jason LeGrow and Aaron Hutchinson. An analysis of fault attacks on CSIDH. *Cryptology ePrint Archive*, 2020.

[11] Jürgen Neukirch. *Algebraic number theory*, volume 322. Springer Science & Business Media, 2013.

[12] Lorenz Panny. *Cryptography on isogeny graphs*. PhD thesis, Ph.D. thesis, TU Eindhoven, 2021.

[13] Joost Renes. Computing isogenies between montgomery curves using the action of (0, 0). In *International Conference on Post-Quantum Cryptography*, pages 229–247. Springer, 2018.

[14] Igor R. Shafarevich. *Basic Algebraic Geometry 1: Varieties in Projective Space*. Springer Berlin, Heidelberg, 2013.

[15] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM review*, 41(2):303–332, 1999.

[16] Daniel Shumow. Isogenies of elliptic curves: A computational approach. *arXiv preprint arXiv:0910.5370*, 2009.

[17] Carl Siegel. Über die classenzahl quadratischer zahlkörper. *Acta Arithmetica*, 1(1):83–86, 1935.

[18] Joseph H. Silverman. *The arithmetic of elliptic curves*, volume 106. Springer, 2009.

[19] Peter Stevenhagen. *Number rings*. Mastermath course, Leiden University, 2017.

[20] Anton Stolbunov. Public-key encryption based on cycles of isogenous elliptic curves. *MA thesis. Saint-Petersburg State Polytechnical University*, 2004.

[21] Andrew Sutherland. *Lecture notes in Elliptic Curves*. MIT, 2017.

[22] John Tate. Endomorphisms of abelian varieties over finite fields. *Inventiones mathematicae*, 2(2):134–144, 1966.

[23] Jacques Vélu. Isogénies entre courbes elliptiques. *CR Acad. Sci. Paris, Séries A*, 273:305–347, 1971.

[24] John Voight. *Quaternion algebras*. Springer Nature, 2021.

[25] Lawrence C Washington. *Elliptic curves: number theory and cryptography*. Chapman and Hall/CRC, 2008.

[26] William C. Waterhouse. Abelian varieties over finite fields. In *Annales scientifiques de l'École normale supérieure*, volume 2, pages 521–560, 1969.