Computational aspects of class group actions and applications to post-quantum cryptography

Houben, M.R.

Computational aspects of class group actions and applications to post-quantum cryptography

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof. dr. ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op woensdag 28 februari 2024
klokke 16.15 uur

door

Marcel Ronald Houben

geboren te Utrecht, Nederland

in 1995

Promotores:	Dr. T.C. Streng	
	Prof. dr. P. Stevenhagen	
Co-promotor:	Dr. W. Castryck	(KU Leuven)

Promotiecommissie:	Prof. dr. N. Budur	(KU Leuven)
	Prof. dr. ir. G.L.A. Derks	
	Dr. C.R. Martindale	(University of Bristol)
	Dr. B.A. Smith	(Inria / École polytechnique)
	Prof. dr. R.M. van Luijk	

This work was carried out at Leiden University and KU Leuven, funded by Research Foundation – Flanders (FWO) through PhD Fellowship Fundamental Research 11C7322N.

Computational aspects of class group actions and applications to post-quantum cryptography

Marc Houben

Examination committee KU Leuven:

Supervisors:

Dr. Wouter Castryck	(KU Leuven)
Dr. Marco Streng	(Universiteit Leiden)
Prof. dr. Willem Veys	(KU Leuven)

Other members:

Prof. dr. Nero Budur	(KU Leuven)
Prof. dr. Ronald van Luijk	(Universiteit Leiden)
Dr. Chloe Martindale	(University of Bristol)
Dr. Benjamin Smith	(Inria/École Polytechnique)
Prof. dr. Walter Van Assche	(KU Leuven)
Prof. dr. ir. Frederik Vercauteren	(KU Leuven)

Dissertation presented
in partial fulfillment
of the requirements
for the degree of
Doctor of Science (PhD):
Mathematics

February 28, 2024

Contents

Preface

Elliptic curves are mathematical objects that boast rich and deep connections to numerous areas of mathematics; most prominently to number theory and arithmetic geometry. They lie at the basis of many both solved and unsolved mathematical problems, such as Fermat's last theorem and the Birch and Swinnerton-Dyer conjecture. In the real world, elliptic curves have become the standard in many cryptographic protocols, and they are used millions of times per second to establish encrypted communication over the internet. The security of these protocols relies on the assumption that a certain computational problem, called the *discrete logarithm problem*, is difficult. Quantum computers break this assumption, and hence much of today's widely employed cryptography is considered potentially unsafe in the near future. An alternative proposal using elliptic curves is *isogeny-based cryptography.* It bases its security on the computational difficulty of finding a non-trivial map, also called an *isogeny*, between a pair of elliptic curves over a large finite field; a problem considered difficult even for quantum computers.

In Chapter 1, we present background material, consisting mainly of standard definitions and results related to elliptic curves and imaginary quadratic orders.

Chapter 2 provides a short introduction to isogeny-based cryptography, in particular to protocols that are based on class group actions.

In Chapter 3, we provide a summary of the ideas and main results of the following four chapters, all of which are papers jointly written with other authors.

Chapter 4 is a paper written together with Wouter Castryck, Frederik Vercauteren, and Benjamin Wesolowski. We analyze situations in which certain maps on elliptic curves called *pairings* can be used to break a computational hardness assumption called the *decisional Diffie–Hellman problem.*

Chapter 5 is a paper written together with Sam van Buuren, Wouter Castryck, Simon-Philipp Merz, Marzio Mula, and Frederik Vercauteren. We analyze special cases in which pairings can be used to recover a secret isogeny between a pair of elliptic curves over a large finite field.

Chapter 6 is a paper written together with Marco Streng. Motivated by a method to speed up the computation of elliptic curves over finite fields with a prescribed number of points, we devise a multivariate generalization of *Hilbert class polynomials.*

Chapter 7 is a paper written together with Wouter Castryck, Thomas Decru, and Frederik Vercauteren. It concerns a method to make isogeny-based protocols more practical by optimizing the underlying algorithms through *radical isogenies.*

In Chapter 8, we conclude by summarizing the main results and implications of our work. We also highlight some topics of further research.

We thank the Research Foundation – Flanders (FWO) for their support through PhD Fellowship Fundamental Research 11C7322N.

Chapter 1

Preliminaries

This chapter consists mainly of standard definitions and results. It is advised to consult it mainly for reference and to skip (at least) to Chapter 2. The reader familiar with isogeny-based cryptography can immediately skip to Chapter 3, where the main contributions of this work are highlighted. Chapters 4, 5, 6, and 7 have been individually published as research papers.

1.1 Elliptic Curves

Our main references for this section are [2, 4].

1.1.1 Definition and Weierstrass models

Let k be a field. An *algebraic group* over k is an algebraic variety G/k together with a specified point $e \in G(k)$, and morphisms of varieties $G \times G \to G$ (multiplication), and $G \to G$ (inversion), which induce a group structure on $G(\bar{k})$ with respect to which e is the identity element. A morphism of algebraic groups is a morphism of algebraic varieties that is also a homomorphism of groups. Projective connected algebraic groups are called *abelian varieties*; their group structure is always commutative. Abelian varieties of dimension one are called *elliptic curves.* Any smooth projective geometrically irreducible algebraic curve C over k of genus one together with a specified point $e \in C(k)$ admits the structure of an elliptic curve with e as an identity element. Every morphism of varieties between elliptic curves that respects the identity element is a morphism (of algebraic groups). An elliptic curve E over k admits a k-rational isomorphism to a smooth projective planar curve with affine model

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6, \tag{1.1}$$

sending the specified point $e \in E(k)$ to the unique point at infinity. A curve of this form is called a *Weierstrass curve.* If the characteristic of k is not 2 or 3, there furthermore exists an isomorphism to one for which $a_1 = a_2 = a_3 = 0$; i.e. to a *short Weierstrass curve*, whose affine model is given by an equation of the form

$$y^2 = x^3 + Ax + B. \tag{1.2}$$

The *j-invariant* of an elliptic curve E/k is an explicit element of k that can be defined by a rational function in the coefficients of its long Weierstrass model [4, III.1]. For the case of a short Weierstrass equation it is $j(E) := (2^8 \cdot 3^3 \cdot A^3)/(4A^3 + 27B^2)$. Two elliptic curves over an algebraically closed field are isomorphic if and only if their j-invariants are equal.

1.1.2 Isogenies between elliptic curves

Morphisms of elliptic curves that are not constant, or equivalently, have finite kernel, are called *isogenies.* Associated to an isogeny $\varphi : E \to E'$ is the field extension $\varphi^* : \overline{k}(E') \hookrightarrow \overline{k}(E)$ of function fields given by the pull-back map $\varphi^* : f \mapsto f \circ \varphi$. The degree $\deg(\varphi)$ of an isogeny is the degree of this field extension, and an isogeny is called *separable* if this is a separable field extension. An isogeny of degree ℓ is also called an *ℓ-isogeny.* For $N \in \mathbf{Z}$ we denote by $[N] : E \to E$, $P \mapsto NP$ the multiplication-by-N map, and its kernel, which consists of the N-torsion points, by $E[N]$. For every isogeny $\varphi : E \to E'$ there exists a unique isogeny $\hat{\varphi} : E' \to E$, called the *dual isogeny*, such that $\varphi \circ \hat{\varphi} = [\deg \varphi]$. The degree of $[N]$ is N^2. For a field $L \supseteq k$, we write $\mathrm{End}_L(E)$ for the set of endomorphisms of E defined over L, and write $\mathrm{End}(E) := \mathrm{End}_{\bar{k}}(E)$ for the *(full) endomorphism ring.* The set $\mathrm{End}_L(E)$ admits a ring structure, where the multiplication is given by composition and the addition is given (pointwise) by the group operation. This ring is isomorphic to either [4, Cor. III.9.4]

(i) the integers $\mathbf{Z}$;

(ii) an order $\mathcal{O}$ in an imaginary quadratic number field K;

(iii) a maximal order $\mathcal{O}$ in $B_{p,\infty}$, the unique quaternion algebra over $\mathbf{Q}$ ramified at p and infinity, where $p > 0$ equals the characteristic of k.

The ring $\mathrm{End}^0_L(E) := \mathrm{End}_L(E) \otimes_{\mathbf{Z}} \mathbf{Q}$ is called the *endomorphism algebra* of E *over* L. We denote by $\mathrm{End}^0(E) := \mathrm{End}_{\bar{k}}(E) \otimes_{\mathbf{Z}} \mathbf{Q}$ the *(full) endomorphism algebra* of E. Any non-integer endomorphism of an elliptic curve is an imaginary quadratic number, whose norm over $\mathbf{Q}$ as an algebraic integer is equal to its degree as an isogeny.

Lemma 1.1.1 *Let $E/\overline{k}$ be an elliptic curve and let H be a finite subgroup of E. Then there exists a separable isogeny $\varphi : E \to E'$ with domain E and kernel H. This isogeny is unique up to post-composition with an isomorphism. The curve E' is necessarily unique up to isomorphism and we denote it by E/H.*

Proof. [4, Prop. III.4.12]. □

There is a set of formulae by Vélu [6] that can be used to explicitly compute a Weierstrass model for E/H, as well as a description of the associated isogeny $\varphi : E \to E/H$.

Proposition 1.1.2 (Vélu, [6]) *Let $E/\overline{k}$ be an elliptic curve in long Weierstrass form*

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

together with a finite subgroup $H \subseteq E$. Denote by $0 \in E$ the point at infinity. Partition $H \setminus \{0\} = H_2 \sqcup H^+ \sqcup H^-$, where $H_2 = H[2] \setminus \{0\}$ and H^+ and H^- are such that for any $P \in H^+$ it holds that $-P \in H^-$. Write $S = H_2 \cup H^+$, and for $P \in S$ define

$$\begin{aligned}
g_P^x &= 3x(P)^2 + 2a_2x(P) + a_4 - a_1y(P),\\
g_P^y &= -2y(P) - a_1x(P) - a_3,\\
u_P &= (g_P^y)^2, \quad v_P = \begin{cases} g_P^x & \textit{if} \quad P \in H_2,\\ 2g_P^x - a_1g_P^y & \textit{else,}\end{cases}\\
v &= \sum_{P\in S} v_P, \quad w = \sum_{P\in S}(u_P + x(P)v_P),\\
A_1 &= a_1, \quad A_2 = a_2, \quad A_3 = a_3,\\
A_4 &= a_4 - 5v, \quad A_6 = a_6 - (a_1^2 + 4a_2)v - 7w.
\end{aligned}$$

Then an equation for $E' = E/H$ is given by

$$E' : y^2 + A_1xy + A_3y = x^3 + A_2x^2 + A_4x + A_6. \tag{1.3}$$

Furthermore, there exists a separable isogeny $\varphi : E \to E'$ with kernel H that satisfies

$$\begin{aligned}
x(\varphi(Q)) &= x(Q) + \sum_{P\in H\setminus\{0\}} (x(Q+P) - x(P))\\
y(\varphi(Q)) &= y(Q) + \sum_{P\in H\setminus\{0\}} (y(Q+P) - y(P)).
\end{aligned}$$

for all $Q \in E$.

1.1.3 Elliptic curves over finite fields

Let $\mathbf{F}_q$ be a finite field with q elements of characteristic $p := \mathrm{char}(\mathbf{F}_q)$, and let $E/\mathbf{F}_q$ be an elliptic curve with long Weierstrass model $E : y^2+a_1xy+a_3y = x^3+a_2x^2+a_4x+a_6$. The q-Frobenius endomorphism $\pi \in \mathrm{End}(E)$ is given by the map $(x, y) \mapsto (x^q, y^q)$. If $N := \#E(\mathbf{F}_q)$, then we define by $t := q + 1 - N$ the *trace* of $E/\mathbf{F}_q$, also called the *trace of Frobenius*. The *characteristic polynomial* of Frobenius is the polynomial $f := X^2 - tX + q$. It always has π as a root. An elliptic curve for which $p \mid t$ is called *supersingular*, otherwise E is called *ordinary*. If $\pi \notin \mathbf{Z}$, which is guaranteed if E is ordinary, then the characteristic polynomial is the minimal polynomial of π as an algebraic integer. The full endomorphism ring $\mathrm{End}(E)$ of an ordinary elliptic curve is an order in an imaginary quadratic number field. For a supersingular elliptic curve it is a maximal order in the quaternion algebra $B_{p,\infty}$ over $\mathbf{Q}$ ramified at p and ∞.

1.2 Orientations and CM

Our main references for this subsection are [1, 3, 5].

1.2.1 Imaginary quadratic orders

An imaginary quadratic order $\mathcal{O}$ is an order (i.e. a full-rank $\mathbf{Z}$-lattice that is also a subring) in an imaginary quadratic number field K. It is always of the form $\mathcal{O} = \mathbf{Z}[\sigma]$ for some $\sigma \in K$; such σ is called a *generator* for $\mathcal{O}$. The *discriminant* $\operatorname{Disc}(\mathcal{O})$ of $\mathcal{O}$ is the discriminant of σ as an imaginary quadratic number. The discriminant is always a negative integer that is congruent to 0 or 1 (mod 4); such integers are also called *imaginary quadratic discriminants.* For every imaginary quadratic discriminant there exists, up to ring isomorphism, a unique imaginary quadratic order with that discriminant. An imaginary quadratic discriminant that corresponds to a maximal order (i.e. the ring of integers of its field K) is called a *fundamental discriminant.* If $\mathcal{O}_1 \subseteq \mathcal{O}_2$ is an inclusion of imaginary quadratic orders, then $\operatorname{Disc}(\mathcal{O}_1) = v^2 \operatorname{Disc}(\mathcal{O}_2)$, where $v = [\mathcal{O}_2 : \mathcal{O}_1]$ is the index of $\mathcal{O}_1$ in $\mathcal{O}_2$ as an additive abelian group. A *proper* $\mathcal{O}$-ideal $\mathfrak{a}$ is an $\mathcal{O}$-ideal for which $\{\beta \in K \mid \beta\mathfrak{a} \subseteq \mathfrak{a}\} = \mathcal{O}$ (cf. [5, Def. 17.9]). We call two proper $\mathcal{O}$-ideals $\mathfrak{a}$ and $\mathfrak{b}$ *equivalent* if there exists $\beta \in K$ for which $\beta\mathfrak{a} = \mathfrak{b}$. The (ideal) *class group* $\operatorname{Cl}(\mathcal{O})$ of $\mathcal{O}$ is the multiplicative group of proper $\mathcal{O}$-ideals up to equivalence.

1.2.2 Elliptic curves over C

Let $\Lambda \subseteq \mathbf{C}$ be a $\mathbf{Z}$-lattice of rank 2. We define a morphism $\Lambda_1 \to \Lambda_2$ of two such lattices as a complex number $\alpha \in \mathbf{C}$ such that $\alpha\Lambda_1 \subseteq \Lambda_2$, that is

$$\operatorname{Hom}(\Lambda_1, \Lambda_2) := \{\alpha \in \mathbf{C} \mid \alpha\Lambda_1 \subseteq \Lambda_2\}.$$

Under this notion of morphism, there is an equivalence of categories

$$\begin{tikzcd}\left\{\text{Elliptic curves over }\mathbf{C}\right\} \arrow[r, shift left] & \left\{\text{Full rank }\mathbf{Z}\text{-lattices }\Lambda \subseteq \mathbf{C}\right\} \arrow[l, shift left]\end{tikzcd}$$

Any lattice is isomorphic to one of the form $\Lambda = \mathbf{Z} + \tau\mathbf{Z} = \langle 1, \tau\rangle$ for some τ in the complex upper half plane $\mathbf{H}$, and any two such lattices $\Lambda_1 = \langle 1, \tau_1\rangle, \Lambda_2 = \langle 1, \tau_2\rangle$ are isomorphic if and only if τ_1 and τ_2 are in the same $\mathrm{SL}_2(\mathbf{Z})$-orbit, i.e. if and only if there exist integers $a, b, c, d \in \mathbf{Z}$ for which $ad - bc = 1$ and such that $\tau_2 = (a\tau_1 + b)/(c\tau_1 + d)$. This happens if and only if τ_1 and τ_2 admit the same value under the *modular j-function* $j : \mathbf{H} \to \mathbf{C}$.

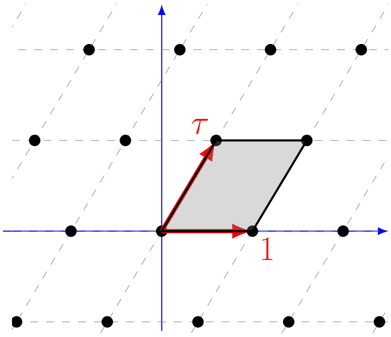


Figure 1.1: A rank two $\mathbf{Z}$-lattice in $\mathbf{C}$ spanned by 1 and $\tau \in \mathbf{H}$.

For most lattices $\Lambda = \langle 1, \tau \rangle$, hence for most elliptic curves over $\mathbf{C}$, we have $\operatorname{End}(\Lambda) \cong \mathbf{Z}$. The exception to this is precisely the special case where τ satisfies a (necessarily imaginary) quadratic equation

$$a\tau^2 + b\tau + c = 0, \quad \text{for some coprime } \ a, b, c \in \mathbf{Z}, a > 0.$$

In that case, $\operatorname{End}(\Lambda) = \mathbf{Z}[a\tau]$, and we say that Λ has *complex multiplication* by the imaginary quadratic order $\mathcal{O} := \mathbf{Z}[a\tau] \subseteq K := \mathbf{Q}(\tau)$. The field $K(j(\tau))$ depends only on $\mathcal{O}$ and is called the *ring class field* of $\mathcal{O}$. The field extension $K(j(\tau))/K$ is abelian and its Galois group is isomorphic to the class group $\operatorname{Cl}(\mathcal{O})$ of $\mathcal{O}$. The *Hilbert class polynomial* associated to τ is

$$\begin{aligned} H_\tau(X) \quad &:= \prod_{\sigma \in \operatorname{Gal}(K(j(\tau))/K)} (X - \sigma(j(\tau))) \\ &= \prod_{\substack{E/\mathbf{C} \text{ ell. curve} \\ \operatorname{End}(E) \cong \mathcal{O}}} (X - j(E)). \end{aligned}$$

The Hilbert class polynomial has integer coefficients and only depends on (the isomorphism class of) $\mathcal{O}$; sometimes it also denoted $H_{\mathcal{O}}(X)$ or $H_D(X)$, where $D := \operatorname{Disc}(\mathcal{O})$.

1.2.3 The CM method

Hilbert class polynomials can be used to construct elliptic curves over finite fields with a prescribed number of points through a procedure known as the *CM method*. Elliptic curves over finite fields always have complex multiplication, in the sense that their endomorphism ring is never isomorphic to $\mathbf{Z}$. Let $E/\mathbf{F}_q$ be an ordinary elliptic curve over a field of characteristic $p > 0$, and let $\pi \in \operatorname{End}(E)$ denote the q-Frobenius endomorphism. Then the minimal polynomial of π is $X^2 - tX + q$, where $N := \#E(\mathbf{F}_q) = q + 1 - t$. Since $\operatorname{End}(E)$ contains the imaginary quadratic order $\mathbf{Z}[\pi]$, it follows that

$$t^2 - 4q = \operatorname{Disc}(\mathbf{Z}[\pi]) = v^2 \operatorname{Disc}(\operatorname{End}(E)), \quad \text{for some } v \in \mathbf{Z}.$$

This motivates the following algorithm

Algorithm 1.2.1 (CM method) Given a prime power $q = p^e$ and an integer $t \in \mathbf{Z}$ coprime to q for which $t^2 - 4q < 0$, find the j-invariant of an elliptic curve $E/\mathbf{F}_q$ with trace of Frobenius t.

1. Find $v \in \mathbf{Z}$ and an imaginary quadratic discriminant D such that $v^2 D = t^2 - 4q$.
2. Compute the Hilbert class polynomial $H_D(X) \in \mathbf{Z}[X]$.
3. Extract a root $j_0 \in \mathbf{F}_q$ of $H_D \pmod{p}$.

The root $j_0 \in \mathbf{F}_q$ corresponds to an $\overline{\mathbf{F}_q}$-isomorphism class of elliptic curves. If one wants to find a curve over $\mathbf{F}_q$ with an exact prescribed trace, one could first construct a curve over $\mathbf{F}_q$ with the j-invariant j_0 output by the above algorithm. Then a twist of this curve will have the desired trace.

1.2.4 Orientations

Let $\mathcal{O} \subseteq K$ be an imaginary quadratic order in an imaginary quadratic number field, and let E be an elliptic curve over a field k. Throughout, we assume that k is perfect. A *K-orientation* on E is an injective ring homomorphism $\iota : K \to \operatorname{End}^0(E)$. An *$\mathcal{O}$-orientation* is a K-orientation for which $\iota(\mathcal{O}) \subseteq \operatorname{End}(E)$. We say that an $\mathcal{O}$-orientation is *primitive* if there does not exist a strict superorder $\mathcal{O}' \supsetneq \mathcal{O}$ in K for which it is also an $\mathcal{O}'$-orientation. This is equivalent to $\iota(\mathcal{O}) = \operatorname{End}(E) \cap \iota(K)$; this last equation associates to each K-oriented elliptic curve a unique order $\mathcal{O}$, called the *primitive order*, with respect to which the orientation is primitive. Given a K-oriented curve (E, ι) and an isogeny $\varphi : E \to E'$, we obtain an induced K-orientation on E' given by

$$\varphi_*(\iota)(\alpha) = \frac{1}{\deg \varphi} \varphi \iota(\alpha) \hat{\varphi}. \tag{1.4}$$

A morphism $(E, \iota) \to (E', \iota')$, also called a *($K$-oriented) isogeny*, of K-oriented elliptic curves is an isogeny $\varphi : E \to E'$ such that $\varphi_*(\iota) = \iota'$.

Lemma 1.2.2 *Let ℓ be a prime number different from* char k*. Let $\varphi : (E, \iota) \to (E', \iota')$ be a K-oriented isogeny of degree ℓ, and let $\mathcal{O}, \mathcal{O}'$ be the respective primitive orders. Then exactly one of the following is true.*

(i) $\mathcal{O} \subsetneq \mathcal{O}'$ and $[\mathcal{O}' : \mathcal{O}] = \ell$, in which case φ is called ascending*;*

(ii) $\mathcal{O} \supsetneq \mathcal{O}'$ and $[\mathcal{O} : \mathcal{O}'] = \ell$, in which case φ is called descending*;*

(iii) $\mathcal{O} = \mathcal{O}'$, in which case φ is called horizontal.

1.3 Bibliography

[1] Leonardo Colò and David Kohel. Orienting supersingular isogeny graphs. *Journal of Mathematical Cryptolology*, 14(1):414–437, 2020.

[2] James S. Milne. *Algebraic groups*, volume 170 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2017.

[3] Hiroshi Onuki. On oriented supersingular elliptic curves. *Finite Fields Appl.*, 69:Paper No. 101777, 18, 2021.

[4] Joseph H. Silverman. *The arithmetic of elliptic curves*, volume 106 of *Graduate Texts in Mathematics*. Springer, second edition, 2009.

[5] Andrew Sutherland. Course notes on elliptic curves, 2017. `https://math.mit.edu/classes/18.783/2017/lectures.html`.

[6] Jacques Vélu. Isogénies entre courbes elliptiques. *Comptes Rendus de l'Académie des Sciences, Série I*, 273:238–241, 1971.

Chapter 2

Isogeny-based cryptography

This chapter is a concise, mostly informal, and highly incomplete introduction to the field of isogeny-based cryptography. Our main focus in this introduction will be on key-exchange protocols; in particular those based on class group actions. Such schemes will be our main concern for the results presented in Chapter 3.

2.1 Diffie–Hellman key exchange

Elliptic curves are widely used in modern-day cryptography. For example, to securely send information in many end-to-end encrypted messaging apps (such as WhatsApp, Signal, and LINE), and to sign transactions in blockchain (such as Bitcoin and Ethereum). The way that messaging apps employ elliptic curves, is through a cryptographic protocol known as a *Diffie–Hellman key exchange.* The idea of such a protocol was first published in 1976 [6]. It is a method for two parties, typically known as *Alice* and *Bob*, to establish a common secret over an insecure channel of communication. Such a common secret is typically subsequently used to encrypt and decrypt information that Alice and Bob would like to send to each other securely (like a WhatsApp message). There are various ways to perform a Diffie–Hellman key exchange, but the main idea is always the same. It can be roughly described as follows. First, Alice and Bob agree publicly on a set X and an element $x_0 \in X$. They also both compute a secret function $X \to X$. Let us say Alice computes (and keeps to herself) the function $f_A : X \to X$, and that Bob does the same for the function $f_B : X \to X$. They each evaluate their secret functions on x_0, and send the result to each other, so that Bob receives $f_A(x_0)$ and Alice receives $f_B(x_0)$. Then, they compute $f_A(f_B(x_0))$ and $f_B(f_A(x_0))$ respectively. If the protocol is designed in such a way that their functions commute, they end up at the same element of X, which is then their common secret.

$$\begin{tikzcd}
\text{Private} & \text{Public} & \text{Private} \\
 & x_0 \in X & \\
f_A & & f_B \\
\text{Alice} \arrow[rr, shift left, "f_A(x_0)"] & & \text{Bob} \arrow[ll, shift left, "f_B(x_0)"] \\
f_A(f_B(x_0)) & & f_B(f_A(x_0))
\end{tikzcd}$$

Figure 2.1: A blueprint for a Diffie–Hellman key-exchange.

As an example, in Elliptic Curve Diffie–Hellman (ECDH), the protocol used by WhatsApp, Alice and Bob publicly agree on an elliptic curve E over a finite field, and a point $P \in E$ of large order n. They hold secret elements $a, b \in \{1, \ldots, n\}$, and they compute and send each other $f_A(P) = aP$ and $f_B(P) = bP$ respectively. Their eventual shared secret is then the point $abP = baP$ on E. For example, in WhatsApp, the elliptic curve used is $E : y^2 = x^3 + 486662x^2 + x$ over the finite field $\mathbf{F}_p$, where $p = 2^{255} - 19$. The public point $P \in E$ is one of the points whose x-coordinate is 9, and has order $n = 2^{252} + 27742317777372353535851937790883648493$.

$$\begin{tikzcd}
\text{Private} & \text{Public} & \text{Private} \\
 & P \in E & \\
a \in \mathbf{Z} & & b \in \mathbf{Z} \\
\text{Alice} \arrow[rr, shift left, "aP"] & & \text{Bob} \arrow[ll, shift left, "bP"] \\
abP & & baP
\end{tikzcd}$$

Figure 2.2: An Elliptic Curve Diffie–Hellman key-exchange.

The security of such a protocol is based on the assumption that it is computationally infeasible to obtain any secret information by only using the publicly available information. More concretely, consider the following set of problems that can be associated to the cryptographic scheme described above.

Problems 2.1.1 (i) *Discrete Logarithm Problem (DLP).* Suppose we are given P and aP. Compute $a \in \{1, \ldots, n\}$.

(ii) *Computational Diffie–Hellman Problem (CDH).* Suppose we are given P, aP, and bP. Compute $abP \in E$.

(iii) *Decisional Diffie–Hellman Problem (DDH).* Suppose we are given either a quadruple $\{P, aP, bP, abP\}$ or a quadruple $\{P, aP, bP, Q\}$, where $Q \in E$ is random. Distinguish, with non-negligible probabilistic advantage, which of the two options it is.[1]

[1] A more formal version of this problem can be found in, e.g. [8, Def. 8.63].

An algorithm to solve problem (i) can be used to solve problem (ii), and an algorithm to solve problem (ii) can be used to solve problem (iii). In a sense, this makes (i) the "hardest" of the problems and (iii) the "easiest". Many security proofs for cryptographic schemes rely on the hypothesis that an analogue of one, or all, of these problems is computationally infeasible. Such a hypothesis is also called a *computational hardness assumption*. In the case of ECDH, it is known that there exists a quantum algorithm that solves (i) in polynomial time. This is known as *Shor's algorithm* [16], and applies to a wide variety of cryptographic schemes that are deployed in practice today, including more classical schemes that are not based on elliptic curves, such as RSA [13]. This means that ECDH and many other protocols used today are insecure in the era of large-scale quantum computing. While it is unknown whether a sufficiently large quantum computer exists, or will exist in the near future, that can be practically used to break schemes that are employed in the real world, this has sparked an area of research known as *post-quantum cryptography*, which looks for ways to encrypt information using algorithms on classical computers that are safe against quantum attacks. One such proposal is *isogeny-based cryptography.*

2.2 Class group action based key exchanges

The primary computational hardness assumption central to essentially all of isogeny-based cryptography is the *isogeny path problem*, and can be described as follows.

Problem 2.2.1 (Isogeny path problem.) *Given a pair of elliptic curves E_0, E_1 over a field k, find an isogeny $\varphi : E_0 \to E_1$.*

Most isogeny-based protocols are not based solely on the pure isogeny path problem as described above, but rather on a version of it that includes some form of extra structure or information. Sometimes, as we will see later, this extra structure turns out to make the protocol insecure. The historically first, and still unbroken, isogeny-based scheme is a Diffie–Hellman key exchange protocol known as CRS, named after Couveignes, Rostovtsev, and Stolbunov [4, 15]. CRS, and later variants such as CSIDH [2] and OSIDH [3], are known as *class group action based key exchange protocols.*

2.2.1 Obtaining a key exchange from a class group action

Let k be a perfect field, and let $\mathcal{O}$ be an imaginary quadratic order. Then there is a free action

$$\mathrm{Cl}(\mathcal{O}) \circlearrowright \{(E, \iota) \mid E/\overline{k} \text{ ell. curve, } \iota : \mathcal{O} \hookrightarrow \mathrm{End}(E) \text{ primitive } \mathcal{O}\text{-orientation}\}/\cong \quad (2.1)$$

of the ideal class group $\mathrm{Cl}(\mathcal{O})$ of $\mathcal{O}$ on the set of primitively $\mathcal{O}$-oriented elliptic curves over $\overline{k}$ up to isomorphism. We will explain in a bit where this comes from, how this action is defined, and how to evaluate it in practice, but let us first see how this gives rise to an idea for a Diffie–Hellman key exchange protocol.

Alice and Bob agree publicly on an elliptic curve E_0/k that is primitively oriented by $\mathcal{O}$. They select secret elements $[\mathfrak{a}]$ and $[\mathfrak{b}]$ of $\mathrm{Cl}(\mathcal{O})$ respectively, and compute $E_A := [\mathfrak{a}]E_0$ and $E_B := [\mathfrak{b}]E_0$. After exchanging E_A and E_B, they both compute $[\mathfrak{b}][\mathfrak{a}]E_0 = E_{AB} = [\mathfrak{a}][\mathfrak{b}]E_0$.

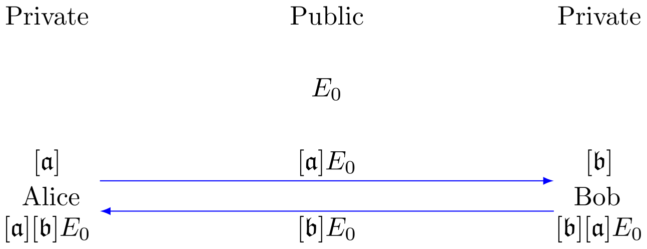


Figure 2.3: A class group action based key exchange.

This gives rise to the following analogues of Problems 2.1.1.

Problems 2.2.2 (i) *Vectorization Problem.* Suppose we are given E_0 and $[\mathfrak{a}]E_0$. Compute $[\mathfrak{a}]$.

(ii) *Computational Diffie–Hellman Problem (CDH).* Suppose we are given E_0, $[\mathfrak{a}]E_0$, and $[\mathfrak{b}]E_0$. Compute $[\mathfrak{a}][\mathfrak{b}]E_0$.

(iii) *Decisional Diffie–Hellman Problem (DDH).* Suppose we are given either a quadruple $\{E_0, [\mathfrak{a}]E_0, [\mathfrak{b}]E_0, [\mathfrak{a}][\mathfrak{b}]E_0\}$ or a quadruple $\{E_0, [\mathfrak{a}]E_0, [\mathfrak{b}]E_0, [\mathfrak{c}]E_0\}$, where $[\mathfrak{c}] \in \mathrm{Cl}(\mathcal{O})$ is a random ideal class. Distinguish, with non-negligible probabilistic advantage, which of the two options it is.

2.2.2 Defining the class group action

We now describe the group action from (2.1) explicitly. Let k be a perfect field of characteristic $p \geq 0$, let $\mathcal{O}$ be an imaginary quadratic order and let E/k be a primitively $\mathcal{O}$-oriented elliptic curve. Note that $\iota(\mathcal{O})$ is then a subring of $\mathrm{End}(E)$. Let $0 \neq \mathfrak{a}$ be an $\mathcal{O}$-ideal such that p does not divide the norm of $\mathfrak{a}$. We define the *kernel of* $\mathfrak{a}$, denoted $E[\mathfrak{a}]$, as

$$\begin{aligned} E[\mathfrak{a}] \quad &:= \quad \bigcap_{\alpha \in \iota(\mathfrak{a})} \ker(\alpha) && (2.2) \\ &= \quad \{P \in E(\overline{k}) \mid \alpha(P) = 0 \quad \forall\, \alpha \in \iota(\mathfrak{a})\}. && (2.3) \end{aligned}$$

The number of elements of $E[\mathfrak{a}]$ equals the norm of the ideal $\mathfrak{a}$. We denote $\mathfrak{a} \cdot E := E/E[\mathfrak{a}]$, and write $\varphi_{\mathfrak{a}} : E \to \mathfrak{a} \cdot E$ for the separable isogeny with kernel $E[\mathfrak{a}]$ (which is unique up to post-composition with an isomorphism by Lemma 1.1.1). Moreover, the isomorphism class of the curve $\mathfrak{a} \cdot E$ together with its orientation induced by (1.4) through $\varphi_{\mathfrak{a}}$ only depends on the ideal class of $\mathfrak{a}$. Since any ideal class contains a representative whose norm is not divisible by p, this defines a group action as in (2.1).

2.2.3 Computing the class group action

We now describe how (part of) the class group action can be explicitly computed in the setting of CRS and CSIDH. Suppose E is defined over a finite field $k = \mathbf{F}_q$ of characteristic p. Let π denote the q-Frobenius endomorphism. Suppose that π is imaginary quadratic (or, equivalently, not an element of $\mathbf{Z}$), and that E is primitively oriented by the imaginary quadratic order $\mathcal{O} := \mathbf{Z}[\pi]$. Denote by $f := X^2 - tX + q$ the characteristic polynomial of Frobenius. If $\ell \neq p$ is a prime number that splits in $\mathcal{O}$, then f splits modulo ℓ, that is

$$f = X^2 - tX + q \equiv (X - \lambda)(X - \mu) \pmod{\ell}.$$

for $\lambda \neq \mu \in (\mathbf{Z}/\ell\mathbf{Z})^\times$. This corresponds to a splitting of the principal $\mathcal{O}$-ideal

$$(\ell) = (\ell, \pi - \lambda)(\ell, \pi - \mu) = \mathfrak{l}\bar{\mathfrak{l}} \tag{2.4}$$

into two ideals $\mathfrak{l} = (\ell, \pi - \lambda)$ and $\bar{\mathfrak{l}} = (\ell, \pi - \mu)$ of norm ℓ. For both of these ideals, the kernel as defined by (2.3) is a subgroup of E of order ℓ, hence corresponds to an ℓ-isogeny with domain E. The orbit of E under the action by the subgroup of $\mathrm{Cl}(\mathcal{O})$ generated by $[\mathfrak{l}]$ is a cycle whose length equals the order of $[\mathfrak{l}] \in \mathrm{Cl}(\mathcal{O})$. We can associate to this cycle a directed graph, whose set of nodes is the orbit of E and whose edges are the ℓ-isogenies corresponding to the ideals $\mathfrak{l}$ and $\bar{\mathfrak{l}}$, i.e. corresponding to the two eigenvalues of Frobenius λ and μ. Since $\mathfrak{l}\bar{\mathfrak{l}} = (\ell)$ is principal, i.e. $[\mathfrak{l}]$ and $[\bar{\mathfrak{l}}]$ are each other's inverses in $\mathrm{Cl}(\mathcal{O})$, the edges belonging to different eigenvalues point in opposite directions along the cycle.

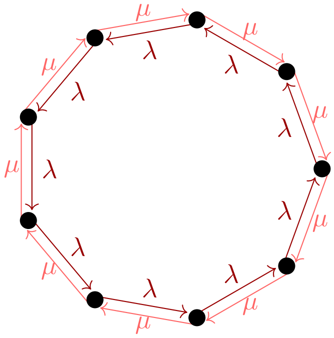


Figure 2.4: Directed cycles associated to ℓ-isogenies corresponding to the eigenvalues of Frobenius λ and μ.

The idea of CRS and CSIDH is now that we compose random walks along these graphs for different small primes ℓ_i. That is, we restrict the class group action to ideals of the form $[\mathfrak{a}] = \prod_i [\mathfrak{l}_i]^{a_i}$, where $\mathfrak{l}_i$ is a prime ideal above ℓ_i and the $a_i \in \mathbf{Z}$ are sampled randomly from predetermined bounded intervals.

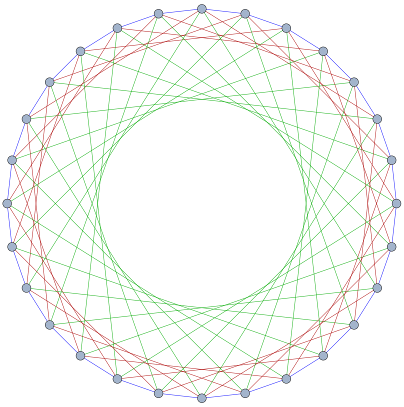


Figure 2.5: A union of three ℓ_i-isogeny graphs.

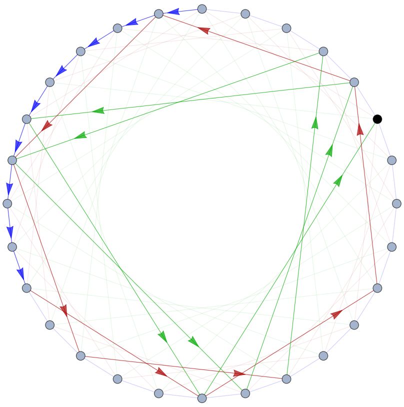


Figure 2.6: An isogeny walk.

If the class group is sufficiently large, these random walks should still give us many different options for, e.g., Alice's public curve $[\mathfrak{a}]E_0$, so that finding an isogeny between the starting and ending curve of the walk exhaustively will remain infeasible.

Now, let us say that we would like to compute the action of the ideal $\mathfrak{l}$ given by (2.4) on E, i.e. one step of the walk in Figure 2.6. According to (2.3), the kernel of $\mathfrak{l}$ is given by

$$E[\mathfrak{l}] = \{P \in E(\overline{k}) \mid P \in E[\ell],\ (\pi - \lambda)(P) = 0\}.$$

Denoting by r the order of $\lambda \in (\mathbf{Z}/\ell\mathbf{Z})^\times$, we see that

$$(\pi - \lambda)(P) = 0 \iff \pi(P) = \lambda P \implies \pi^r(P) = \lambda^r P = P \implies P \in E(\mathbf{F}_{q^r}).$$

It follows that $E[\mathfrak{l}] \subseteq E[\ell](\mathbf{F}_{q^r})$. One way to compute $[\mathfrak{l}]E$ is thus to sample an ℓ-torsion point $P \in E(\mathbf{F}_{q^r})$ whose eigenvalue under the action of Frobenius is λ, and then to compute the codomain of the ℓ-isogeny with kernel $\langle P \rangle$ using Vélu's formulae (1.3). All of this can be done in time polynomial in ℓ and $\log(q)$. In practice, the computational complexity depends heavily (although polynomially) on r, as field arithmetic in larger fields is more expensive. The optimal situation for efficient evaluation of ℓ-isogeny walks is thus the case where the multiplicative orders of the eigenvalues of Frobenius modulo ℓ are as small as possible, i.e. $\lambda = 1, \mu = -1$. This is equivalent to

$$t \equiv 0 \pmod{\ell}, \qquad \text{and} \qquad q + 1 \equiv 0 \pmod{\ell}.$$

Demanding this for many primes ℓ automatically forces $t = 0$ by the Chinese remainder theorem, i.e. the curve E to be supersingular. This gives rise to CSIDH (Commutative Supersingular Isogeny Diffie–Hellman).

Example 2.2.3 (CSIDH-512) Let p be the prime number

$$p := 4 \cdot \underbrace{(3 \cdot 5 \cdot \ldots \cdot 373)}_{\text{73 consecutive primes}} \cdot 587 - 1 \approx 2^{511}. \tag{2.5}$$

Let $E_0/\mathbf{F}_p$ be the supersingular elliptic curve given by $E_0 : y^2 = x^3 + x$. Then $\mathcal{O} := \mathrm{End}_{\mathbf{F}_p}(E_0) = \mathbf{Z}[\pi]$, where $\pi : E_0 \to E_0$ denotes the p-Frobenius, and $\mathcal{O} \hookrightarrow \mathrm{End}(E_0)$ is a primitive orientation. We denote by $\ell_1, \ldots, \ell_{74}$ the odd prime factors of $p+1$, and by $\mathfrak{l}_i := (\ell_i, \pi - 1)$ the $\mathcal{O}$-ideal above ℓ_i corresponding to Frobenius eigenvalue $+1$. This gives rise to the following Diffie–Hellman key exchange procedure.

(i) Alice samples a random element $(a_1, \cdots, a_{74}) \in \{-5, \ldots, 5\}^{74}$, computes $E_A := \prod_i [\mathfrak{l}_\mathfrak{i}]^{a_i} E_0$, and sends E_A to Bob.

(ii) Bob samples a random element $(b_1, \cdots, b_{74}) \in \{-5, \ldots, 5\}^{74}$, computes $E_B := \prod_i [\mathfrak{l}_\mathfrak{i}]^{b_i} E_0$, and sends E_B to Alice.

(iii) Alice computes $\prod_i [\mathfrak{l}_\mathfrak{i}]^{a_i} E_B = \prod_i [\mathfrak{l}_\mathfrak{i}]^{a_i + b_i} E_0$.

(iv) Bob computes $\prod_i [\mathfrak{l}_\mathfrak{i}]^{b_i} E_A = \prod_i [\mathfrak{l}_\mathfrak{i}]^{a_i + b_i} E_0$. ☆

CRS follows the same protocol as CSIDH, but with an ordinary starting curve $E_0/\mathbf{F}_q$. The computational performance of CRS depends heavily on the trace of E_0; one for which the eigenvalues of Frobenius have small multiplicative order modulo many primes ℓ_i is typically better. Other than essentially by exhaustive search, currently no method is known for computing an ordinary elliptic curve over a finite field with both a favorable trace and a large class group (the latter requirement rules out the use of the CM method 1.2.1, since (Hilbert) class polynomials become impractically large; more on this in Chapter 6). Moreover, as explained above, if sufficiently many eigenvalue pairs are "optimal", i.e. ± 1, this forces the curve to be supersingular. As such, all known instantiations of CRS that offer cryptographic levels of security are several orders of magnitude slower than CSIDH, and the protocol is widely considered impractical. However, the CRS scheme is still interesting from a theoretical standpoint, since it is conceivable that the structure of supersingular curves (particularly their additional endomorphisms) might some day be used in an attack against CSIDH.

2.2.4 OSIDH

Oriented Supersingular-Isogeny Diffie–Hellman [3] (OSIDH) is another class group action based protocol, which can be roughly described as follows. Let E_0 be an elliptic curve over a (large) finite prime field $\mathbf{F}_p$ oriented by an imaginary quadratic order $\mathcal{O}$ of class number one. Let $n \in \mathbf{Z}_{>0}$ and let ℓ be a (small) prime number. The idea is for Alice and Bob to act by (secret) elements of the class group of the order $\mathcal{O}_n := \mathbf{Z} + \ell^n \mathcal{O}$ on length-n chains of descending ℓ-isogenies, starting from a given one $E_0 \to E_1 \to \ldots \to E_n$. More specifically, Alice acts by an ideal class $[\mathfrak{a}] = \prod_i [\mathfrak{q}_i]^{a_i}$, for some prime ideals $\mathfrak{q}_i$ of small norm coprime to ℓ and exponents $-r \leq a_i \leq r$ (note the similarity to the ideal class in CSIDH and CRS), to obtain $[\mathfrak{a}] \cdot (E_0 \to E_1 \to \ldots \to$

$E_n) = F_0 \to F_1 \to \ldots \to F_n$. Bob does the same with an ideal class $[\mathfrak{b}] = \prod_i [\mathfrak{l}_i]^{b_i}$ of the same form to obtain $[\mathfrak{b}] \cdot (E_0 \to E_1 \to \ldots \to E_n) = G_0 \to G_1 \to \ldots \to G_n$. Then, instead of exchanging the full descending chains $(F_k)_{0 \leq k \leq n}$ and $(G_k)_{0 \leq k \leq n}$ (which would be insecure; see e.g. [3, Section 5.1]), Alice and Bob publish F_n and G_n together with the action of $[\mathfrak{l}_i]^j$ for all i and all $-r \leq j \leq r$ on F_n and G_n respectively. This is sufficient for Alice and Bob to both be able to compute $[\mathfrak{b}][\mathfrak{a}]E_n = [\mathfrak{a}][\mathfrak{b}]E_n$ (see [3, Section 5.2] for more details). There exist exponential-time attacks against OSIDH [5] that are practical for a large set of parameter choices; in particular for the original proposal of [3] that claimed a security level equivalent to CSIDH-512.

2.3 SIDH

Supersingular Isogeny Diffie–Hellman [7] (SIDH) is an isogeny-based key exchange protocol that does not rely on class group actions. For a long time, SIDH was considered the most promising post-quantum candidate for isogeny-based cryptography. Its major advantage compared to class group action based key exchanges was in its efficiency, and in the fact that the best-known quantum attacks had exponential complexity, whereas CRS and CSIDH are known to admit subexponential quantum attacks [9, 10, 12]. In 2022, classical polynomial time attacks against SIDH were found [1, 11, 14].

Example 2.3.1 (SIKEp503) Let p be the prime number

$$p := 2^{250} \cdot 3^{159} - 1 = 2^a \cdot 3^b - 1. \tag{2.6}$$

Let $E_0/\mathbf{F}_p$ be the supersingular elliptic curve given by $E_0 : y^2 = x^3 + x$. Let $E_0[2^a] = \langle P_A, Q_A \rangle$ and $E_0[3^b] = \langle P_B, Q_B \rangle$.

(i) Alice samples a random integer $m_A \in \{1, \ldots, 2^a\}$, computes $\varphi_A := E_0 \to E_A := E_0/\langle P_A + m_A Q_A \rangle$ and sends $E_A, \varphi_A(P_B), \varphi_A(Q_B)$ to Bob.

(ii) Bob samples a random integer $m_B \in \{1, \ldots, 3^b\}$, computes $\varphi_B := E_0 \to E_B := E_0/\langle P_B + m_B Q_B \rangle$ and sends $E_B, \varphi_B(P_A), \varphi_B(Q_A)$ to Alice.

(iii) Alice computes $E_B/\langle \varphi_B(P_A) + m_A \varphi_B(Q_A) \rangle = E_0/\langle P_A + m_A Q_A, P_B + m_B Q_B \rangle$.

(iv) Bob computes $E_A/\langle \varphi_A(P_B) + m_B \varphi_A(Q_B) \rangle = E_0/\langle P_A + m_A Q_A, P_B + m_B Q_B \rangle$.

☆

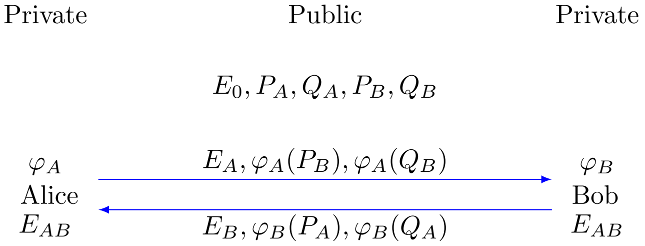

Private
Public
Private
E_0, P_A, Q_A, P_B, Q_B
φ_A
Alice
E_{AB}
$E_A, \varphi_A(P_B), \varphi_A(Q_B)$
$E_B, \varphi_B(P_A), \varphi_B(Q_A)$
φ_B
Bob
E_{AB}

Figure 2.7: Supersingular Isogeny Diffie–Hellman

2.4 Bibliography

[1] Wouter Castryck and Thomas Decru. An efficient key recovery attack on SIDH. In *Eurocrypt 2023 Pt. 5*, volume 14008 of *Lecture Notes in Computer Science*, pages 423–447. Springer, 2023.

[2] Wouter Castryck, Tanja Lange, Chloe Martindale, Lorenz Panny, and Joost Renes. CSIDH: An efficient post-quantum commutative group action. In *Asiacrypt 2018 Pt. 3*, volume 11274 of *Lecture Notes in Computer Science*, pages 395–427. Springer, 2018.

[3] Leonardo Colò and David Kohel. Orienting supersingular isogeny graphs. *Journal of Mathematical Cryptolology*, 14(1):414–437, 2020.

[4] Jean-Marc Couveignes. Hard homogeneous spaces, 2006. Unpublished article, available at `https://eprint.iacr.org/2006/291`.

[5] Pierrick Dartois and Luca De Feo. On the security of OSIDH. In *PKC (1)*, volume 13177 of *Lecture Notes in Computer Science*, pages 52–81. Springer, 2022. `https://ia.cr/2021/1681`.

[6] W. Diffie and M. Hellman. New directions in cryptography. *IEEE Transactions on Information Theory*, 22(6):644–654, 1976.

[7] 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.

[8] J. Katz and Y. Lindell. *Introduction to Modern Cryptography: Principles and Protocols*. Chapman & Hall/CRC, second edition, 2014.

[9] Greg Kuperberg. A subexponential-time quantum algorithm for the dihedral hidden subgroup problem. *SIAM Journal on Computing*, 35(1):170–188, 2005.

[10] Greg Kuperberg. Another subexponential-time quantum algorithm for the dihedral hidden subgroup problem. In *8th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2013)*, volume 22 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 20–34. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2013.

[11] Luciano Maino, Chloe Martindale, Lorenz Panny, Giacomo Pope, and Benjamin Wesolowski. A direct key recovery attack on SIDH. In *Eurocrypt 2023 Pt. 5*, volume 14008 of *Lecture Notes in Computer Science*, pages 448–471. Springer, 2023.

[12] Oded Regev. A subexponential time algorithm for the dihedral hidden subgroup problem with polynomial space, 2004.

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

[14] Damien Robert. Breaking SIDH in polynomial time. In *Eurocrypt 2023 Pt. 5*, volume 14008 of *Lecture Notes in Computer Science*, pages 472–503. Springer, 2023.

[15] Alexander Rostovtsev and Anton Stolbunov. Public-key cryptosystem based on isogenies, 2006. Unpublished article, available at `https://eprint.iacr.org/2006/145`.

[16] Peter W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. *Proceedings 35th Annual Symposium on Foundations of Computer Science*, pages 124–134, 1994.

Chapter 3

Main Results

In this chapter, we summarize and highlight the main ideas and results of the thesis. The full and precise versions of these results appear in joint works presented in the later chapters.

3.1 Pairing-based attacks on class group action based cryptography

This section accompanies two joint research works, corresponding to Chapters 4 and 5. The first is a joint work on breaking the decisional Diffie–Hellman problem for class group action based schemes, together with Wouter Castryck, Frederik Vercauteren, and Benjamin Wesolowski. The second is a joint work on weak instances of the CRS protocol, together with Sam van Buuren, Wouter Castryck, Simon-Philipp Merz, Marzio Mula, and Frederik Vercauteren. We start this section by an introduction of our main tool: *self-pairings* on elliptic curves. Then, we introduce the *isogeny interpolation problem*; a partial solution to this problem turned out to break the SIDH scheme (Example 2.3.1). In Section 3.1.3, we show how these two ideas come together, as we highlight simplified versions of the ideas and main results of the two papers mentioned above.

3.1.1 Self-pairings

Let E/k be an elliptic curve over a field k of characteristic $p \geq 0$ and let $m \in \mathbf{Z}_{>0}$ be such that $p \nmid m$. We denote by $\mu_m \subseteq \overline{k}^\times$ the m-th roots of unity. Pairings are bilinear maps that send a pair of points on (subgroups of) two isogenous elliptic curves over k to the unit group $\overline{k}^\times$ of the algebraic closure of the field of definition of the curves. The typical example is the *Weil-pairing* $e_m : E[m] \times E[m] \to \mu_m \subseteq \overline{k}^\times$ [11, III.8]. It is an alternating, Galois-invariant, non-degenerate bilinear map, which is compatible with isogenies $\varphi : E \to E'$ in the sense that

$$e_m(\varphi(P), \varphi(Q)) = e_m(P, Q)^{\deg \varphi} \tag{3.1}$$

for all $P, Q \in E[m]$. We study the following notion related to pairings.

Definition 3.1.1 A *self-pairing* on a subgroup G of an elliptic curve E/k is a map $f : G \to \overline{k}^{\times}$ such that

$$f(\lambda P) = f(P)^{\lambda^2} \tag{3.2}$$

for all $\lambda \in \mathbf{Z}$ and all $P \in G$. △

Note that any bilinear map $e : G \times G \to \overline{k}^{\times}$ gives rise to a self-pairing $f : G \to \overline{k}^{\times}$, $P \mapsto e(P, P)$.

Example 3.1.2 Let $e_m : E[m] \times E[m] \to \mu_m$ denote the Weil pairing. If $\tau \in \mathrm{End}(E)$ is any endomorphism, then we obtain a self-pairing

$$f : E[m] \to \mu_m,\ P \mapsto e_m(P, \tau(P)), \tag{3.3}$$

which we call the (τ-)*twisted Weil self-pairing.* ☆

Example 3.1.3 Let E/K be an elliptic curve over a number field K. Then the canonical (Néron–Tate) height $\hat{h}_K : E \to \mathbf{R}$ satisfies (3.2), but is not a self-pairing, since $\mathbf{R} \not\subseteq \overline{K}^{\times}$. ☆

Similar to Equation (3.1), self-pairings have a notion of compatibility with isogenies.

Definition 3.1.4 Let E, E' be elliptic curves over k with self-pairings $f : G \to \overline{k}^{\times}$, $f' : G' \to \overline{k}^{\times}$ on subgroups $G \subseteq E$, $G' \subseteq E'$. Let $\varphi : E \to E'$ be an isogeny. We say that the self-pairings f, f' are *compatible* with φ if

$$\varphi(G) \subseteq G', \quad \text{and} \quad f'(\varphi(P)) = f(P)^{\deg(\varphi)} \tag{3.4}$$

for all $P \in G$. △

Example 3.1.5 Consider the twisted Weil self-pairing f of Example 3.1.2, and let $\varphi : E \to E$ be any endomorphism that commutes with τ. Then $\varphi(E[m]) \subseteq E[m]$ and

$$f(\varphi(P)) = e_m(\varphi(P), \tau(\varphi(P))) = e_m(\varphi(P), \varphi(\tau(P))) = e_m(P, \tau(P))^{\deg(\varphi)} = f(P)^{\deg(\varphi)},$$

hence f is compatible with φ. ☆

Definition 3.1.6 For a self-pairing $f : G \to \overline{k}^{\times}$ on a finite subgroup G, the *order* of f is the smallest positive integer m such that $f(G) \subseteq \mu_m$. △

3.1.2 Isogeny interpolation

Let $\varphi : E_0 \to E_1$ be an isogeny of degree d. It is an elementary result that φ is fixed once its images under $N > 4d$ points are known.

Lemma 3.1.7 *[14, Lemma 3.1]. Let $\varphi_1, \varphi_2 : E_0 \to E_1$ be isogenies of degree $\leq d$. Suppose that $\#\ker(\varphi_1 - \varphi_2) > 4d$. Then $\varphi_1 = \varphi_2$.*

The *isogeny interpolation problem* asks to recover an isogeny given sufficiently many image points. This problem can be effectively solved in certain cases.

Theorem 3.1.8 *Let E_0, E_1 be elliptic curves over a finite field of characteristic $p > 0$. Let $\varphi : E_0 \to E_1$ be an isogeny of known degree d coprime to p. Let $N \in \mathbf{Z}_{>0}$ such that $N^2 > 4d$ and $\gcd(N, d) = 1$. Suppose that we are given either of the following:*

(i) the images $\varphi(P), \varphi(Q)$ for a basis P, Q of $E[N]$; or

(ii) the image $\varphi(P)$ for a point $P \in E[N^2]$ of order N^2.

Then one can recover φ in polynomial time.

Proof. Case (i) is by Damien Robert [9, Thm. 1.1], following ideas from Castryck, Decru [4], Maino, and Martindale [8]. Case (ii) follows from case (i) by a reduction argument first proposed by Luca De Feo. A sketch of this argument appears in our joint work in Chapter 5; see below Remark 5.6.1. □

This result allows to break SIDH in polynomial time. For instance, in Example 2.3.1, by applying the theorem to $\varphi = \varphi_A, P = P_B, Q = Q_B, d = 2^a$, and $N = 3^b$.

3.1.3 Main contributions

Weak instances of CRS

We now highlight results of joint work with Sam van Buuren, Wouter Castryck, Simon-Philipp Merz, Marzio Mula, and Frederik Vercauteren. The full version of this work can be found in Chapter 5.

The main obstruction in applying the Isogeny Interpolation Theorem 3.1.8 to schemes that are based on class group actions, such as CSIDH and CRS, is that in such schemes no image points under the secret isogeny are shared. We studied whether it is possible to use self-pairings to obtain information about image points anyway, in an attempt to make class group action schemes vulnerable to the same attacks that broke SIDH. In what follows, we assume for ease of exposition that m is an *odd* integer. The following is a relatively straightforward result about self-pairings.

Lemma 3.1.9 *Let $f : G \to \overline{k}^\times$ be a self-pairing and let $P \in G$ of odd order m. Then $f(P)$ is an m-th root of unity.*

Proof. See Lemma 5.4.6. □

We call a self-pairing on a subgroup G for which $\#G = m$ *primitive* if its image contains a primitive m-th root of unity. By the lemma, this implies that G is cyclic. The main idea of how self-pairings could be used to obtain information about image points is as follows.

Idea 3.1.10 Let $\varphi : E \to E'$ be an unknown isogeny of known degree d. Suppose that we have primitive self-pairings $f : C \to \mu_m$ and $f' : C' \to \mu_m$ on cyclic subgroups $C = \langle P \rangle \subseteq E$ and $C' = \langle P' \rangle \subseteq E'$ of order m that are compatible with φ. Then, since $\varphi(P) \in C'$, it follows that $\varphi(P) = \lambda P'$ for some $\lambda \in \mathbf{Z}$. Now, since

$$f(P)^d = f'(\varphi(P)) = f'(P')^{\lambda^2}, \tag{3.5}$$

we can determine $\lambda^2 \pmod{m}$ from the values of $f(P)$, $f'(P')$, and d, by a discrete logarithm computation in μ_m.

Knowing $\lambda^2 \pmod{m}$, the idea is then to guess $\lambda \pmod{m}$, and hence $\varphi(P)$. Then, if m is large, smooth, and square, we can recover φ using case (ii) of Theorem 3.1.8. The next natural question is when such self-pairings exist. If the self-pairings are assumed to be compatible with isogenies coming from a class group action (i.e. if the above attack strategy applies to CSIDH and CRS), then our main result gives a complete classification.

Theorem 3.1.11 (van Buuren, Castryck, Houben, Merz, Mula, Vercauteren) *Let k be a field of characteristic $p \geq 0$, let $\mathcal{O}$ be an imaginary quadratic order of discriminant D, and let $m \in \mathbf{Z}_{>0}$ be odd and such that $p \nmid m$. Primitive self-pairings of order m compatible with $\mathcal{O}$-oriented isogenies (through the recipe of Section 2.2.2) exist if and only if $m \mid D$.*

Proof. See Prop 5.4.8 and Section 5.5. □

In CSIDH and CRS, we have $\mathcal{O} = \mathbf{Z}[\pi]$, which has discriminant $t^2 - 4q$. For CSIDH we have that $t = 0$ and $q = p$ is prime, hence the discriminant does not contain large smooth square factors, and CSIDH remains completely insusceptible to isogeny interpolation combined with self-pairings. We show, however, that exceptionally weak instances of CRS admit polynomial time key-recovery attacks. See Example 5.6.4 for an explicit such weak instance.

On the Decisional Diffie–Hellman problem

We now highlight results of joint work with Wouter Castryck, Frederik Vercauteren, and Benjamin Wesolowski. The full version of this work can be found in Chapter 4.

The pairing-based attack described above solves the Vectorization Problem 2.2.2(i). It was shown by Castryck, Sotáková, and Vercauteren [6] that there are cases in which the Decisional Diffie–Hellman Problem 2.2.2(iii) can be solved in classical polynomial time using Tate pairings. We present a new approach based on the Weil pairing that is more general, conceptually simpler, and oftentimes more efficient than the previous method.

In what follows, we denote by $m \in \mathbf{Z}_{>0}$ an odd *prime number*. Suppose $f : G \to \mu_m$ is a self-pairing on a subgroup of an elliptic curve E. We define an equivalence relation on $\mu_m \setminus \{1\}$ by setting $x \sim y \iff \exists \lambda \in \mathbf{Z} : y = x^{\lambda^2}$. This partitions $\mu_m \setminus \{1\}$ into two equivalence classes S_1, S_2. There are now four options for the image of G under

f. Indeed, $f(G)$ equals either $\{1\}$, μ_m, $\{1\} \cup S_1$, or $\{1\} \cup S_2$. In the last two cases, we call f *ramified.*

Idea 3.1.12 Suppose that $\varphi : E \to E'$ is an unknown isogeny of (unknown) degree d coprime to m. Suppose that we have ramified self-pairings $f : G \to \mu_m$ and $f' : G' \to \mu_m$ on subgroups $G \subseteq E$ and $G' \subseteq E'$ that are compatible with φ. Let $P \in E$ and $P' \in E'$ be such that $f(P)$ and $f'(P')$ are primitive m-th roots of unity. Since f' is ramified, we find that

$$f(P)^d = f'(\varphi(P)) \sim f'(P').$$

It follows that we can determine whether d is a quadratic or non-quadratic residue modulo m by computing whether or not $f(P) \sim f'(P')$.

To check whether $f(P) \sim f'(P')$, a discrete logarithm computation in μ_m followed by a Legendre symbol computation suffices. Indeed,

$$\left(\frac{\log_{f(P)} f'(P')}{m} \right) = \begin{cases} 1 & \text{if } f(P) \sim f'(P'); \\ -1 & \text{if } f(P) \not\sim f'(P'). \end{cases}$$

Equivalently,

$$\left(\frac{d}{m} \right) = \left(\frac{\log_{f(P)} f'(P')}{m} \right).$$

Our main result classifies when ramified self-pairings compatible with isogenies coming from a class group action exist.

Theorem 3.1.13 (Castryck, Houben, Vercauteren, Wesolowski) *Let k be a field, let $\mathcal{O}$ be an imaginary quadratic order of discriminant D, and let $m \in \mathbf{Z}_{>0}$ be an odd prime number different from* $\operatorname{char} k$. *Ramified self-pairings of order m compatible with $\mathcal{O}$-oriented isogenies (through the recipe of Section 2.2.2) exist if and only if $m \mid D$. In that case, an explicit family of ramified self-pairings is given by the twisted Weil pairing*

$$f : E[m] \to \mu_m,\ P \mapsto e_m(P, \sigma(P)),$$

for some generator σ of $\mathcal{O} = \mathbf{Z}[\sigma]$ of norm coprime to m.

Proof. See Theorem 4.1.1. □

This leads to the following attack strategy against the Decisional Diffie–Hellman Problem 2.2.2(iii). Using ramified self-pairings of order m compatible with the class group action, we can determine whether the norm of (a representative of) a connecting ideal class $[\mathfrak{a}]$, or equivalently, the degree of an isogeny $\varphi_{\mathfrak{a}} : E_0 \to [\mathfrak{a}]E_0$, is a square or not modulo m. A Diffie–Hellman quadruple $E_0, [\mathfrak{a}]E_0, [\mathfrak{b}]E_0, [\mathfrak{a}][\mathfrak{b}]E_0$ now yields the verifiable equality

$$\left(\frac{N(\mathfrak{a})}{m} \right) \left(\frac{N(\mathfrak{b})}{m} \right) = \left(\frac{N(\mathfrak{a}\mathfrak{b})}{m} \right).$$

However, if $\mathcal{O}$ has ideal classes of both square and non-square norm modulo m, this equality should fail with probability 50% when $[\mathfrak{a}][\mathfrak{b}]$ is replaced by a random ideal class $[\mathfrak{c}] \in \mathrm{Cl}(\mathcal{O})$, thus giving a non-negligible distinguishing advantage.

3.2 Generalized class polynomials

This section accompanies joint work with Marco Streng on *generalized class polynomials*, corresponding to Chapter 6. After introducing (classical) class polynomials, we present and motivate the definition of a multivariate analogue. We then highlight the main results of our joint paper.

3.2.1 Class polynomials

Recall from Section 1.2.2) that the *Hilbert class polynomial* associated to an imaginary quadratic number τ in the complex upper half plane $\mathbf{H}$ is defined as

$$H_\tau(X) = \prod_{\sigma \in \mathrm{Gal}(K(j(\tau))/K)} (X - \sigma(j(\tau))) \in \mathbf{Z}[X]. \tag{3.6}$$

Hilbert class polynomials can be used to construct elliptic curves over finite fields with a prescribed number of points through the CM method; Algorithm 1.2.1. The bottleneck in this algorithm is in the computation of the Hilbert class polynomial; the main reason being that its coefficients are typically large, as illustrated by the following example.

Example 3.2.1 Let $\tau \in \mathbf{H}$ be an imaginary quadratic number of discriminant $D = -103$. Then

$$\begin{aligned} H_\tau(X) \quad = \quad & X^5 + 70292286280125X^4 + 85475283659296875X^3 \\ & + 4941005649165514137656250000X^2 \\ & + 13355527720114165506172119140625X \\ & + 28826612937014029067466156005859375. \end{aligned}$$

☆

For larger discriminants the situation gets worse rather quickly. For "typical" discriminants of size 10^9 the total size is already in the gigabytes [13]. One possible idea to remedy this, is to replace the j-function in (3.6) by a different modular function, in the hope that the resulting polynomial will have smaller coefficients. The resulting more general notion of class polynomial is captured by the following definition.

Definition 3.2.2 Let f be a modular function and $\tau \in \mathbf{H}$ imaginary quadratic. If $f(\tau) \in K(j(\tau))$ then we call (f, τ) a *class invariant*, and we define the associated class polynomial by

$$H_\tau[f](X) = \prod_{\sigma \in \mathrm{Gal}(K(f(\tau))/K)} (X - \sigma(f(\tau))).$$

△

Note that $K(f(\tau))/K$ is indeed automatically Galois, since $K(j(\tau))/K$ is an abelian extension.

Example 3.2.3 Let $\mathfrak{f}(z) = \zeta_{48}^{-1}\eta(\frac{z+1}{2})/\eta(z)$, where ζ_{48} is a primitive 48-th root of unity, and

$$\eta(z) = q^{1/24}\prod_{n=1}^{\infty}(1-q^n), \quad \text{where} \quad q = \exp(2\pi i z) \tag{3.7}$$

is the Dedekind η-function. Let τ be as in Example 3.2.1. Then

$$H_\tau[\mathfrak{f}](X) = X^5 + 2X^4 + 3X^3 + 3X^2 + X - 1. \tag{3.8}$$

☆

The modular function $\mathfrak{f}$ from the example is known as *Weber's function* (well, one of three such functions [15, §34]). It is related to the modular j-function by the equation

$$(\mathfrak{f}^{24} - 1)^3 - j\mathfrak{f}^{24} = 0. \tag{3.9}$$

For any τ of discriminant $\equiv 1 \pmod 8$, we have that $(\mathfrak{f}, \tau)$ is a class invariant. The resulting class polynomials can be used in place of Hilbert class polynomials in the CM method; the only extra step one needs is to compute a j-invariant from an "$\mathfrak{f}$-invariant" using (3.9). The phenomenon that the Weber function yields smaller class polynomials can be explained through the following definition.

Definition 3.2.4 The *reduction factor* of a modular function f of level N is

$$r(f) = \frac{\deg(j : X(N) \to \mathbf{P}^1)}{\deg(f : X(N) \to \mathbf{P}^1)}.$$

△

At imaginary quadratic $\tau \in \mathbf{H}$, the value of the modular j-function $j(\tau)$ is an algebraic number (in fact, an algebraic integer). We denote by $h(j(\tau))$ its logarithmic height. For a (possibly multivariate) nonzero polynomial F over $\mathbf{C}$, we denote by $|F|_\infty \in \mathbf{R}_{>0}$ the maximum of the absolute values of its coefficients.

Proposition 3.2.5 *Let f be a modular function that has a q-expansion with coefficients in a number field, and let $\tau_1, \tau_2, \ldots \in \mathbf{H}$ be a sequence of imaginary quadratic* *numbers such that $h(j(\tau_i)) \to \infty$. Suppose that $K(f(\tau_i)) = K(j(\tau_i))$ for all i. Then*

$$\frac{\log |H_{\tau_i}[j](X)|_\infty}{\log |H_{\tau_i}[f](X)|_\infty} \to r(f). \tag{3.10}$$

Proof. This result follows from [7, Prop. B.3.5]; see the argument on the bottom of page 9 of [2]. □

Essentially, the proposition says that for a modular function f, asymptotically, the bitsize of the largest coefficient of its class polynomials are a factor $r(f)$ less than that of Hilbert class polynomials. For Weber's function $\mathfrak{f}$, the reduction factor is 72, which means that asymptotically we would require about 72 times fewer digits to write down the largest coefficient of a class polynomial for $\mathfrak{f}$ when compared to j. No modular function with a reduction factor larger than 72 is known. In fact, according to the following result, we cannot do much better.

Theorem 3.2.6 (Bröker–Stevenhagen, 2008) *Let f be a modular function. Then $r(f) \leq 32768/325 \approx 100.82$.*

Proof. See [2, Thm. 4.1]. □

The upper bound on $r(f)$ can be further improved to 96 if one assumes Selberg's eigenvalue conjecture [10].

3.2.2 Main contributions

We now highlight results of joint work with Marco Streng. The full version of this work can be found in Chapter 6.

Since the reduction factors of class polynomials are limited by the Bröker-Stevenhagen bound, we considered a multivariate extension that we call *generalized class polynomials*. A univariate polynomial over a field k can be seen as a function on the projective line $\mathbf{P}^1(k)$ whose poles are restricted to the unique point at infinity. Class polynomials can thus be described, up to a multiplicative constant, by their divisor as a function on $\mathbf{P}^1$.

$$\operatorname{div} H_\tau[f] = \sum_{\sigma \in G} (\sigma(f(\tau))) - \#G(\infty), \qquad \text{where } G = \operatorname{Gal}(K(f(\tau))/K).$$

In other words, for a class invariant (f, τ), the class polynomial represents a function on $\mathbf{P}^1$ that has a simple zero at every element of the Galois orbit of $f(\tau)$, and a pole at the unique point at infinity; see Figure 3.1.

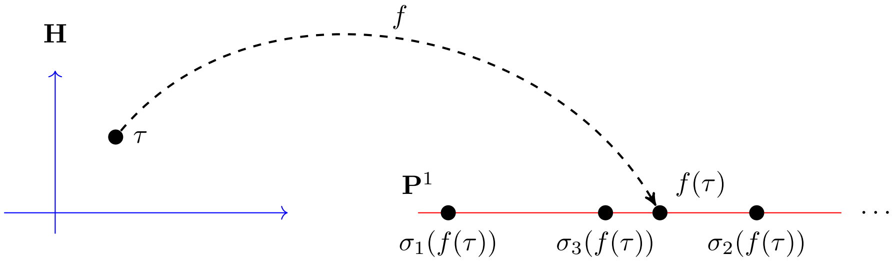


Figure 3.1: The Galois orbit of a class invariant.

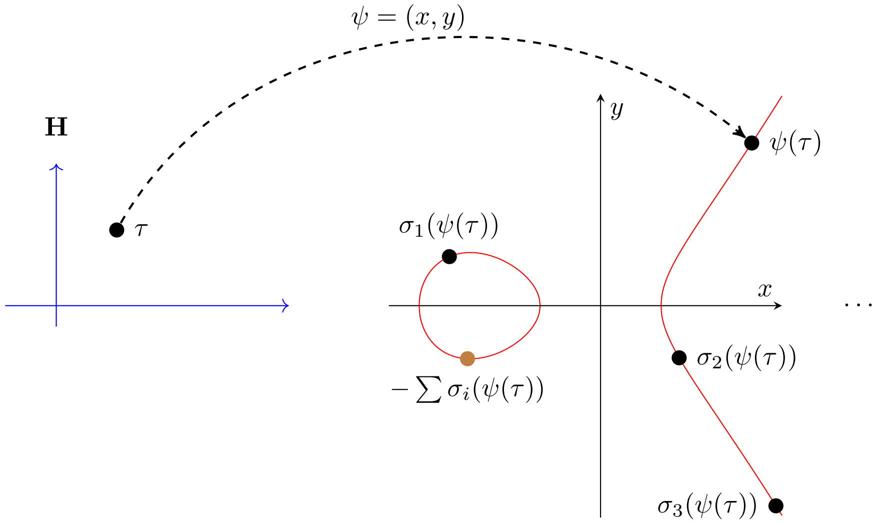


Figure 3.2: The Galois orbit of a pair of class invariants satisfying the equation of an elliptic curve.

Now suppose that we are given, instead of one class invariant, a pair $(x, \tau), (y, \tau)$ of class invariants. Since any pair of modular functions is algebraically dependent, the modular functions x, y satisfy the equation of a (possibly singular) planar curve. Let us suppose for simplicity that this is an elliptic curve E given by a Weierstrass equation

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

Then $\psi(\tau) := (x(\tau), y(\tau))$ defines a point on $E(K(j(\tau))$, hence we can again consider its Galois orbit. Setting $G := \mathrm{Gal}(K(x(\tau), y(\tau))/K)$, the *generalized class polynomial* $H_\tau[E] \in K[X, Y]$ is now defined, uniquely up to a non-zero multiplicative scalar, by its divisor

$$\operatorname{div} H_\tau[E] = \sum_{\sigma\in G}\Big(\sigma(\psi(\tau))\Big) + \underbrace{\Big(-\sum_{\sigma\in G}\sigma(\psi(\tau))\Big)}_{\text{sum on } E} - (\#G+1)\Big(\infty\Big),$$

where $\infty \in E$ denotes the unique point at infinity. Note that the extra term (the negative of the sum of the points in the Galois orbit) is now necessary to ensure that the resulting divisor is principal; see Figure 3.2.

Example 3.2.7 Consider the modular curve $E = X^0_+(119)$; i.e. the quotient of $X^0(119) = \mathbf{H}/\Gamma^0(119)$ by the Fricke-Atkin-Lehner involution $\tau \mapsto -119/\tau$. Then E is an elliptic curve, and a modular parametrization is given by

$$E : y^2 + 3xy - y = x^3 - 3x^2 + x,$$

where x, y are modular functions for $\Gamma^0(119)$ with q-expansions

$$\begin{aligned} x &= q^{-2} + q^{-1} + 1 + q + 2q^2 + 2q^3 + 3q^4 + 3q^5 + 4q^6 + 5q^7 + \dots \\ y &= q^{-3} + 1 + 2q + 2q^2 + 4q^3 + 4q^4 + 7q^5 + 9q^6 + 12q^7 + \dots \end{aligned}$$

where $q = \exp(2\pi i\tau/119)$. Let τ be an imaginary quadratic number of discriminant -103. Then, further depending on τ, there are two options for the generalized class polynomial:[1]

$$\begin{aligned} H_{\tau_1}[E] &= X^3 + 2X^2 + XY + 2X + Y, \\ H_{\tau_2}[E] &= X^3 - 2X^2 - XY + X + 2Y + 1. \end{aligned}$$

One can compare this with Examples 3.2.1, 3.2.3. ☆

We may expect to estimate the size reduction of the coefficients of generalized class polynomials compared to Hilbert class polynomials through the following generalization of Definition 3.2.4.

Definition 3.2.8 The *reduction factor* of a modular curve C is

$$r(C) := \frac{\deg(j : X(N) \to \mathbf{P}^1)}{\deg(\psi : X(N) \to C)},$$

where ψ is any covering of C by the modular curve $X(N)$ for some $N \in \mathbf{Z}_{>0}$. △

For the case of rational elliptic curves with a finite number of points, this indeed correctly measures the expected asymptotic size reduction.

[1] For $X_0^+(119)$, the number of distinct class polynomials per discriminant is always at most two.

Theorem 3.2.9 (Houben, Streng) *Assume C is an elliptic curve over $\mathbf{Q}$ of rank 0, and that the map $\psi = (x, y) : \mathbf{H} \to \mathbf{C}$ consists of a pair of modular functions corresponding to Weierstrass coordinates of C and whose q-expansions are rational. If $\tau \in \mathbf{H}$ ranges over a sequence of imaginary quadratic points for which $K(\psi(\tau)) = K(j(\tau))$ and*

$$\frac{h(j(\tau))}{\log\log(\#\operatorname{Cl}(\mathcal{O}))} \to \infty, \tag{3.11}$$

then

$$\frac{\log|H_\tau[j]|_\infty}{\log|H_\tau[C]|_\infty} \to r(C). \tag{3.12}$$

Proof. See Theorem 6.3.4. □

For the case of the modular curve in Example 3.2.7, the reduction factor is 72; equal to the one for Weber's function. We did not find any elliptic curve with a reduction factor better than 72. Though this might seem somewhat disappointing, we believe there are several interesting conclusions and challenges for further work. For example:

(i) Weber's function is only known to yield class invariants for discriminants $\equiv 1 \pmod 8$. The generalized class polynomials associated to $X_+^0(119)$ are the first known to yield class invariants of reduction factor ≥ 72 for discriminants $\not\equiv 1 \pmod 8$.

(ii) The coordinate function x on $X_+^0(119)$ yields previously unknown univariate class polynomials. Its reduction factor of 36 already beats all previously known class invariants along a subset of imaginary quadratic discriminants of positive density (defined by a congruence condition). As a result, we expect that the further study of generalized class polynomials could provide new insights into the univariate case as well.

(iii) We know that there exist higher genus curves whose reduction factors exceed the Bröker-Stevenhagen bound. For example, the modular curve $X_+^0(239)$ has genus 3 and reduction factor $r(X_+^0(239)) = 120$. It remains to study whether the analogue of Theorem 3.2.9 holds for this curve.

It should be further noted that Theorem 3.2.9 only provides an asymptotic and concludes nothing about the speed of convergence. Some practical height reduction factors for $X_+^0(119)$, i.e. the left hand side of (3.12), for fundamental discriminants of prime class number are plotted in Figure 3.3.

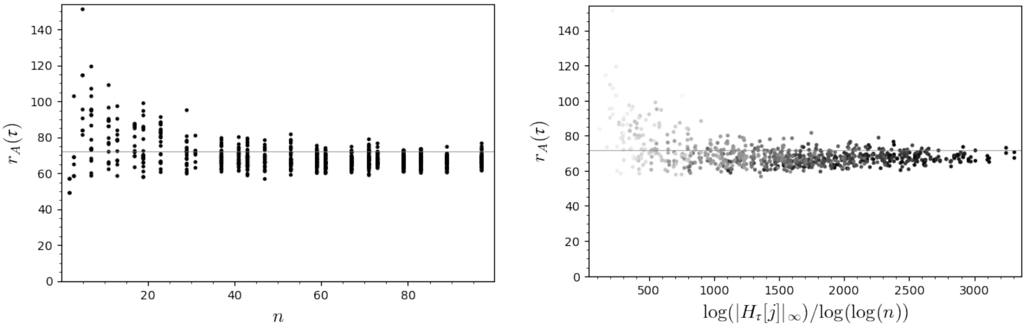


Figure 3.3: Practical reduction factors for $H_\tau[X_+^0(119)]$ for fundamental discriminants D with $\gcd(D, 119) = 1$ and prime class number $n < 100$. In the graph on the right, the x-axis plots the parameter from (3.11), and class polynomials with lower class number correspond to points with a lighter shade.

3.3 Radical isogenies

This section accompanies joint work with Wouter Castryck, Thomas Decru, and Frederik Vercauteren on *horizontal racewalking using radical isogenies*, corresponding to Chapter 7. After introducing and motivating the study of radical isogenies, we summarize and highlight the main contributions of our joint paper.

3.3.1 Computing isogeny chains

One of the main disadvantages of isogeny-based cryptography compared to other post-quantum proposals, such as lattice-based cryptography, is that it is relatively slow. Therefore, there has been continued interest in optimizing the algorithms underlying the evaluation of isogeny-based protocols. Most isogeny-based protocols rely on computing isogenies of small degree between elliptic curves. This often takes the form of *isogeny walks*, such as in Figure 2.6, which typically consist of chains of isogenies of small degree. For example, in CSIDH-512, see Example 2.2.3, one computes isogenies of degree up to 587 in chains of length up to 5. Since isogenies of smaller degree are typically easier to evaluate, a straightforward optimization to the CSIDH protocol is to skew the possible lengths of the chains; i.e. to take longer chains using isogenies of small degree and shorter chains using isogenies of larger degree. This happens, for example, in CSURF-512 [3]. The topic of our research is an attempt to make the evaluation of chains of isogenies of a *given* degree faster. For this, we would like to be able to extend isogeny chains efficiently, i.e., assuming we have computed a chain

$$E_0 \xrightarrow{\varphi_1} E_1 \xrightarrow{\varphi_2} \cdots \xrightarrow{\varphi_k} E_k \tag{3.13}$$

of isogenies of degree N of length $k \geq 1$ such that the composition of the isogenies is cyclic[2] of degree N^k, we would like to efficiently compute an isogeny $\varphi_k : E_k \to E_{k+1}$ of degree N that cyclically extends the chain. More precisely, we study the following problem.

Problem 3.3.1 *Let $E/\mathbf{F}_q$ be an elliptic curve and let $N \in \mathbf{Z}_{>1}$ coprime to* $\operatorname{char} \mathbf{F}_q$. *Let $P \in E(\mathbf{F}_q)[N]$ and consider the cyclic isogeny $\varphi : E \to E' = E/\langle P \rangle$ of degree N. Find P' on $E'(\overline{\mathbf{F}_q})$ such that the composition $E \xrightarrow{\varphi} E' \to E'/\langle P' \rangle$ is a cyclic isogeny of degree N^2.*

One possible method, if say $E'[N](\mathbf{F}_q) \cong (\mathbf{Z}/N\mathbf{Z})$, is to sample a random point $Q \in E'(\mathbf{F}_q)$, and to multiply by a suitable cofactor $P' := (\#E'(\mathbf{F}_q)/N)Q$. This results in a suitable point P' of order N with probability $\phi(N)/N$, where ϕ is Euler's totient function. However, this is non-deterministic, and relatively slow, since we must multiply a point on $E'(\mathbf{F}_q)$ by the (in case of cryptographic applications) large integer $\#E'(\mathbf{F}_q)/N \approx q$. Alternatively, one could

(i) find the j-invariant of $E'/\langle P' \rangle$ by extracting a root of the modular polynomial $\Phi_N(j(E'), X)$ different from $j(E)$ over $\mathbf{F}_q$; or

(ii) extract a root over $\mathbf{F}_q$ of the N-division polynomial on $E'(\mathbf{F}_q)$,

but these root-finding algorithms are typically even slower. A different approach is suggested by *radical isogenies*, first introduced by Castryck, Decru, and Vercauteren [5]. We will illustrate the idea with an example, which will make use of the following notion.

Definition 3.3.2 Let k be a field of characteristic $p \geq 0$ and let E/k be an elliptic curve. Let $P \in E(k)$ be a point of order $N \geq 4$ such that $p \nmid N$. Then there exist unique $b, c \in k$ such that E admits an isomorphism $\varphi : E \to E_{b,c}$ to the Weierstrass curve

$$E_{b,c} : y^2 + (1-c)xy - by = x^3 - bx^2 \tag{3.14}$$

for which $\varphi(P) = (0,0)$ [12, Lemma 2.1]. Such a Weierstrass model is called the *Tate normal form* of the pair (E, P). △

Example 3.3.3 (Radical 5-isogenies) Consider Problem 3.3.1 for the case $N = 5$. Using the Tate normal form, any elliptic curve with a point of order 5 can be written as

$$E : y^2 - (1-b)xy - by = x^3 - bx^2, \text{ where } P = (0,0).$$

for some value of the parameter b. Instead of specifying this parameter, we consider it as a formal variable and write down the general equation for $E/\langle P \rangle$ using Vélu's

[2] That is, the kernel of the isogeny is a cyclic subgroup of E_0. For example, if N is prime, this is equivalent to the condition that $\ker \varphi_i \neq \ker \hat{\varphi}_{i-1}$ for all $i = 2, \ldots, k$.

formulae (1.3):

$$y^2 + (1-b)xy - by = x^3 - bx^2 - 5b(b^2+2b-1)x - b(b^4+10b^3-5b^2+15b-1).$$

By finding an appropriate root of the 5-division polynomial on this curve, still written in terms of the formal variable b, we can obtain a formula for the coordinates of a 5-torsion point $P' = (x_0', y_0')$ on $E/\langle P\rangle$ that cyclically extends the isogeny $E \to E/\langle P\rangle$.

$$\begin{aligned} x_0' &= 5\alpha^4 + (b-3)\alpha^3 + (b+2)\alpha^2 + (2b-1)\alpha - 2b, \\ y_0' &= 5\alpha^4 + (b-3)\alpha^3 + (b^2-10b+1)\alpha^2 + (13b-b^2)\alpha - b^2 - 11b, \end{aligned}$$

where $\alpha = \sqrt[5]{b}$. Putting the curve-point pair $(E/\langle P\rangle, P')$ in Tate normal form, we obtain the following Weierstrass model for (the isomorphic curve) $E' \cong E/\langle P\rangle$:

$$E' : y^2 - (1-b')xy - b'y = x^3 - b'x^2, \text{ where } b' = \alpha\frac{\alpha^4+3\alpha^2+4\alpha^2+2\alpha+1}{\alpha^4-2\alpha^3+4\alpha^2-3\alpha+1},$$

thus obtaining an equation for the corresponding Tate-normal-form parameter b' of the curve E'. Now, computing a chain of 5-isogenies of elliptic curves over $\mathbf{F}_q$ amounts to iteratively computing b' from b. If $\gcd(5, q-1) = 1$ then this is deterministic, and, for fields $\mathbf{F}_q$ of cryptographic size, faster than any other known method of computing chains of 5-isogenies. ☆

In general, if (b, c) denote the Tate normal form parameters of a curve E together with a point of order $N > 3$, there exists a formal expression (depending on N) for the Tate normal form parameters (b', c') of a next curve E' in an N-isogeny chain. This expression is an algebraic function of b, c, and $\alpha = \sqrt[N]{\rho(b,c)}$, where $\rho(b,c)$ is another explicit algebraic function of b and c; cf. the formula for b' in Example 3.3.3. Such an expression is known as a *radical isogeny formula* and its general existence was shown in [5, Thm. 5].

3.3.2 Main contibutions

We now highlight results of joint work with Wouter Castryck, Thomas Decru, and Frederik Vercauteren. The full version of this work can be found in Chapter 7.

Though radical isogeny formulae always exist, they are not always easy to find. The approach suggested by Example 3.3.3 to use the N-division polynomial on E' was employed in [5], but proved computationally infeasible for $N > 13$. We developed an alternative way to compute radical isogeny formulae, which allowed to extend the range from $N \leq 13$ up to all primes $N \leq 41$. In addition to that, we rewrote and simplified the formulae up to degree $N = 19$. As an example, we compare expressions for the old and new radical 8-isogeny formulae below.

Example 3.3.4 (Old radical 8-isogeny formula, from [5])

$$\begin{aligned} A' \quad = \quad & \frac{-A^3+6A^2-12A+8}{A^2}\alpha^7 + \frac{4A^3-24A^2+48A-32}{A^3+4A^2-4A}\alpha^6 + \\ & \frac{-4A^3+24A^2-48A+32}{A^3+4A^2-4A}\alpha^5 + \frac{2A^3-12A^2+24A-16}{A^3+4A^2-4A}\alpha^4 + \\ & \frac{A-2}{A}\alpha^3 + \frac{-2A^2+4A}{A^2+4A-4}\alpha^2 + \frac{3A^2-4}{A^2+4A-4}\alpha + \frac{-A^2+2A}{A^2+4A-4}, \end{aligned}$$

where $\alpha = \sqrt[8]{(-A^3+A^2)/(A^4-8A^3+24A^2-32A+16)}$. ☆

This is equivalent to the following.

Example 3.3.5 (New radical 8-isogeny formula, from Chapter 7)

$$A' = \frac{-2A(A-2)\alpha^2 - A(A-2)}{(A-2)^2\alpha^4 - A(A-2)\alpha^2 - A(A-2)\alpha + A},$$

where $\alpha = \sqrt[8]{-A^2(A-1)/(A-2)^4}$. ☆

To evaluate radical isogenies, specifically to obtain α, one needs to compute an N-th root over $\mathbf{F}_p$. If p is odd, then for even degrees N such an N-th root is never unique. It turns out that choosing an incorrect root sometimes yields an N-isogeny that does not come from the class group action, i.e. an isogeny that is not *horizontal* in the sense of Lemma 1.2.2. We conjectured, and proved for $N \leq 14$, a simple criterion to select the right root, which allows for faster deterministic computation of isogeny walks in even degree. The combined optimizations and improvements to the radical isogeny formulae led to a speed up of 12% over the previous implementation of CSIDH-512 using radical isogenies (which in turn obtained a speed up of 19% over an implementation of 512-bit CSIDH without radical isogenies [1]). Using radical 16-isogenies, we obtained about a factor of 3 speed up for the computation of long chains of 2-isogenies over 512-bit prime fields.

3.4 Bibliography

[1] Daniel J. Bernstein, Luca De Feo, Antonin Leroux, and Benjamin Smith. Faster computation of isogenies of large prime degree. In *ANTS XIV—Proceedings of the Fourteenth Algorithmic Number Theory Symposium*, volume 4 of *Open Book Ser.*, pages 39–55. Math. Sci. Publ., Berkeley, CA, 2020.

[2] Reinier Bröker and Peter Stevenhagen. Constructing elliptic curves of prime order. In *Computational arithmetic geometry*, volume 463 of *Contemp. Math.*, pages 17–28. Amer. Math. Soc., Providence, RI, 2008.

[3] Wouter Castryck and Thomas Decru. CSIDH on the surface. In Jintai Ding and Jean-Pierre Tillich, editors, *PQCrypto 2020*, volume 12100 of *Lecture Notes in Computer Science*, pages 111–129. Springer, 2020.

[4] Wouter Castryck and Thomas Decru. An efficient key recovery attack on SIDH. In *Eurocrypt 2023 Pt. 5*, volume 14008 of *Lecture Notes in Computer Science*, pages 423–447. Springer, 2023.

[5] Wouter Castryck, Thomas Decru, and Frederik Vercauteren. Radical isogenies. In *Proceedings of Asiacrypt 2020 Part II*, volume 12492 of *Lecture Notes in Computer Science*, pages 493–519. Springer, 2020.

[6] Wouter Castryck, Jana Sotáková, and Frederik Vercauteren. Breaking the decisional Diffie-Hellman problem for class group actions using genus theory. In *Crypto 2020 Pt. 2*, volume 12171 of *Lecture Notes in Computer Science*, pages 92–120. Springer, 2020.

[7] Marc Hindry and Joseph H. Silverman. *Diophantine geometry*, volume 201 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2000. An introduction.

[8] Luciano Maino, Chloe Martindale, Lorenz Panny, Giacomo Pope, and Benjamin Wesolowski. A direct key recovery attack on SIDH. In *Eurocrypt 2023 Pt. 5*, volume 14008 of *Lecture Notes in Computer Science*, pages 448–471. Springer, 2023.

[9] Damien Robert. Breaking SIDH in polynomial time. In *Eurocrypt 2023 Pt. 5*, volume 14008 of *Lecture Notes in Computer Science*, pages 472–503. Springer, 2023.

[10] Atle Selberg. On the estimation of Fourier coefficients of modular forms. In *Proc. Sympos. Pure Math., Vol. VIII*, pages 1–15. Amer. Math. Soc., Providence, R.I., 1965.

[11] Joseph H. Silverman. *The arithmetic of elliptic curves*, volume 106 of *Graduate Texts in Mathematics*. Springer, second edition, 2009.

[12] Marco Streng. Generators of the group of modular units for $\Gamma^1(N)$ over the rationals. *Ann. H. Lebesgue*, 6:95–116, 2023.

[13] Andrew V. Sutherland. Accelerating the CM method. *LMS J. Comput. Math.*, 15:172–204, 2012.

[14] David Urbanik and David Jao. SoK: The problem landscape of SIDH. In *Proceedings of the 5th ACM on ASIA Public-Key Cryptography Workshop*, APKC '18, page 53–60, New York, NY, USA, 2018. Association for Computing Machinery.

[15] H. Weber. *Lehrbuch der Algebra*, volume III. Chelsea Publishing Company, 1902.

Chapter 4

On the decisional Diffie–Hellman problem for class group actions on oriented elliptic curves

This chapter consists of a paper written together with Wouter Castryck, Frederik Vercauteren, and Benjamin Wesolowski. It has been published as

Wouter Castryck, Marc Houben, Frederik Vercauteren, and Benjamin Wesolowski. On the decisional Diffie-Hellman problem for class group actions on oriented elliptic curves. *Res. Number Theory*, 8(4):Paper No. 99, 18, 2022. `https://doi.org/10.1007/s40993-022-00399-6`.

All authors of this paper contributed equally to the work.

Compared to the published version, we have fixed a few minor typographical and mathematical errors. We also improved the complexity estimate of the second step in Algorithm 1 based on a suggestion by Marco Streng. Additionally, we more concretely specified the input size in the complexity statements of Section 4.5, following a suggestion by Chloe Martindale. The numbering (of e.g. theorems and definitions) in the published version is different.

Acknowledgements

We thank the anonymous reviewers for several helpful comments, and Daniel J. Bernstein for suggesting to use Kedlaya–Umans factorization in the proof of Theorem 4.4.1.

Abstract

We show how the Weil pairing can be used to evaluate the assigned characters of an imaginary quadratic order $\mathcal{O}$ in an unknown ideal class $[\mathfrak{a}] \in \mathrm{Cl}(\mathcal{O})$ that connects two given $\mathcal{O}$-oriented elliptic curves (E, ι) and $(E', \iota') = [\mathfrak{a}](E, \iota)$. When specialized to ordinary elliptic curves over finite fields, our method is conceptually simpler and often somewhat faster than a recent approach due to Castryck, Sotáková and Vercauteren, who rely on the Tate pairing instead. The main implication of our work is that it breaks the decisional Diffie–Hellman problem for practically all oriented elliptic curves that are acted upon by an even-order class group. It can also be used to better handle the worst cases in Wesolowski's recent reduction from the vectorization problem for oriented elliptic curves to the endomorphism ring problem, leading to a method that always works in sub-exponential time.

4.1 Introduction

This paper is primarily concerned with the Decisional Diffie–Hellman problem (DDH) for ideal class groups acting on oriented elliptic curves through isogenies. In order to state this problem precisely, we fix an order $\mathcal{O}$ in an imaginary quadratic number field K along with an algebraically closed field k. A (primitive) $\mathcal{O}$-*orientation* on an elliptic curve E over k is an injective ring homomorphism $\iota : \mathcal{O} \hookrightarrow \mathrm{End}(E)$ that cannot be extended to a superorder $\mathcal{O}' \supsetneq \mathcal{O}$ in K. The set

$$\mathcal{E}\ell\ell_{\mathcal{O}}(k) = \{\, (E, \iota) \mid E \text{ an elliptic curve over } k \text{ and } \iota \text{ an } \mathcal{O}\text{-orientation on } E \,\}/\cong,$$

if non-empty, comes equipped with a free action

$$\mathrm{Cl}(\mathcal{O}) \times \mathcal{E}\ell\ell_{\mathcal{O}}(k) \longrightarrow \mathcal{E}\ell\ell_{\mathcal{O}}(k) : ([\mathfrak{a}], (E, \iota)) \longmapsto [\mathfrak{a}](E, \iota) \tag{4.1}$$

by the ideal class group of $\mathcal{O}$, see Section 4.2 for details (including what it means for two $\mathcal{O}$-oriented elliptic curves (E, ι) and (E', ι') to be isomorphic). Now assume that a party, say Eve, has unlimited access to samples from $\mathcal{E}\ell\ell_{\mathcal{O}}(k)^3$ that are consistently of either of the following two forms:

$$\begin{array}{ll} \big(\, [\mathfrak{a}](E, \iota),\ [\mathfrak{b}](E, \iota),\ [\mathfrak{a}][\mathfrak{b}](E, \iota) \,\big) & \quad [\mathfrak{a}], [\mathfrak{b}] \xleftarrow{\$} \mathrm{Cl}(\mathcal{O}), \\ \big(\, [\mathfrak{a}](E, \iota),\ [\mathfrak{b}](E, \iota),\ [\mathfrak{c}](E, \iota) \,\big) & \quad [\mathfrak{a}], [\mathfrak{b}], [\mathfrak{c}] \xleftarrow{\$} \mathrm{Cl}(\mathcal{O}), \end{array}$$

for some fixed and publicly known (E, ι). Then Eve successfully solves DDH if she can guess, with non-negligible advantage, from which of these two distributions her triples were sampled.

The hardness of the decisional Diffie–Hellman problem is a natural security foundation for cryptographic constructions based on ideal class group actions, which can be traced back to the works of Couveignes [11] and Rostovtsev–Stolbunov [24, 28] and which have attracted much attention lately, in the context of post-quantum cryptography. Here, one lets k be an algebraic closure of a finite field, in which case all curves in $\mathcal{E}\ell\ell_{\mathcal{O}}(k)$ can be defined over a common finite subfield $F \subseteq k$. While the initial focus was on ordinary elliptic curves, whose orientations ι are just ring isomorphisms, most of the latest work is concerned with supersingular elliptic curves, whose endomorphism rings are orders in a quaternion algebra and therefore leave room for a wide range of orientations. Here, we highlight supersingular elliptic curves defined over a finite prime field $\mathbf{F}_p$, which are naturally oriented by an order in $\mathbf{Q}(\sqrt{-p})$. The corresponding ideal class group actions underpin CSIDH [6] and spin-offs such as [1, 15, 2, 20], and tend to yield more practical cryptosystems than in the ordinary case. More generally oriented supersingular elliptic curves made their first cryptographic appearance in the OSIDH protocol due to Colò and Kohel [10]. To date, this protocol remains largely theoretical, but it has attracted a good amount of recent interest, see e.g. [13, 22, 31].

Our paper revisits the recent work [8], which presents an efficient solution to DDH for essentially all ordinary elliptic curves over finite fields whose endomorphism ring has an even class number. In more detail, as soon as there exists a non-trivial assigned character $\chi : \mathrm{Cl}(\mathcal{O}) \to \{\pm 1\}$ of sufficiently small modulus m, the attack from [8] allows Eve to compute $\chi([\mathfrak{a}])$ merely from the knowledge of (E, ι) and $(E', \iota') = [\mathfrak{a}](E, \iota)$, i.e., without knowing $[\mathfrak{a}]$ itself. This indeed suffices to break DDH, since it allows her to check whether $\chi([\mathfrak{c}]) = \chi([\mathfrak{a}])\chi([\mathfrak{b}])$, which is true for $[\mathfrak{c}] = [\mathfrak{a}][\mathfrak{b}]$, but for uniformly random $[\mathfrak{c}]$ it fails with probability $1/2$.

Unfortunately, the method from [8] is specific to ordinary curves: the attack proceeds by extending the base field and navigating to the floors of the m-isogeny volcanoes[1] of (E, ι) and (E, ι'), with the goal of enforcing non-trivial cyclic rational m^∞-torsion, and then recovering the character value using two Tate pairing computations. Beyond ordinary curves, it is generally impossible to turn the rational m^∞-torsion cyclic using an isogeny walk, so this strategy fails. For supersingular elliptic curves over $\mathbf{F}_p$ with $p \equiv 1 \bmod 4$ equipped with their natural $\mathbf{Z}[\sqrt{-p}]$-orientation, where it suffices to consider the assigned character of modulus $m = 4$, an ad-hoc fix was given in [8, Thm. 10], but it is unclear how this fix would generalize.

Contribution

We give an alternative method for computing assigned character values $\chi([\mathfrak{a}])$ purely from (E, ι) and $(E', \iota') = [\mathfrak{a}](E, \iota)$, using the Weil pairing rather than the Tate pairing. Our approach deals with arbitrary orientations and works over arbitrary fields. Moreover, it simplifies and often speeds up the attack from [8] in the case of ordinary elliptic curves over finite fields, as it avoids the need for navigating through isogeny volcanoes. It also naturally incorporates the previously ad-hoc case of supersingular elliptic curves over prime fields.

The main result is easy enough to be stated right away; we recall that for an odd prime divisor $m \mid \mathrm{Disc}(\mathcal{O})$, the assigned character of modulus m is defined as

$$\chi_m : \mathrm{Cl}(\mathcal{O}) \to \{\pm 1\} : [\mathfrak{a}] \mapsto \left(\frac{N(\mathfrak{a})}{m} \right) \tag{4.2}$$

where it is assumed that $[\mathfrak{a}]$ is represented by an ideal $\mathfrak{a}$ of norm coprime to m (see our conventions further down) and $\left(\frac{\cdot}{m}\right)$ is the Legendre symbol.

Theorem 4.1.1 *Let $\mathcal{O}$ be an imaginary quadratic order and let $(E, \iota), (E', \iota')$ be $\mathcal{O}$-oriented elliptic curves connected by an ideal class $[\mathfrak{a}] \in \mathrm{Cl}(\mathcal{O})$. Let $m \mid \mathrm{Disc}(\mathcal{O})$ be an odd prime divisor different from $\mathrm{char}\, k$ and consider the assigned character $\chi_m : \mathrm{Cl}(\mathcal{O}) \to \{\pm 1\}$ of modulus m. Then $\mathcal{O}$ admits a generator σ (i.e. $\mathcal{O} = \mathbf{Z}[\sigma]$) of norm coprime to m, and for any such σ there exist points $P \in E[m]$, $P' \in E'[m]$ such that $\iota(\sigma)(P)$ is not a multiple of P, and likewise for P'. Moreover*

$$\chi_m([\mathfrak{a}]) = \left(\frac{a}{m} \right)$$

with $a = \log_{e_m(P, \iota(\sigma)(P))} e_m(P', \iota'(\sigma)(P'))$, regardless of the choice of such σ, P, P'.

[1] Or rather 2-isogeny volcanoes in case $m \in \{4, 8\}$.

The condition that σ be a generator of $\mathcal{O}$ can be relaxed to $\sigma \in \mathcal{O} \setminus (\mathbf{Z} + m\mathcal{O})$. A proof of Theorem 4.1.1, along with its adaptations covering assigned characters with even modulus, can be found in Section 4.3. Since these results apply to arbitrary fields, they may be of independent theoretical interest.

Applications and implications

From a cryptographic viewpoint, the most important consequence is that DDH should be considered broken by classical computers for essentially all elliptic curves over finite fields that are oriented by an imaginary quadratic order $\mathcal{O}$ with even class number; see Section 4.4 for a more in-depth discussion.

As a more surprising application, we prove in Section 4.5 that the new method allows to significantly improve reductions between computational problems underlying isogeny-based cryptography. On one hand, we have the problem of computing endomorphism rings of supersingular elliptic curves. It is of foundational importance to the field, as its presumed hardness is necessary for the security of essentially all isogeny-based cryptosystems [17, 7, 16]. Oriented versions of this Endomorphism Ring Problem were introduced in [31]. On the other hand, many cryptosystems relate directly to the presumably hard inversion problem for the action of the class group $\mathrm{Cl}(\mathcal{O})$ on oriented supersingular curves: the Vectorization Problem. It was proved in [31] that the vectorization problem reduces to the endomorphism ring problem in polynomial time in the length of the instance and in $\#(\mathrm{Cl}(\mathcal{O})[2])$. Unfortunately, the dependence on $\#(\mathrm{Cl}(\mathcal{O})[2])$ means that the reduction is, in the worst case, exponential in the size of the input, since $\#(\mathrm{Cl}(\mathcal{O})[2])$ could be as large as $D^{1/\log\log D}$, where $D = |\operatorname{Disc}(\mathcal{O})|$. We improve this result, by proving in Section 4.5 that there is a reduction from the vectorization problem to the endomorphism ring problem that, in the worst case, is sub-exponential in the length of the input.

Conventions

Throughout, all ideal classes $[\mathfrak{a}] \in \mathrm{Cl}(\mathcal{O})$ are assumed to be represented by an ideal $\mathfrak{a}$ of norm coprime to $p \operatorname{Disc}(\mathcal{O})$, where $p = \max\{1, \operatorname{char} k\}$. Such a representative always exists, see e.g. [12, Cor. 7.17]. For an $\mathcal{O}$-oriented elliptic curve (E, ι) and a point $P \in E$, we will sometimes write $\sigma(P)$ instead of $\iota(\sigma)(P)$ if ι is clear from the context. Likewise, for $[\mathfrak{a}] \in \mathrm{Cl}(\mathcal{O})$ we will sometimes write $[\mathfrak{a}]E$ for the first component of $[\mathfrak{a}](E, \iota)$.

Paper organization

Section 4.2 provides background: it gives the full list of assigned characters of an imaginary quadratic order and it recalls how its ideal class group acts on oriented elliptic curves. Our main Section 4.3 contains a proof of Theorem 4.1.1, as well as statements and proofs for the even-modulus counterparts. Section 4.4 discusses the algorithmic aspects of these results, along with their implications for the decisional Diffie–Hellman problem. Finally, in Section 4.5 we present our improved reduction from the vectorization problem for oriented elliptic curves to the endomorphism ring problem.

4.2 Background

4.2.1 Assigned characters

The following is a very brief summary of the relevant parts of [12, I.§3 & II.§7], to which we refer for more details. From genus theory, we know that each order $\mathcal{O}$ in an imaginary quadratic field comes equipped with an explicit list of group homomorphisms $\mathrm{Cl}(\mathcal{O}) \to \{\pm 1\}$, called the *assigned characters*, whose joint kernel is $\mathrm{Cl}(\mathcal{O})^2$. Writing

$$\mathrm{Disc}(\mathcal{O}) = -2^f d = -2^f m_1^{f_1} m_2^{f_2} \cdots m_r^{f_r}$$

for distinct odd prime numbers $m_1, \ldots, m_r$ and exponents $f \geq 0$, $f_1, \ldots, f_r \geq 1$, this list consists of

$$\begin{array}{ll}
\chi_{m_1}, \ldots, \chi_{m_r} & \text{if } f = 0, \\
\chi_{m_1}, \ldots, \chi_{m_r}, \delta & \text{if } f = 2 \text{ and } d \equiv 1 \bmod 4, \\
\chi_{m_1}, \ldots, \chi_{m_r} & \text{if } f = 2 \text{ and } d \equiv 3 \bmod 4, \\
\chi_{m_1}, \ldots, \chi_{m_r}, \delta\epsilon & \text{if } f = 3 \text{ and } d \equiv 1 \bmod 4, \\
\chi_{m_1}, \ldots, \chi_{m_r}, \epsilon & \text{if } f = 3 \text{ and } d \equiv 3 \bmod 4, \\
\chi_{m_1}, \ldots, \chi_{m_r}, \delta & \text{if } f = 4, \\
\chi_{m_1}, \ldots, \chi_{m_r}, \delta, \epsilon & \text{if } f \geq 5.
\end{array}$$

Here χ_{m_i} is defined as in (4.2) and

$$\delta : \mathrm{Cl}(\mathcal{O}) \to \{\pm 1\} : [\mathfrak{a}] \mapsto (-1)^{\frac{N(\mathfrak{a})-1}{2}}, \quad \epsilon : \mathrm{Cl}(\mathcal{O}) \to \{\pm 1\} : [\mathfrak{a}] \mapsto (-1)^{\frac{N(\mathfrak{a})^2-1}{8}}.$$

Observe that $\delta\epsilon$ can be described in one go as

$$\delta\epsilon : \mathrm{Cl}(\mathcal{O}) \to \{\pm 1\} : [\mathfrak{a}] \mapsto (-1)^{\frac{(N(\mathfrak{a})+2)^2-9}{8}}.$$

We write $\mu \in \{r, r+1, r+2\}$ for the total number of assigned characters.[2]

Because the joint kernel is $\mathrm{Cl}(\mathcal{O})^2$, any character of $\mathrm{Cl}(\mathcal{O})$ whose order divides 2 can be written as a product of pairwise distinct assigned characters. As it turns out, there is a unique non-trivial combination that produces the trivial character:

$$\chi_{m_1}^{f_1 \bmod 2} \chi_{m_2}^{f_2 \bmod 2} \cdots \chi_{m_r}^{f_r \bmod 2} \delta^{\frac{d+1}{2} \bmod 2} \epsilon^{f \bmod 2} = 1. \tag{4.3}$$

Therefore, by combining assigned characters we obtain $2^{\mu-1}$ distinct characters. Necessarily, this quantity equals the cardinality of $\mathrm{Cl}(\mathcal{O}) / \mathrm{Cl}(\mathcal{O})^2 \cong \mathrm{Cl}(\mathcal{O})[2]$.

[2] Note that two different assigned characters may define the same map $\mathrm{Cl}(\mathcal{O}) \to \{\pm 1\}$. Thus, formally, the definition of an assigned character should include its symbol (e.g. χ_{m_1}) as appearing in the list above.

Example 4.2.1 For a prime number $p \equiv 1 \bmod 4$, the ring $\mathbf{Z}[\sqrt{-p}]$ has two assigned characters: δ and χ_p. By (4.3) these are in fact equal to each other, and non-trivial. If $p \equiv 3 \bmod 4$ then $\mathbf{Z}[\sqrt{-p}]$ has only one assigned character, namely χ_p, and it is trivial. ☆

We often make reference to the *modulus* m of an assigned character χ, which is an important complexity parameter for our attack. This is simply defined to be

$$\begin{cases} m_i & \text{if } \chi = \chi_{m_i}, \\ 4 & \text{if } \chi = \delta, \\ 8 & \text{if } \chi = \epsilon, \delta\epsilon. \end{cases}$$

Note that $\chi([\mathfrak{a}]) = \chi([\mathfrak{a}'])$ as soon as $N(\mathfrak{a}) \equiv N(\mathfrak{a}') \bmod m$. Typically m is the smallest positive integer with this property, but not always (e.g., as in the case of $m_i = p$ in both examples above).

4.2.2 Class group action

We now recall how the ideal class group of $\mathcal{O}$ acts on $\mathcal{E}\!\ell\!\ell_{\mathcal{O}}(k)$. This is part of the theory of complex multiplication, which is classical for $k = \mathbf{C}$, while for k an algebraic closure of a finite field this was elaborated in [30, §3.9-12]; see also [22] for the specifics of the supersingular case. For arbitrary k, we refer to Milne's course notes [21, §7].

If ι is an $\mathcal{O}$-orientation on an elliptic curve E over k, then we can linearly extend it to a map $K \hookrightarrow \mathrm{End}^0(E)$, where $\mathrm{End}^0(E) = \mathrm{End}(E) \otimes_{\mathbf{Z}} \mathbf{Q}$ denotes the endomorphism algebra. To each isogeny $\varphi : E \to E'$ we can naturally attach an embedding

$$\iota_{\mathbf{Q}} : K \hookrightarrow \mathrm{End}^0(E') : \sigma \mapsto \frac{1}{\deg \varphi} \varphi \circ \iota(\sigma) \circ \hat{\varphi},$$

whose restriction to the preimage $\mathcal{O}'$ of $\mathrm{End}(E')$ is an orientation that is called the *induced orientation*, denoted by $\varphi_* \iota$. We are primarily interested in isogenies φ for which $\mathcal{O}' = \mathcal{O}$, in which case φ is said to be *horizontal* with respect to ι. Two $\mathcal{O}$-oriented elliptic curves $(E, \iota), (E', \iota')$ are called *isomorphic*, denoted $(E, \iota) \cong (E', \iota')$, if there exists an isomorphism $\varphi : E \to E'$ such that $\iota' = \varphi_* \iota$.

The default way to construct a horizontal isogeny is by considering an invertible ideal $\mathfrak{a} \subseteq \mathcal{O}$ of norm coprime to $\max\{1, \operatorname{char} k\}$ and attaching to it the finite subgroup

$$E[\mathfrak{a}] = \bigcap_{\alpha \in \mathfrak{a}} \ker \iota(\alpha).$$

Then the separable degree-$N(\mathfrak{a})$ isogeny $\varphi_{\mathfrak{a}} : E \to E'$ with kernel $E[\mathfrak{a}]$ is horizontal. In particular E' comes naturally equipped with an $\mathcal{O}$-orientation $\iota' = \varphi_{\mathfrak{a}*} \iota$. The pair (E', ι') is well-defined up to isomorphism and only depends on the class of $\mathfrak{a}$ inside $\mathrm{Cl}(\mathcal{O})$; we write $[\mathfrak{a}](E, \iota) := (E', \iota')$. This defines the map (4.1), which turns out to be a free group action.

Remark 4.2.2 In general the action is not transitive, where one subtlety is reflected in [22, Prop. 3.3]; see also the example in [22, §3.1] and the proof of [26, Thm. 4.5]. This has no consequences for the current paper, since we are working in a single orbit, namely that of the starting curve (E, ι). $\diamondsuit$

4.3 Evaluating characters using the Weil pairing

In this section we prove Theorem 4.1.1 and discuss its analogues for the assigned characters $\delta, \epsilon, \delta\epsilon$. In all cases it is assumed that $p = \max\{1, \operatorname{char} k\}$ is coprime to the modulus of the character under consideration. If p is an odd prime then χ_p, if it appears in the list of assigned characters, can be computed from the other characters using the relation (4.3); see for instance Example 4.2.1 where we had $\chi_p = \delta$. If $p = 2$ then the same conclusion holds for δ, ϵ or $\delta\epsilon$, because in even characteristic at most one of these three characters can appear in the list of assigned characters.[3]

4.3.1 Preliminaries

Lemma 4.3.1 *Let $\mathcal{O}$ be an imaginary quadratic order and let m be an odd prime number. Then $\mathcal{O} = \mathbf{Z}[\sigma]$ for some $\sigma \in \mathcal{O}$ of norm coprime to m.*

Proof. Let $\tau \in \mathcal{O}$ be a generator of $\mathcal{O}$, suppose of norm divisible by m. Then for any $k \in \mathbf{Z}$,

$$N(\tau + k) = N(\tau) + k(\operatorname{tr}(\tau) + k) \equiv k(\operatorname{tr}(\tau) + k) \bmod m.$$

Since $m \geq 3$ we can thus always find $k \in \mathbf{Z}$ such that $m \nmid N(\tau + k)$. $\square$

Lemma 4.3.2 *Let $\mathcal{O}$ be an imaginary quadratic order of even discriminant. Then $\mathcal{O} = \mathbf{Z}[\sigma]$ for some $\sigma \in \mathcal{O}$ of odd norm.*

Proof. Let $\tau \in \mathcal{O}$ be a purely imaginary generator of $\mathcal{O}$, e.g. $\tau = \sqrt{\operatorname{Disc}(\mathcal{O})/4}$, where $\operatorname{Disc}(\mathcal{O})$ is the discriminant of $\mathcal{O}$. Then $N(\tau + 1) = N(\tau) + \operatorname{tr}(\tau) + 1 = N(\tau) + 1$, hence we can take $\sigma = \tau$ or $\sigma = \tau + 1$. $\square$

Lemma 4.3.3 *Let $\mathcal{O}$ be an imaginary quadratic order, let (E, ι) be an $\mathcal{O}$-oriented elliptic curve over k, let $m \neq \operatorname{char} k$ be a prime number, and let $\sigma \in \mathcal{O}$ be a generator. Then there exists a $P \in E[m]$ such that $\iota(\sigma)(P)$ is not a multiple of P.*

Proof. The endomorphism $\iota(\sigma)$ of E induces an $\mathbf{F}_m$-linear map $E[m] \to E[m]$. Suppose to the contrary that every $P \in E[m]$ is an eigenvector. This can only happen if the map has the full m-torsion $E[m]$ as an eigenspace. Thus there exists $\lambda \in \mathbf{Z}$ such

[3] If (E, ι) is an $\mathcal{O}$-oriented elliptic curve over an algebraically closed field k with $\operatorname{char} k = 2$, then $2^5 \nmid \operatorname{Disc}(\mathcal{O})$. Indeed, if we would have $2^5 \mid \operatorname{Disc}(\mathcal{O})$ then E is necessarily supersingular, hence it concerns $y^2 + y = x^3$, the unique supersingular elliptic curve in characteristic 2. Its endomorphism ring is isomorphic to the ring of Hurwitz quaternions H, and it is easy to check that every embedding $\mathcal{O} \hookrightarrow H$ can be extended to an embedding $\mathcal{O}' \hookrightarrow H$ with $\operatorname{Disc}(\mathcal{O}') = \operatorname{Disc}(\mathcal{O})/4$. See [22, Prop. 3.2] for a generalization of this observation.

that $E[m] \subseteq \ker(\iota(\sigma - \lambda))$. It then follows that $\iota_{\mathbf{Q}}((\sigma-\lambda)/m) \in \mathrm{End}(E)$, and hence that $\sigma - \lambda \in m\mathcal{O}$ by the fact that ι is a primitive embedding, i.e. it cannot be extended to a strict superorder of $\mathcal{O}$. Since $\mathbf{Z} + m\mathcal{O} \subsetneq \mathcal{O}$ this contradicts the assumption that σ generates $\mathcal{O}$. □

4.3.2 Evaluating the characters χ_m

We now prove Theorem 4.1.1.

Proof of Theorem 4.1.1. The existence of σ, P, P' follows from Lemma 4.3.1 combined with Lemma 4.3.3. The endomorphism $\iota(\sigma)$ of E induces an $\mathbf{F}_m$-linear map $E[m] \to E[m]$. Since $m \mid \mathrm{Disc}(\mathcal{O}) = \mathrm{tr}(\sigma)^2 - 4N(\sigma)$ and $m \nmid N(\sigma)$, its characteristic polynomial has a nonzero double root, say $\alpha \in \mathbf{F}_m^\times$. Consequently, we can extend to a basis P_0, P of $E[m]$ for which the matrix of σ is in upper-triangular form $\left(\begin{smallmatrix} \alpha & \beta \\ 0 & \alpha \end{smallmatrix}\right)$ for some $\beta \in \mathbf{F}_m^\times$. With respect to this basis any $Q \in E[m]$ that is not an eigenvector of σ is of the form $Q = \lambda P_0 + \mu P$ where $\mu \neq 0$. We see that

$$e_m(Q, \sigma(Q)) = e_m(\lambda P_0 + \mu P, (\alpha\lambda + \beta\mu)P_0 + \alpha\mu P) = e_m(P, \beta P_0)^{\mu^2} = e_m(P, \sigma(P))^{\mu^2},$$

showing that $e_m(P, \sigma(P))$ is independent of the choice of P, up to raising to powers that are nonzero squares modulo m. Then, of course, the same conclusion applies to $e_m(P', \sigma(P'))$.

Recall our convention from the introduction, namely that we assume that the norm of $\mathfrak{a}$, which equals the degree of the corresponding isogeny $\varphi = \varphi_{\mathfrak{a}} : E \to E'$, is coprime to m. In particular, $P_0 \notin \ker \varphi$. By definition of the class group action, $\iota' = \varphi_* \iota$ satisfies

$$\iota'(\sigma)(\varphi(P)) = \left(\frac{1}{\deg \varphi} \varphi \iota(\sigma) \hat{\varphi}\right)(\varphi(P)) = \varphi(\iota(\sigma)(P)) = \beta\varphi(P_0) + \alpha\varphi(P),$$

showing that $\varphi(P)$ is not an eigenvector for $\iota'(\sigma)$ acting on $([\mathfrak{a}]E)[m]$. So we see that $e_m(\varphi(P), \iota'(\sigma)(\varphi(P)))$ is obtained from $e_m(P', \iota'(\sigma)(P'))$ by raising it to a nonzero square mod m. To conclude, we observe that

$$e_m(\varphi(P), \iota'(\sigma)(\varphi(P))) = e_m(\varphi(P), \varphi(\iota(\sigma)(P))) = e_m(P, \iota(\sigma)(P))^{\deg \varphi}. \qquad \square$$

4.3.3 Evaluating δ, ϵ or $\delta\epsilon$

We now present the analogues of Theorem 4.1.1 for the even-modulus characters δ, ϵ and $\delta\epsilon$. We first focus on δ, which, as we saw in Section 4.2.1, is an assigned character if and only if we can write $\mathrm{Disc}(\mathcal{O}) = -4 \cdot d$ for some $d \equiv 0, 1 \bmod 4$.

Proposition 4.3.4 *Assume* $\mathrm{char}\, k \neq 2$. *Let $\mathcal{O}$ be an imaginary quadratic order of discriminant $-4 \cdot d$ where $d \equiv 0, 1 \bmod 4$, and let (E, ι), (E', ι') be $\mathcal{O}$-oriented elliptic curves over k connected by an ideal class $[\mathfrak{a}] \in \mathrm{Cl}(\mathcal{O})$. Then $\mathcal{O}$ admits an odd-norm* *generator σ, and for any such σ there exist points $P \in E[4]$, $P' \in E'[4]$ such that $\iota(\sigma)(2P) \neq 2P$ and $\iota'(\sigma)(2P') \neq 2P'$. Moreover*

$$\delta([\mathfrak{a}]) = (-1)^{\frac{a-1}{2}},$$

with $a = \log_{e_4(P,\iota(\sigma)(P))} e_4(P', \iota'(\sigma)(P'))$, for any such choice of σ, P, P'.

Proof. The existence of σ, P, P' follows from Lemma 4.3.2 and Lemma 4.3.3. Note that the assumption on the discriminant of $\mathcal{O}$ shows that the character δ indeed exists, and that this implies that $N(\sigma) \equiv 1 \bmod 4$ (since the principal ideal class $[(\sigma)]$ lies in the kernel of δ). By upper-triangularizing the action of σ on $E[2]$ as in the proof of Theorem 4.1.1, we see that there exists a $P_0 \in E[4]$ such that the matrix M_σ of σ acting on $E[4]$ with respect to the basis P_0, P is of the form

$$M_\sigma \equiv \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \bmod 2.$$

Since $N(\sigma) \equiv 1 \bmod 4$ this means that M_σ is of the form either $\left(\begin{smallmatrix} \alpha & \beta \\ 0 & \alpha \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} \alpha & \beta \\ 2 & -\alpha \end{smallmatrix}\right)$, with α, β odd. Any Q with the property that $\sigma(2Q) \neq 2Q$ is of the form $\lambda P_0 + \mu P$ where μ is odd. If M_σ is of the first form we get

$$e_4(Q, \sigma(Q)) = e_4(\lambda P_0 + \mu P, (\alpha\lambda + \beta\mu)P_0 + \alpha\mu P) = e_4(P, \beta P_0)^{\mu^2} = e_4(P, \sigma(P))^{\mu^2}.$$

If M_σ is of the second form we again get

$$\begin{aligned} e_4(Q, \sigma(Q)) &= e_4(\lambda P_0 + \mu P, (\alpha\lambda + \beta\mu)P_0 + (2\lambda - \alpha\mu)P) \\ &= e_4(P, \beta P_0)^{\mu^2} e_4(P, P_0)^{2(\lambda\alpha\mu - \lambda^2)} = e_4(P, \sigma(P))^{\mu^2} \end{aligned}$$

where the last equality uses that λ, μ, α are odd. From $\mu^2 \equiv 1 \bmod 4$ it follows that $e_4(P, \sigma(P))$ does not depend on the choice of P. Then, of course, the same is true for $e_4(P', \sigma(P'))$.

By our convention we assume that the norm of $\mathfrak{a}$, and hence the degree of the corresponding isogeny $\varphi = \varphi_{\mathfrak{a}} : E \to E'$, is odd. In particular, $2P_0 \notin \ker \varphi$ and

$$\iota'(\sigma)(\varphi(2P)) = \left(\frac{1}{\deg \varphi} \varphi \iota(\sigma) \hat{\varphi}\right)(\varphi(2P)) = \varphi(\iota(\sigma)(2P)) = \varphi(2P_0) + \varphi(2P)$$

is different from $\varphi(2P)$. Thus we find that $e_4(P', \sigma(P'))$ equals

$$e_4(\varphi(P), \iota'(\sigma)(\varphi(P))) = e_4(\varphi(P), \varphi(\iota(\sigma)(P))) = e_4(P, \iota(\sigma)(P))^{\deg \varphi},$$

which concludes the proof. □

Next, we discuss the modulus-8 characters ϵ and $\delta\epsilon$. Note that by Section 4.2.1, we have that ϵ is an assigned character if and only if either $2^5 \mid \mathrm{Disc}(\mathcal{O})$ or $\mathrm{Disc}(\mathcal{O}) = -2^3 \cdot d$ with $d \equiv 3 \bmod 4$. Similarly, $\delta\epsilon$ is an assigned character if and only if either $2^5 \mid \mathrm{Disc}(\mathcal{O})$ or $\mathrm{Disc}(\mathcal{O}) = -2^3 \cdot d$ with $d \equiv 1 \bmod 4$.

Proposition 4.3.5 *Assume* $\operatorname{char} k \neq 2$*, let* $\mathcal{O}$ *be an imaginary quadratic order of discriminant* $\operatorname{Disc}(\mathcal{O}) \equiv -2^f d$ *with* d *odd and* $f \geq 3$*, and consider* $\mathcal{O}$*-oriented elliptic curves* (E, ι)*,* (E', ι') *over* k *connected by an ideal class* $[\mathfrak{a}] \in \operatorname{Cl}(\mathcal{O})$*. Assume that* ϵ*, resp.* $\delta\epsilon$*, appears among the assigned characters of* $\mathcal{O}$*. Then* $\mathcal{O}$ *admits an odd-norm generator* σ*, and for any such* σ *there exist points* $P \in E[8]$*,* $P' \in E'[8]$ *such that* $\iota(\sigma)(4P) \neq 4P$ *and* $\iota'(\sigma)(4P') \neq 4P'$*. Moreover* $\epsilon([\mathfrak{a}])$*, resp.* $\delta\epsilon([\mathfrak{a}])$*, can be computed as*

$$\epsilon([\mathfrak{a}]) = (-1)^{\frac{a^2-1}{8}}, \quad \textit{resp.} \quad \delta\epsilon([\mathfrak{a}]) = (-1)^{\frac{(a+2)^2-9}{8}},$$

with

$$a = \log_{e_8(P, \iota(\sigma)(P))} e_8(P', \iota'(\sigma)(P')),$$

and for any such choice of σ, P, P'*.*

Proof. As in the previous proof, the existence of σ, P, P' follows from Lemma 4.3.2 and Lemma 4.3.3. The main difference with the foregoing proofs is that if $Q \in E[8]$ is another point satisfying $\sigma(4Q) \neq 4Q$, then $e_8(Q, \sigma(Q))$ relates more subtly to $e_8(P, \sigma(P))$. Namely, we will argue that

$$e_8(Q, \sigma(Q)) \in \left\{ e_8(P, \sigma(P)), e_8(P, \sigma(P))^{N(\sigma)} \right\}, \tag{4.4}$$

and then of course the same again applies to $e_8(P', \sigma(P'))$. This will then lead to the conclusion that

$$e_8(P', \sigma(P')) \in \left\{ e_8(P, \sigma(P))^{\deg \varphi}, e_8(P, \sigma(P))^{N(\sigma) \deg \varphi} \right\},$$

which is indeed sufficient, since the principal ideal class $[(\sigma)]$ has trivial character values. More explicitly, if ϵ exists then we must have $N(\sigma) \bmod 8 \in \{1, 7\}$, while if $\delta\epsilon$ exists then we have $N(\sigma) \bmod 8 \in \{1, 3\}$.

In order to prove (4.4), note that, since $N(\sigma) \equiv 1 \bmod 2$,

$$\operatorname{tr}(\sigma)^2 + 4 \equiv \operatorname{tr}(\sigma)^2 - 4 \cdot N(\sigma) = \operatorname{Disc}(\mathcal{O}) \equiv 0 \bmod 8,$$

so that $\operatorname{tr}(\sigma) \equiv 2 \bmod 4$. It follows that the characteristic polynomial of σ modulo 4 is $X^2 + 2X + N(\sigma)$, hence we can extend to a basis P_0, P of $E[8]$ such that the matrix of $\iota(\sigma)$ acting on $E[8]$ is of the form

$$M_\sigma \equiv \begin{cases} \begin{pmatrix} \alpha & \beta \\ 0 & \alpha \end{pmatrix} \bmod 4 & \text{if } N(\sigma) \equiv 1 \bmod 4, \\ \begin{pmatrix} \alpha & \beta \\ 2 & \alpha \end{pmatrix} \bmod 4 & \text{if } N(\sigma) \equiv 3 \bmod 4, \end{cases}$$

with α, β odd. It follows that

$$M_\sigma^2 \equiv \begin{cases} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \bmod 4 & \text{if } N(\sigma) \equiv 1 \bmod 4, \\ \begin{pmatrix} 3 & 2 \\ 0 & 3 \end{pmatrix} \bmod 4 & \text{if } N(\sigma) \equiv 3 \bmod 4. \end{cases}$$

In any case we can record that

$$e_8(P, \sigma^2(P))^2 = e_8(P, P_0)^4 = -1. \tag{4.5}$$

Now, with respect to the basis $P, \sigma(P)$, the matrix of $\iota(\sigma)$ acting on $E[8]$ is congruent to $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ mod 2. Any other $Q = \lambda P + \mu\sigma(P)$ such that $\sigma(4Q) \neq 4Q$ thus has exactly one of λ, μ odd. We now proceed to showing (4.4). If μ is odd then we can write $\sigma(Q) = \lambda' P + \mu' \sigma(P)$ with λ' odd, so since

$$e_8(Q, \sigma(Q))^{N(\sigma)} = e_8(\sigma(Q), \sigma^2(Q))$$

we may reduce to the case where λ is odd (and μ is even). For odd λ, we have

$$e_8(Q, \sigma(Q)) = e_8(\lambda^{-1}Q, \sigma(\lambda^{-1}Q))^{\lambda^2} = e_8(\lambda^{-1}Q, \sigma(\lambda^{-1}Q)),$$

hence we may further reduce to the case where $\lambda = 1$. Now note that

$$\begin{aligned} e_8(P + \mu\sigma(P), \sigma(P) + \mu\sigma^2(P)) &= e_8(P, \sigma(P)) e_8(\sigma(P), \sigma^2(P))^{\mu^2} e_8(P, \sigma^2(P))^{\mu} \\ &= e_8(P, \sigma(P)) e_8(P, \sigma(P))^{4\frac{\mu^2}{4}N(\sigma)} e_8(P, \sigma^2(P))^{2\frac{\mu}{2}} \\ &= e_8(P, \sigma(P)) \cdot (-1)^{\frac{\mu^2}{4}} \cdot (-1)^{\frac{\mu}{2}} \\ &= e_8(P, \sigma(P)), \end{aligned}$$

where in the third equality we used (4.5). □

Remark 4.3.6 If $\mathcal{O}$ is an imaginary quadratic order of discriminant $\mathrm{Disc}(\mathcal{O}) \equiv 0 \bmod 2^5$, then both ϵ and $\delta\epsilon$ and hence $\delta = (\delta\epsilon)\epsilon$ exist, so that $N(\sigma) \equiv 1 \bmod 8$. In this case there is a well-defined group homomorphism $\gamma : \mathrm{Cl}(\mathcal{O}) \to (\mathbf{Z}/8\mathbf{Z})^\times : [\mathfrak{a}] \mapsto N(\mathfrak{a}) \bmod 8$ through which $\delta, \epsilon, \delta\epsilon$ factor. This is the only situation where one can get finer-than-binary modular information about $N(\mathfrak{a})$ modulo a prime power; the above proof shows that we can recover $\gamma([\mathfrak{a}])$ at once as

$$\log_{e_8(P, \iota(\sigma)(P))} e_8(P', \iota'(\sigma)(P')).$$

◊

Remark 4.3.7 In the statements of Theorem 4.1.1, Proposition 4.3.4 and Proposition 4.3.5, the condition that σ be a generator of $\mathcal{O}$ can in fact be relaxed to

$\sigma \in \mathcal{O} \setminus (\mathbf{Z} + m\mathcal{O})$ if m is odd and to $\sigma \in \mathcal{O} \setminus (\mathbf{Z} + 2\mathcal{O})$ if m is even, without modifying the proofs. $\Diamond$

Wrapping up, we have given justification for Algorithm 1 below, evaluating an assigned character $\chi : \mathrm{Cl}(\mathcal{O}) \to \{\pm 1\}$ of modulus m coprime to $\max\{1, \operatorname{char} k\}$ in an unknown ideal class $[\mathfrak{a}]$ connecting two given $\mathcal{O}$-oriented curves (E, ι) and (E', ι'). Here, by the field of definition of (E, ι), (E', ι') we mean any (e.g., the smallest) subfield $F \subseteq k$ over which the curves E, E' and the endomorphisms in $\iota(\mathcal{O}), \iota'(\mathcal{O})$ are defined.

Algorithm 1: Evaluating an assigned character in an unknown ideal class

```
Input:
    $\mathcal{O}$-oriented curves $(E,\iota)$, $(E',\iota')$ in the same orbit with field of definition $F$
    an assigned character $\chi$ of $\mathrm{Cl}(\mathcal{O})$ with modulus $m$ coprime to $\max\{1, \operatorname{char} F\}$
Output:
    $\chi([\mathfrak{a}]) \in \{\pm 1\}$, where $[\mathfrak{a}] \in \mathrm{Cl}(\mathcal{O})$ is such that $(E',\iota') = [\mathfrak{a}](E,\iota)$
1: Find a generator $\sigma$ of $\mathcal{O}$ of norm coprime to $m$.
2: Base-change to the smallest extension $\mathcal{F} \supseteq F$ over which all points in $E[m]$ are
   defined; necessarily, then also all of $E'[m]$ is defined over $\mathcal{F}$.
3: Find a point $P \in E(\mathcal{F})$ such that $E[m] = \langle P, \iota(\sigma)(P) \rangle$ and compute
   $\zeta = e_m(P, \iota(\sigma)(P))$.
4: Likewise, find a point $P' \in E'(\mathcal{F})$ such that $E'[m] = \langle P', \iota'(\sigma)(P') \rangle$ and compute
   $\zeta' = e_m(P', \iota'(\sigma)(P'))$.
5: Inside $\mu_m \subseteq \mathcal{F}^\times$, compute $a = \log_\zeta \zeta'$.
6: If $m$ is an odd prime then recover $\chi([\mathfrak{a}])$ as $\left(\frac{a}{m}\right)$, else recover $\chi([\mathfrak{a}])$ as
       $(-1)^{\frac{a-1}{2}}, \quad (-1)^{\frac{a^2-1}{8}}, \quad (-1)^{\frac{(a+2)^2-9}{8}},$
   depending on whether $\chi = \delta, \epsilon, \delta\epsilon$, respectively.
```

4.4 Complexity and consequences for DDH

Running Algorithm 1 in practice comes with challenges that are specific to our field of definition F. Nevertheless, before going into a more detailed analysis of our main case of interest, namely where F is a finite field, let us add some general comments to its six numbered steps:

1. Very easy, by following the proof of Lemma 4.3.1 or Lemma 4.3.2.

2. The degree of $\mathcal{F}/F$ is $O(m^2)$.[4]

[4] Indeed, by going to at most a quadratic extension, we may assume F is such that $\iota(\mathcal{O}) \subseteq \mathrm{End}_F(E)$. The orientation endows $E[m]$ with the structure of an $\mathcal{O}$-module, and $\mathrm{Aut}_{\mathcal{O}}(E[m]) \cong (\mathcal{O}/m\mathcal{O})^\times$. Since endomorphisms coming from $\mathcal{O}$ are defined over F, hence commute with $\mathrm{Gal}(\mathcal{F}/F)$, we thus obtain a group homomorphism $\mathrm{Gal}(\mathcal{F}/F) \to \mathrm{Aut}_{\mathcal{O}}(E[m])$, which is injective by definition of $\mathcal{F}$. The result follows since $\#(\mathcal{O}/m\mathcal{O})^\times = O(m^2)$.

3.–4. For m an odd prime, the proof of Theorem 4.1.1 shows that the set of m-torsion points that are independent of their image under σ has size $m^2 - m$. So it suffices to try $O(1)$ random points $P \in E[m]$, compute $\iota(\sigma)(P)$ and check whether $e_m(P, \iota(\sigma)(P))$ is a primitive mth root of unity (i.e., not 1).[5]

5. Pollard-ρ type algorithms allow us to compute the discrete logarithm using $O(\sqrt{m})$ operations in μ_m.

6. Trivial.

Theorem 4.4.1 *Let $\mathcal{O} = \mathbf{Z}[\sigma]$ be an imaginary quadratic order and consider two $\mathcal{O}$-oriented elliptic curves (E, ι) and (E', ι') that belong to the same orbit under the action of $\mathrm{Cl}(\mathcal{O})$, say given in Weierstrass form and connected by an unknown ideal class $[\mathfrak{a}]$. Assume that $E, E', \iota(\mathcal{O}), \iota'(\mathcal{O})$ are all defined over a finite field $\mathbf{F}_q$. Let χ be an assigned character of $\mathcal{O}$ with modulus m coprime to q. There exists a randomized algorithm for computing $\chi([\mathfrak{a}])$ that is expected to use*

$$\widetilde{O}(m^3 \log^2 q) \tag{4.6}$$

bit operations and $O(1)$ calls to $\iota(\sigma), \iota'(\sigma)$.

Proof. If we write $f_E(x, y)$ for the defining Weierstrass polynomial of E and $\Psi_{E,m}(x)$ for its m-division polynomial, then the field $\mathcal{F}$ can be constructed as (a quadratic extension of) the splitting field of the resultant $r_{E,m}(x) = \mathrm{res}_y(f_E, \Psi_{E,m})$, whose degree is $O(m^2)$. The division polynomial $\Psi_{E,m}(x)$ can be computed recursively and the resultant $r_{E,m}(x)$ can be factored using Kedlaya–Umans [19]. Using fast arithmetic, this takes a combined time of (4.6). Note that we obtain all points in $E[m]$ as a by-product; once we know $\mathcal{F}$ we can sample points from $E'[m]$ faster. The Weil pairings can be computed using Miller's algorithm, taking $O(\log m)$ operations in $\mathcal{F}$, and Pollard-ρ takes an expected $O(\sqrt{m})$ operations in $\mathcal{F}$, so these costs are dominated by (4.6), again assuming fast arithmetic. Finally, while the norm of the given generator σ may not be coprime to m, from the proofs of Lemma 4.3.1 and Lemma 4.3.2 we see that we can instead work with $\sigma + k$, for some positive integer k bounded by m. Since $\iota(\sigma + k) = \iota(\sigma) + [k]$, the overhead this causes is clearly absorbed by (4.6); and similarly for $\iota'(\sigma + k)$. □

The effectivity of this algorithm co-depends on how easy it is to evaluate $\iota(\sigma)$ and $\iota'(\sigma)$, which is a separate discussion that is captured by the notion of *efficient representations*, see Section 4.5.1 and [32] for more details. One special but interesting case is where $\iota(\sigma)$ equals $\pi_{\mathbf{F}_q}$, or is easily derived from it, whose cost is quasi-quadratic in $m \log q$. So, in this case, the overall cost remains estimated by (4.6). This matches with the asymptotic runtime of the Tate pairing attack from [8], as estimated in [8,

[5] Alternatively, one may opt for a more deterministic approach by computing and analyzing a matrix of $\iota(\sigma)$ acting on $E[m]$, in which case two evaluations of $\iota(\sigma)$ will do. Note however that writing down a matrix of $\iota(\sigma)$ comes at the cost of computing some discrete logarithms.

§5.1].[6]

While the Weil pairing attack is conceptually simpler (no descent of the isogeny volcano needed), in general one should expect the Tate pairing attack to run faster in practice. The main reason is that there it suffices to work over a field $\mathcal{F}$ such that E admits an $\mathcal{F}$-rational point of order m, rather than requiring all m-torsion to be $\mathcal{F}$-rational (in turn, this is because the Tate pairing admits non-trivial self-pairing values, in contrast with the Weil pairing). The degree of such an extension field is bounded by $O(m)$, rather than by $O(m^2)$. But the comparison turns in favour of the Weil pairing as soon as $E[m] \subseteq E(\mathbf{F}_q)$, where no field extension is needed. Note that, here, it makes more sense to measure the cost of a call to $\iota(\sigma), \iota'(\sigma)$ by the cost of evaluating $(\pi_{\mathbf{F}_q} - 1)/m^s$, where s is maximal such that $E[m^s] \subseteq E(\mathbf{F}_q)$; see [25, Lem. 1]. For this we need s successive point divisions by m; the cost of such a division is dominated by that of finding a root of a polynomial of degree m^2, which can be done in time

$$\widetilde{O}(m^2 \log^2 q), \tag{4.7}$$

see [23, §2]. This now becomes the dominant cost of the attack. The asymptotic cost of the Tate pairing also drops to (4.7) in this case, but the Weil pairing attack comes with less overhead.

All this aside, let us re-emphasize that the Weil pairing approach works in far greater generality: for arbitrary orientations and over any field admitting explicit computation. A proof-of-concept implementation of the new method can be found at `https://github.com/KULeuven-COSIC/oriented_DDH`. At the time of publication, this implementation handles the case of $\mathbf{Z}[\sqrt{-p}]$-oriented elliptic curves in characteristic $p \equiv 1 \bmod 4$. We intend to extend the repository in due course, by also covering the higher-degree group actions that were described in [9].

Consequences for DDH

If $\mathrm{Cl}(\mathcal{O})$ admits a non-trivial assigned character whose modulus m is sufficiently small, say polynomially bounded by $\log \mathrm{Disc}(\mathcal{O})$, and if it satisfies $\gcd(m, q) = 1$, then we can use this character to distinguish between random triples and Diffie–Hellman triples with probability $1/2$, as explained in the introduction. So, in this case, we can consider the decisional Diffie–Hellman problem broken for $\mathcal{O}$-oriented elliptic curves over $\mathbf{F}_q$. More generally, if $\mathrm{Cl}(\mathcal{O})$ admits $s \geq 1$ independent such characters (meaning that one cannot use the relation (4.3) to rewrite one of the characters in terms of the others), then we can distinguish with probability $1 - 1/2^s$.

A sufficient condition for the existence of such a character is that $\mathrm{Disc}(\mathcal{O})$ has at least two small odd prime factors different from $p = \mathrm{char}\,\mathbf{F}_q$.[7] Heuristically, we expect that this applies to a density 1 subset of all imaginary quadratic orders when ordered

[6] Here and below, for simplicity, the height $h \approx \mathrm{val}_m(\mathrm{tr}(\pi_{\mathcal{F}})^2 - 4\#\mathcal{F})$ of the m-isogeny volcano of E over $\mathcal{F}$ is estimated by $O(1)$.

[7] In serious cryptographic applications, one can ignore the phrase "different from $p = \mathrm{char}\,\mathbf{F}_q$". Indeed, if $p \mid \mathrm{Disc}(\mathcal{O})$ then E and E' are necessarily supersingular, so if moreover p is small then we can compute $\mathrm{End}(E)$ and $\mathrm{End}(E')$ by navigating through all $O(p)$ nodes of the supersingular isogeny graph. As a result, one is skating on very thin ice (see Section 4.5).

by the absolute value of their discriminant. This can be backed up using Mertens' third theorem; or see [29, III.§6] for more dedicated tools.

As discussed in [8, §6], one can thwart the attack by restricting the class-group action to $\mathrm{Cl}(\mathcal{O})^2$, or at least to a subgroup of $\mathrm{Cl}(\mathcal{O})$ on which all assigned characters of small modulus have trivial evaluations. However, this may have practical consequences in terms of key generation and key validation. Moreover, we do not rule out that the attack can be modified to work for characters whose order is a larger power of 2, e.g., in view of [3, 27]. Quantumly, it is known that 2^r-torsion subgroups, for any small fixed value of r, do not contribute to the hardness of the vectorization problem anyway [5]. Therefore, the cleanest way out is to follow the recommendation from [8, §6], namely to only work with orientations by imaginary quadratic orders whose class number is odd. There may be constructive reasons to deviate from this, e.g., as in the OSIDH protocol [10] where one uses orders of large prime power conductor in an imaginary quadratic field with class number one (such orders always have even class number).

Remark 4.4.2 It is interesting to view Theorem 4.4.1 against the *classical* decisional Diffie–Hellman problem, namely for exponentiation in a group $G = \langle g \rangle$ of some large prime order m. Note that exponentiation defines a free and transitive action of $(\mathbf{Z}/m\mathbf{Z})^\times$ on the set of generators of G. The Legendre symbol

$$\chi : (\mathbf{Z}/m\mathbf{Z})^\times \to \{\pm 1\} : a \mapsto \left(\frac{a}{m}\right)$$

is the unique quadratic character, of modulus m, and if one could cook up an efficient classical way for computing $\chi(a)$ merely from the knowledge of g and g^a, then this would break DDH in this setting. This would be a spectacular result; in general, to the best of our knowledge, we cannot do significantly better than computing a using Pollard-ρ and then evaluating χ at a. This should be compared to steps 5. and 6. from Algorithm 1. In other words, one could say that classical DDH is not weakened by the existence of χ because its modulus is large. ◊

4.5 Reductions to endomorphism ring computation

In this section, we prove that our main result Theorem 4.1.1 allows to significantly improve reductions between computational problems underlying isogeny-based cryptography. It was proved in [31] that two such families of problems are tightly connected: there are computational reductions from action inversion problems (called Effective $\mathcal{O}$-Vectorization or Effective $\mathcal{O}$-Uber) to endomorphism ring computation problems (called $\mathcal{O}$-EndRing and $\mathcal{O}$-EndRing*). However, these reductions are exponential in the worst case. In this section, we apply Theorem 4.1.1 to obtain reductions that are sub-exponential in the worst case, and even polynomial in many regimes of interest. All results in this section that start with (ERH), such as Theorem 4.5.7, assume the extended Riemann hypothesis — precisely, the Riemann hypothesis for Hecke L-functions.

4.5.1 The supersingular endomorphism ring problem

In this section, we assume that the field k is an algebraic closure of a finite field of characteristic p, and that p does not split in $\mathcal{O}$, nor does it divide the conductor of $\mathcal{O}$. Then, the set $\mathscr{E\ell\ell}_{\mathcal{O}}(k)$ is non-empty and all curves in it are supersingular; this set is often denoted by $\mathrm{SS}_{\mathcal{O}}(p)$ in the literature [22, Prop. 3.2]. Recall that a curve E/k is supersingular if and only if its endomorphism ring $\mathrm{End}(E)$ is isomorphic to a maximal order in the quaternion algebra

$$B_{p,\infty} = \left(\frac{-q,-p}{\mathbf{Q}}\right) = \mathbf{Q} + \mathbf{Q}i + \mathbf{Q}j + \mathbf{Q}ij,$$

with the multiplication rules $i^2 = -q$, $j^2 = -p$, and $ji = -ij$, where q is a positive integer that depends on p.

Given a supersingular elliptic curve E over k, the endomorphism ring problem EndRing consists in computing four endomorphisms that form a basis of $\mathrm{End}(E)$. There is flexibility in how these endomorphisms can be represented, but we always assume that it is an *efficient representation.* As in [32], we say that an isogeny $\varphi : E \to E'$ is given in an efficient representation if there is an algorithm to evaluate $\varphi(P)$ for any $P \in E(\mathbf{F}_{p^r})$ in time polynomial in the length of the representation of φ and in $r\log(p)$. We also assume that an efficient representation of φ has length $\Omega(\log(\deg(\varphi)))$.

This endomorphism ring problem is of foundational importance to isogeny-based cryptography: it is presumed to be hard, and this hardness is *necessary* (and sometimes sufficient) for the security of essentially all isogeny-based protocols [17, 7, 16]. It does not, however, capture well the notion of orientation, which plays an important role in many protocols. Therefore, the following oriented variants were introduced in [31]. Computationally, an $\mathcal{O}$-orientation ι is represented by a generator σ of $\mathcal{O}$ (i.e., $\mathcal{O} = \mathbf{Z}[\sigma]$) together with an efficient representation of the endomorphism $\iota(\sigma)$.

Problem 4.5.1 ($\mathcal{O}$-EndRing) *Given $(E,\iota) \in \mathscr{E\ell\ell}_{\mathcal{O}}(k)$, find a basis of* $\mathrm{End}(E)$.

Problem 4.5.2 ($\mathcal{O}$-EndRing*) *Given an $\mathcal{O}$-orientable curve E, find a basis of* $\mathrm{End}(E)$*, and an $\mathcal{O}$-orientation of E expressed in this basis.*

Clearly, $\mathcal{O}$-EndRing reduces to $\mathcal{O}$-EndRing*.

4.5.2 Action inversion problems

Many cryptosystems relate, directly or more subtly, to an inversion problem for the action of $\mathrm{Cl}(\mathcal{O})$ on $\mathscr{E\ell\ell}_{\mathcal{O}}(k)$. In essence, given (E,ι) and (E',ι') in $\mathscr{E\ell\ell}_{\mathcal{O}}(k)$, find a class $[\mathfrak{a}]$ such that $(E',\iota') \cong [\mathfrak{a}](E,\iota)$ (or decide that it does not exist). This is called the vectorization problem. It is too weak for many practical purposes, because knowledge of the class $[\mathfrak{a}]$ is not sufficient to efficiently apply its action on any other $\mathcal{O}$-oriented curve. Therefore, the following stronger problem was introduced in [31].

Problem 4.5.3 (Effective $\mathcal{O}$-Vectorization) *Given three $\mathcal{O}$-oriented supersingular curves $(E,\iota),(E',\iota'),(F,\jmath) \in \mathcal{E}\ell\ell_{\mathcal{O}}(k)$, find an $\mathcal{O}$-ideal $\mathfrak{a}$ (or decide that it does not exist) such that $(E',\iota') \cong [\mathfrak{a}](E,\iota)$, and an efficient representation of $\varphi_{\mathfrak{a}} : (F,\jmath) \to [\mathfrak{a}](F,\jmath)$.*

The security of many cryptosystems directly reduces to this problem, such as CSIDH [6], CSI-FiSh [1], CSURF [4], or other generalizations [9].

One can define a similar problem where no orientation is provided for E'. Then, one cannot require $(E',\iota') \cong [\mathfrak{a}](E,\iota)$ anymore, but one can still ask for $E' \cong [\mathfrak{a}]E$. The resulting *Uber isogeny problem* was introduced in [14].

Problem 4.5.4 (Effective $\mathcal{O}$-Uber) *Given two $\mathcal{O}$-oriented curves $(E,\iota),(F,\jmath) \in \mathcal{E}\ell\ell_{\mathcal{O}}(k)$ and an $\mathcal{O}$-orientable curve E', find an $\mathcal{O}$-ideal $\mathfrak{a}$ such that $E' \cong [\mathfrak{a}]E$, and an efficient representation of $\varphi_{\mathfrak{a}} : (F,\jmath) \to [\mathfrak{a}](F,\jmath)$.*

This Effective $\mathcal{O}$-Uber problem is significantly harder than the Effective $\mathcal{O}$-Vectorization problem. In fact, most isogeny-based cryptosystems reduce to an instance of Effective $\mathcal{O}$-Uber [14], even cryptosystems such as SIDH [18] which, at first sight, do not seem to involve any orientation.

4.5.3 Action inversion reduces to endomorphism ring

Strengthening and generalizing a result of [7], it was proved in [31] that Effective $\mathcal{O}$-Vectorization reduces to $\mathcal{O}$-EndRing, and that Effective $\mathcal{O}$-Uber reduces to $\mathcal{O}$-EndRing*. Both reductions are in polynomial time in the length of the instance, and in $\#(\mathrm{Cl}(\mathcal{O})[2])$. Unfortunately, the dependence on $\#(\mathrm{Cl}(\mathcal{O})[2])$ means that the reduction is, in the worst case, exponential in the size of the input, since $\#(\mathrm{Cl}(\mathcal{O})[2])$ could be as large as $D^{1/\log\log D}$, where $D = |\operatorname{Disc}(\mathcal{O})|$. The issue is the following: given two oriented curves (E,ι) and (E',ι') as in the definition of Effective $\mathcal{O}$-Vectorization, the reductions first find a class $[\mathfrak{a}]^2$ such that $(E',\iota') \cong [\mathfrak{a}](E,\iota)$. Finding $[\mathfrak{a}]$ from $[\mathfrak{a}]^2$ is a square root computation. There are $\#(\mathrm{Cl}(\mathcal{O})[2])$ square roots of $[\mathfrak{a}]^2$, but only one is the correct class $[\mathfrak{a}]$. In [31], one simply does an exhaustive search. Now, thanks to Theorem 4.1.1, there is a much more efficient way to find the correct square root, which in the worst case is sub-exponential in $\operatorname{Disc}(\mathcal{O})$. This is the following proposition. Recall the L-notation

$$L_x(\alpha) = \exp\left(O\left((\log x)^{\alpha}(\log\log x)^{1-\alpha}\right)\right)$$

for sub-exponential complexities.

Proposition 4.5.5 (ERH) *Given $\mathcal{O}$ of discriminant $-D$, the factorization $D = \prod_{i=1}^{\omega(D)} \ell_i^{e_i}$ (with $\ell_i < \ell_{i+1}$), two $\mathcal{O}$-oriented elliptic curves $(E,\iota),(E',\iota') \in \mathcal{E}\ell\ell_{\mathcal{O}}(k)$, a basis of* $\operatorname{End}(E)$, *and an $\mathcal{O}$-ideal $\mathfrak{a}$ for which there exists an ideal class $[\mathfrak{c}]$ such that $[\mathfrak{a}] = [\mathfrak{c}]^2$ and $(E',\iota') = [\mathfrak{c}](E,\iota)$, one can find a representative for the ideal class $[\mathfrak{c}]$ in* *probabilistic polynomial time in the length of the input and in*[8]

$$\min\left(2^{\omega(D)}, \max_i\left(\ell_i \mid \ell_i \leq 2^{\omega(D)-i}\right)\right) \ll \min\left(L_D(1/2), \#(\mathrm{Cl}(\mathcal{O})[2]), \ell_{\omega(D)}\right).$$

Here, by "probabilistic polynomial time in the length of the input", we mean probabilistic polynomial time in $\log p$*,* $\log D$*,* $\log N(\mathfrak{a})$*, the lengths of* ι *and* ι'*, and in the length of the basis of* $\mathrm{End}(E)$*.*

Before proving it, let us recall the following proposition from [31].

Proposition 4.5.6 (ERH, [31, Proposition 9]) *Given* $(E,\iota) \in \mathcal{E}\ell\ell_{\mathcal{O}}(k)$*, a basis of* $\mathrm{End}(E)$*, and an* $\mathcal{O}$*-ideal* $\mathfrak{a}$*, one can compute* $[\mathfrak{a}](E,\iota)$ *and an efficient representation of* $\varphi_{\mathfrak{a}} : (E,\iota) \to [\mathfrak{a}](E,\iota)$ *in probabilistic polynomial time in the length of the input. That is, in* $\log|\mathrm{Disc}(\mathcal{O})|$*,* $\log p$*,* $\log N(\mathfrak{a})$*, and in the length of* ι *and of the basis of* $\mathrm{End}(E)$*.*

Proof of Proposition 4.5.5. Let $B > 0$ be a bound to be tuned later. Consider the sets of prime numbers

$$\begin{aligned} P_1 &= \{\ell \mid \ell \text{ is an odd prime factor of } \mathrm{Disc}(\mathcal{O}) \text{ and } \ell \leq B\}, \text{ and} \\ P_2 &= \{\ell \mid \ell \text{ is an odd prime factor of } \mathrm{Disc}(\mathcal{O}) \text{ and } \ell > B\}. \end{aligned}$$

For each $\ell \in P_1$, compute $\chi_\ell([\mathfrak{c}])$ in time $\ell^{O(1)}$ using Theorem 4.4.1 and the fact that $(E',\iota') = [\mathfrak{c}](E,\iota)$. Now, with [3], one can compute square roots in $\mathrm{Cl}(\mathcal{O})$ in polynomial time, so we get an ideal $\mathfrak{a}$ such that $[\mathfrak{a}]$ and $[\mathfrak{c}]$ differ by a two-torsion factor. From [3], one also gets a basis of $\mathrm{Cl}(\mathcal{O})[2]$, so we can ensure that $\chi_\ell([\mathfrak{a}]) = \chi_\ell([\mathfrak{c}])$ for each $\ell \in P_1$. The solution is now of the form $[\mathfrak{c}] = [\mathfrak{a}][\mathfrak{b}]$ where $[\mathfrak{b}]$ is in the subgroup G of $\mathrm{Cl}(\mathcal{O})[2]$ of classes such that $\chi_\ell([\mathfrak{b}]) = 1$ for all $\ell \in P_1$. Therefore, the number of remaining candidates for the class $[\mathfrak{c}]$ is $\#G \leq 2^{\#P_2+1}$. These can be enumerated (from the basis of $\mathrm{Cl}(\mathcal{O})[2]$, deduce a basis of the subgroup G) and checked for correctness in polynomial time using Proposition 4.5.6 and the provided basis of $\mathrm{End}(E)$. Overall, the running time is polynomial in $\log p$, $\log|\mathrm{Disc}(\mathcal{O})|$, B, and $2^{\#P_2}$. The running time follows by choosing $B = \min\left(2^{\omega(D)}, \max_i\left(\ell_i \mid \ell_i \leq 2^{\omega(D)-i}\right)\right)$.

Let us prove the last inequality. First, $2^{\omega(D)} \ll \#(\mathrm{Cl}(\mathcal{O})[2])$, so $B \ll \#(\mathrm{Cl}(\mathcal{O})[2])$. Second, if $\{\ell_i \mid \ell_i \leq 2^{\omega(D)-i}\}$ is empty, then $2^{\omega(D)-1} < \ell_1 \leq \ell_{\omega(D)}$ so $2^{\omega(D)} \ll \ell_{\omega(D)}$. If it is not empty, clearly $\max_i\left(\ell_i \mid \ell_i \leq 2^{\omega(D)-i}\right) \ll \ell_{\omega(D)}$. In both cases, we deduce $B \ll \ell_{\omega(D)}$. Lastly, it remains to see that $B \ll L_D(1/2)$. Suppose there exists j such that $\ell_j = \max_i\left(\ell_i \mid \ell_i \leq 2^{\omega(D)-i}\right)$. We have $\log_2(\ell_j) \leq \omega(D) - j$, and

$$\log_2(D) \geq \sum_{i=j+1}^{\omega(D)} \log_2(\ell_i) \geq (\omega(D) - j)\log_2(\ell_j) \geq \log_2(\ell_j)^2.$$

[8] With the convention that $\max(\emptyset) = +\infty$.

We deduce that $\ell_j \leq 2^{\log_2(D)^{1/2}}$, hence $B \ll L_D(1/2)$. If there exists no such j, then

$$\log_2(D) \geq \sum_{i=1}^{\omega(D)} \log_2(\ell_i) \geq \sum_{i=1}^{\omega(D)} (\omega(D) - i) = \Theta(\omega(D)^2),$$

so $2^{\omega(D)} = L_D(1/2)$, hence $B \ll L_D(1/2)$. □

The main result of this section is the following theorem.

Theorem 4.5.7 (ERH, reduction of Effective $\mathcal{O}$-Vectorization to $\mathcal{O}$-EndRing) *Given an order $\mathcal{O}$ of discriminant $-D$, the factorization $D = \prod_{i=1}^{\omega(D)} \ell_i^{e_i}$ (with $\ell_i < \ell_{i+1}$), three $\mathcal{O}$-oriented elliptic curves (E,ι), (E',ι'), $(F,\jmath) \in \mathcal{E}\ell\ell_{\mathcal{O}}(k)$, together with bases of* $\mathrm{End}(E)$, $\mathrm{End}(E')$ *and* $\mathrm{End}(F)$, *one can compute (or assert that it does not exist) an $\mathcal{O}$-ideal $\mathfrak{c}$ such that $(E',\iota') = [\mathfrak{c}](E,\iota)$ and an efficient representation of $\varphi_{\mathfrak{c}} : (F,\jmath) \to [\mathfrak{c}](F,\jmath)$ in probabilistic polynomial time in the length of the input and in*

$$\min\left(2^{\omega(D)}, \max_i\left(\ell_i \mid \ell_i \leq 2^{\omega(D)-i}\right)\right) \ll \min\left(L_D(1/2), \#(\mathrm{Cl}(\mathcal{O})[2]), \ell_{\omega(D)}\right).$$

Here, by "probabilistic polynomial time in the length of the input", we mean probabilistic polynomial time in $\log p$, $\log D$, *the lengths of ι, ι', and $\jmath$, and in the lengths of the bases of* $\mathrm{End}(E)$, $\mathrm{End}(E')$, *and* $\mathrm{End}(F)$.

Remark 4.5.8 This improves the result of [31, Thm. 2] in two ways. First, the worst case is now sub-exponential: when D is primorial, the running time of [31, Thm. 2] could reach about $D^{1/\log\log D}$, while it is now always at most $L_D(1/2)$. Second, Theorem 4.5.7 is now very efficient for a new important family of discriminants: when almost all prime divisors of D are small, no matter how many there are. In particular, primorial numbers (the worst case of [31, Thm. 2]) now benefit from a polynomial time algorithm. ♢

Proof. Thanks to Proposition 4.5.5, the proof is a straightforward adaptation of the proof of [31, Thm. 2]. Suppose we are given $(E,\iota), (E',\iota') \in \mathcal{E}\ell\ell_{\mathcal{O}}(k)$, together with $\mathrm{End}(E)$ and $\mathrm{End}(E')$. Consider the involution $\tau_p : \mathcal{E}\ell\ell_{\mathcal{O}}(k) \to \mathcal{E}\ell\ell_{\mathcal{O}}(k)$ defined in [31, Def. 7] as $\tau_p(E,\iota) = (E^{(p)}, (\phi_p)_*\bar{\iota})$, where $\bar{\iota}$ is the conjugate of ι (i.e., $\bar{\iota}(\alpha) = \iota(\overline{\alpha})$ for any $\alpha \in \mathcal{O}$), and $\phi_p : E \to E^{(p)}$ is the Frobenius isogeny.

Then, per [31, Prop. 11], one can compute $\mathfrak{a}$ and $\mathfrak{b}$ such that $\tau_p(E,\iota) = [\mathfrak{a}](E,\iota)$ and $\tau_p(E',\iota') = [\mathfrak{b}](E',\iota')$ in polynomial time. From [31, Lem. 10], the ideal class of $\mathfrak{c}$ is one of the $\#(\mathrm{Cl}(\mathcal{O})[2])$ square roots of $[\mathfrak{a}\overline{\mathfrak{b}}]$. Therefore, the ideal $\mathfrak{c}$ can be found by Proposition 4.5.5 within the claimed running time. Finally, compute an efficient representation of $\varphi_{\mathfrak{c}} : (F,\jmath) \to [\mathfrak{c}](F,\jmath)$ in polynomial time with Proposition 4.5.6. □

Corollary 4.5.9 (ERH) *Given an order $\mathcal{O}$ of discriminant $-D$, and the factorization $D = \prod_{i=1}^{\omega(D)} \ell_i^{e_i}$ (with $\ell_i < \ell_{i+1}$),* Effective $\mathcal{O}$-Uber *reduces to* $\mathcal{O}$-EndRing* *in* *probabilistic polynomial time in the length of the instance and in*

$$\min\left(2^{\omega(D)}, \max_i\left(\ell_i \mid \ell_i \leq 2^{\omega(D)-i}\right)\right) \ll \min\left(L_D(1/2), \#(\mathrm{Cl}(\mathcal{O})[2]), \ell_{\omega(D)}\right).$$

Here, by "probabilistic polynomial time in the length of the instance", we mean probabilistic polynomial time in $\log p$*,* $\log D$*, and in the lengths of the orientations.*

Proof. Again, this is a straightforward adaptation of [31, Cor. 4]. Suppose we are given $(E,\iota),(F,\jmath) \in \mathcal{E}\ell\ell_{\mathcal{O}}(k)$ and an $\mathcal{O}$-orientable elliptic curve E'. Solving $\mathcal{O}$-EndRing*, one can find ε-bases of $\mathrm{End}(E)$, $\mathrm{End}(F)$ and $\mathrm{End}(E')$, and an $\mathcal{O}$-orientation ι' of E'. The result follows from Theorem 4.5.7. □

4.6 Bibliography

[1] Ward Beullens, Thorsten Kleinjung, and Frederik Vercauteren. CSI-FiSh: Efficient isogeny based signatures through class group computations. In *Asiacrypt (1)*, volume 11921 of *Lecture Notes in Computer Science*, pages 227–247. Springer, 2019. `https://ia.cr/2018/485`.

[2] Dan Boneh, Dmitry Kogan, and Katharine Woo. Oblivious pseudorandom functions from isogenies. In *Asiacrypt (2)*, volume 12492 of *Lecture Notes in Computer Science*, pages 520–550. Springer, 2020. `https://ia.cr/2020/1532`.

[3] Wieb Bosma and Peter Stevenhagen. On the computation of quadratic 2-class groups. *J. Théor. Nombres Bordeaux*, 8(2):283–313, 1996.

[4] Wouter Castryck and Thomas Decru. CSIDH on the surface. In Jintai Ding and Jean-Pierre Tillich, editors, *PQCrypto 2020*, volume 12100 of *Lecture Notes in Computer Science*, pages 111–129. Springer, 2020.

[5] Wouter Castryck, Ann Dooms, Carlo Emerencia, and Alexander Lemmens. A fusion algorithm for solving the hidden shift problem in finite abelian groups. In *PQCrypto*, volume 12841 of *Lecture Notes in Computer Science*, pages 133–153. Springer, 2021. `https://eprint.iacr.org/2021/562`.

[6] Wouter Castryck, Tanja Lange, Chloe Martindale, Lorenz Panny, and Joost Renes. CSIDH: An efficient post-quantum commutative group action. In *Asiacrypt 2018 Pt. 3*, volume 11274 of *Lecture Notes in Computer Science*, pages 395–427. Springer, 2018.

[7] Wouter Castryck, Lorenz Panny, and Frederik Vercauteren. Rational isogenies from irrational endomorphisms. In *Eurocrypt (2)*, volume 12106 of *Lecture Notes in Computer Science*, pages 523–548. Springer, 2020. `https://ia.cr/2019/1202`.

[8] Wouter Castryck, Jana Sotáková, and Frederik Vercauteren. Breaking the decisional Diffie-Hellman problem for class group actions using genus theory. In *Crypto (2)*, volume 12171 of *Lectures Notes in Computer Science*, pages 92–120. Springer, 2020. `https://ia.cr/2020/151`.

[9] Mathilde Chenu and Benjamin Smith. Higher-degree supersingular group actions. In *MathCrypt*, J. Math. Cryptol. (to appear), 2021. `https://ia.cr/2021/955`.

[10] Leonardo Colò and David Kohel. Orienting supersingular isogeny graphs. *Journal of Mathematical Cryptolology*, 14(1):414–437, 2020.

[11] Jean-Marc Couveignes. Hard homogeneous spaces, 2006. Unpublished article, available at `https://eprint.iacr.org/2006/291`.

[12] David A. Cox. *Primes of the form $x^2 + ny^2$: Fermat, class field theory, and complex multiplication.* Pure and Applied Mathematics. Wiley, second edition, 2013.

[13] Pierrick Dartois and Luca De Feo. On the security of OSIDH. In *PKC (1)*, volume 13177 of *Lecture Notes in Computer Science*, pages 52–81. Springer, 2022. https://ia.cr/2021/1681.

[14] Luca De Feo, Cyprien Delpech de Saint Guilhem, Tako Boris Fouotsa, Péter Kutas, Antonin Leroux, Christophe Petit, Javier Silva, and Benjamin Wesolowski. Séta: Supersingular encryption from torsion attacks. In *Asiacrypt (4)*, volume 13093 of *Lecture Notes in Computer Science*, pages 249–278. Springer, 2021. https://ia.cr/2019/1291.

[15] Luca De Feo and Michael Meyer. Threshold schemes from isogeny assumptions. In *PKC (2)*, volume 12111 of *Lecture Notes in Computer Science*, pages 187–212. Springer, 2020. https://ia.cr/2019/1288.

[16] Tako Boris Fouotsa, Péter Kutas, Simon-Philipp Merz, and Yan Bo Ti. On the isogeny problem with torsion point information. In *PKC (1)*, volume 13177 of *Lecture Notes in Computer Science*, pages 142–161. Springer, 2022. https://ia.cr/2021/153.

[17] Steven D. Galbraith, Christophe Petit, Barak Shani, and Yan Bo Ti. On the security of supersingular isogeny cryptosystems. In *Asiacrypt (1)*, volume 10031 of *Lecture Notes in Computer Science*, pages 63–91. Springer, 2016. https://ia.cr/2016/859.

[18] David Jao and Luca De Feo. Towards quantum-resistant cryptosystems from supersingular elliptic curve isogenies. In *PQCrypto*, volume 7071 of *Lecture Notes in Computer Science*, pages 19–34. Springer, 2011. https://ia.cr/2011/506.

[19] Kiran S. Kedlaya and Christopher Umans. Fast polynomial factorization and modular composition. In *IEEE FOCS 2008*, pages 146–155, 2008. http://users.cms.caltech.edu/~umans/papers/KU08-final.pdf.

[20] Yi-Fu Lai, Steven D. Galbraith, and Cyprien Delpech de Saint Guilhem. Compact, efficient and UC-secure isogeny-based oblivious transfer. In *Eurocrypt (1)*, volume 12696 of *Lecture Notes in Computer Science*, pages 213–241. Springer, 2021. https://ia.cr/2020/1012.

[21] James S. Milne. Complex multiplication (v0.10), 2020. https://www.jmilne.org/math/CourseNotes/cm.html.

[22] Hiroshi Onuki. On oriented supersingular elliptic curves. *Finite Fields Appl.*, 69:Paper No. 101777, 18, 2021.

[23] Michael Rabin. Probabilistic algorithms in finite fields. *SIAM J. Comput.*, 9(2):273–280, 1980. http://publications.csail.mit.edu/lcs/pubs/pdf/MIT-LCS-TR-213.pdf.

[24] Alexander Rostovtsev and Anton Stolbunov. Public-key cryptosystem based on isogenies. IACR Cryptology ePrint Archive 2006/145, 2006. https://ia.cr/2006/145.

[25] Hans-Georg Rück. A note on elliptic curves over finite fields. *Math. Comp.*, 49(179):301–304, 1987.

[26] René Schoof. Nonsingular plane cubic curves over finite fields. *J. Combin. Theory Ser. A*, 46(2):183–211, 1987.

[27] Peter Stevenhagen. Rédei-matrices and applications. In *Number theory (Paris, 1992–1993)*, volume 215 of *London Math. Soc. Lecture Note Ser.*, pages 245–259. Cambridge Univ. Press, Cambridge, 1995.

[28] Anton Stolbunov. Cryptographic schemes based on isogenies. 2012. PhD thesis. `https://ntnuopen.ntnu.no/ntnu-xmlui/bitstream/handle/11250/262577/529395_FULLTEXT01.pdf`.

[29] Gérald Tenenbaum. *Introduction to analytic and probabilistic number theory*, volume 163 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, third edition, 2015. Translated from the 2008 French edition by Patrick D. F. Ion.

[30] William C. Waterhouse. Abelian varieties over finite fields. *Ann. Sci. École Norm. Sup.*, 2:521–560, 1969.

[31] Benjamin Wesolowski. Orientations and the supersingular endomorphism ring problem. In *Eurocrypt (3)*, volume 13277 of *Lecture Notes in Computer Science*, pages 345–371. Springer, 2022.

[32] Benjamin Wesolowski. The supersingular isogeny path and endomorphism ring problems are equivalent. In *IEEE FOCS 2021*, pages 1100–1111, 2022.

Chapter 5

Weak instances of class group action based cryptography via self-pairings

This chapter consists of a paper written together with Wouter Castryck, Simon-Philipp Merz, Marzio Mula, Sam van Buuren, and Frederik Vercauteren. It has been published as

Wouter Castryck, Marc Houben, Simon-Philipp Merz, Marzio Mula, Sam van Buuren, and Frederik Vercauteren. Weak instances of class group action based cryptography via self-pairings. In *Advances in Cryptology – CRYPTO 2023*, pages 762–792, Lecture Notes in Computer Science, vol 14083. Springer, Cham. https://doi.org/10.1007/978-3-031-38548-3_25.

All authors of this paper contributed equally to the work.

Compared to the published version we added Section 5.8, which appears as Appendix A in the eprint version [5]. We also fixed some typos. The numbering (of e.g. theorems and definitions) in the published version is different.

Acknowledgements

We owe thanks to Luca De Feo, Damien Robert, Katherine Stange and the anonymous reviewers for various helpful comments, discussions and suggestions. We thank Abel Laval for pointing out a typo in Equation (5.7).

Abstract

In this paper we study non-trivial self-pairings with cyclic domains that are compatible with isogenies between elliptic curves oriented by an imaginary quadratic order $\mathcal{O}$. We prove that the order m of such a self-pairing necessarily satisfies $m \mid \Delta_{\mathcal{O}}$ (and even $2m \mid \Delta_{\mathcal{O}}$ if $4 \mid \Delta_{\mathcal{O}}$ and $4m \mid \Delta_{\mathcal{O}}$ if $8 \mid \Delta_{\mathcal{O}}$) and is not a multiple of the field characteristic. Conversely, for each m satisfying these necessary conditions, we construct a family of non-trivial cyclic self-pairings of order m that are compatible with oriented isogenies, based on generalized Weil and Tate pairings.

As an application, we identify weak instances of class group actions on elliptic curves assuming the degree of the secret isogeny is known. More in detail, we show that if $m^2 \mid \Delta_{\mathcal{O}}$ for some prime power m then given two primitively $\mathcal{O}$-oriented elliptic curves (E, ι) and $(E', \iota') = [\mathfrak{a}](E, \iota)$ connected by an unknown invertible ideal $\mathfrak{a} \subseteq \mathcal{O}$, we can recover $\mathfrak{a}$ essentially at the cost of a discrete logarithm computation in a group of order m^2, assuming the norm of $\mathfrak{a}$ is given and is smaller than m^2. We give concrete instances, involving ordinary elliptic curves over finite fields, where this turns into a polynomial time attack.

Finally, we show that these self-pairings simplify known results on the decisional Diffie–Hellman problem for class group actions on oriented elliptic curves.

5.1 Introduction

Isogeny based cryptography using class group actions was originally proposed in the works of Couveignes [13] and Rostovtsev–Stolbunov [32] (CRS), and both use ordinary elliptic curves. In particular, let $\mathcal{O}$ be an order in an imaginary quadratic number field K, then there is a natural action of the ideal-class group $\mathrm{Cl}(\mathcal{O})$ on the set of ordinary elliptic curves (up to isomorphism) over a finite field $\mathbf{F}_q$ whose endomorphism ring is isomorphic to $\mathcal{O}$. Since it is difficult to construct ordinary elliptic curves with many small rational subgroups and large enough $\mathrm{Cl}(\mathcal{O})$, computing the class group action in CRS is rather slow. CSIDH [7, 3] significantly improved the efficiency of the CRS approach by considering the set of supersingular elliptic curves over a large prime field $\mathbf{F}_p$ and restricting to the $\mathbf{F}_p$-rational endomorphisms. These form a subring of the full endomorphism ring which again is isomorphic to an order $\mathcal{O}$ in an imaginary quadratic number field. Since $\#E(\mathbf{F}_p) = p + 1$ for such supersingular elliptic curves, it now becomes trivial to force the existence of small rational subgroups by choosing p such that $p + 1$ has small prime factors. The OSIDH protocol by Colò and Kohel [12] (and more rigorously by Onuki [27]) extended this even further by using oriented elliptic curves: here one considers elliptic curves together with an $\mathcal{O}$-orientation, which is simply an injective ring homomorphism $\iota : \mathcal{O} \hookrightarrow \mathrm{End}(E)$. OSIDH provides a convenient unifying framework for CRS and CSIDH, but also contains many new families of potential cryptographic interest. While the original Colò–Kohel proposal does not seem viable [15], a more recent proposal [16] looks promising.

A different approach to isogeny based cryptography is taken by SIDH [21], which relies on random walks in the isogeny graph of supersingular elliptic curves over $\mathbf{F}_{p^2}$. To make the protocol work however, it needs to reveal the action of the secret isogeny $\phi : E \to E'$ on a basis of $E[m]$, where m typically is a power of 2 or 3. This extra information was recently exploited in a series of papers [30, 4, 23] resulting in a polynomial time attack on SIDH. This attack not only showed that SIDH is totally insecure, but also added a very powerful technique to the isogeny toolbox: it is possible to recover a secret isogeny $\phi : E \to E'$ between two elliptic curves E and E', all defined over a finite field $\mathbf{F}_q$, in polynomial time if the following information is available:

- the action of ϕ on a basis of $E[m]$ is given where m is sufficiently smooth,
- the degree $d = \deg(\phi)$ is known and coprime with m,
- $m^2 > d$.

The origins of this paper trace back to the simple question: to what extent can the above technique be applied to the class group action setting and are there weak instances where this results in a polynomial time attack? To illustrate which problems need to be solved, we will focus on the CSIDH setting (the more general oriented case is deferred to later sections). In particular, assume E and E' are two supersingular elliptic curves over $\mathbf{F}_p$ connected by a secret isogeny $\phi : E \to E' := [\mathfrak{a}]E$ with $\ker(\phi) = E[\mathfrak{a}]$ and $\mathfrak{a} \subseteq \mathcal{O}$ an invertible ideal. To be able to apply the above technique to recover ϕ, we need to know the degree of ϕ and its action on a basis of $E[m]$ for some smooth m.

Whether the degree of ϕ is known depends on how the class group action is implemented, e.g. in side-channel protected implementations, the degree is sometimes fixed and thus known. For example, this may be the case for the "dummy-free" constant-time variant of CSIDH that was proposed in [9]. In CSIDH variants that employ dummy computations to achieve constant-time, fault attacks that skip isogeny computations could allow an attacker to determine whether an isogeny was a dummy computation or not, and thus deduce information about the private key. In the dummy-free approach the parity of each secret exponent e_i in CSIDH is fixed and sampled from an interval $[-e, e]$. For $e = 1$, which was suggested both in [9] and in [10], the degree of any secret isogeny is thus fixed to a publicly known value, i.e. the product of all the split primes used in the CSIDH group action. In the remainder of the paper, we will assume the degree of ϕ is known. Note that by construction, the degree is automatically smooth, so this does not impose a further restriction.

Determining the action of the secret isogeny ϕ on a basis of $E[m]$ for a chosen m is a somewhat more challenging task, since we only have E, E' and the degree of ϕ at our disposal. To make partial progress, note that we can choose $m = \ell^r$ for some small odd prime ℓ not dividing $d = \deg(\phi)$ that splits in $\mathbf{Q}(\sqrt{-p})$. Then $E[m]$ is spanned by two eigenspaces $\langle P\rangle, \langle Q\rangle$ of the Frobenius endomorphism π_p corresponding to two different eigenvalues. Since ϕ commutes with π_p, $E'[m]$ will also be spanned by two eigenspaces $\langle P'\rangle, \langle Q'\rangle$ of π_p on E' corresponding to these same eigenvalues, so we already have that $\langle P'\rangle = \langle\phi(P)\rangle$ and $\langle Q'\rangle = \langle\phi(Q)\rangle$. In particular, there exist units $\lambda, \mu \in \mathbf{Z}/m\mathbf{Z}$ such that $P' = \lambda\phi(P)$ and $Q' = \mu\phi(Q)$. Using the independence of the points P and Q (resp. P' and Q') and compatibility of the classical Weil pairing e_m with isogenies, we obtain

$$e_m(P', Q') = e_m(\lambda\phi(P), \mu\phi(Q)) = e_m(P, Q)^{\lambda\mu d}\,.$$

By computing a discrete logarithm (note that ℓ is assumed small, so computing the discrete logarithm is easy), we can therefore eliminate one variable, say μ, since d is assumed known, so we are left with determining λ. It is tempting to use the same trick again by pairing P' with itself, which would lead to

$$e_m(P', P') = e_m(\lambda\phi(P), \lambda\phi(P)) = e_m(P, P)^{\lambda^2 d}\,.$$

Unfortunately, the classical Weil pairing e_m results in a trivial self-pairing, i.e. we always have $e_m(P, P) = 1$. What we thus require is a non-trivial self-pairing f_m compatible with isogenies, which implies $f_m(\phi(P)) = f_m(P)^d$, and thus $f_m(P') = f_m(P)^{\lambda^2 d}$, with both sides of order m say. We thus recover λ up to sign and as such we can recover $\pm\phi$. The existence of non-trivial self-pairings therefore is crucial to the success of the attack.

Contributions

- We give a self-contained overview of generalized Weil [20] and Tate [2] pairings, filling some gaps in the existing literature and relating both pairings by extending a result in [20]. Although these generalized pairings are more powerful than the classical Weil and Tate pairings, they do not seem to be well known in the cryptographic community.

- We formally define a cyclic self-pairing of order m on an elliptic curve E to be a homogeneous degree-2 function $f_m : C \to \mu_m$ with cyclic domain $C \subseteq E$ such that $\mathrm{im}(f_m)$ spans μ_m. We derive necessary conditions for the existence of non-trivial cyclic self-pairings of order m on $\mathcal{O}$-oriented elliptic curves that are compatible with oriented isogenies. In particular, we show that m cannot be a multiple of the field characteristic and that $m \mid \Delta_{\mathcal{O}}$, with $\Delta_{\mathcal{O}}$ the discriminant of $\mathcal{O}$ (and even $2m \mid \Delta_{\mathcal{O}}$ if $4 \mid \Delta_{\mathcal{O}}$ and $4m \mid \Delta_{\mathcal{O}}$ if $8 \mid \Delta_{\mathcal{O}}$). Note that our results only apply to self-pairings compatible with isogenies, which is required to make the above attack work. This is in stark contrast to considering an individual elliptic curve, where non-trivial cyclic self-pairings of order m always exist (as soon as m is not a multiple of the field characteristic), e.g. by choosing any cyclic order-m subgroup $C = \langle P \rangle$ and simply defining $f_m(\lambda P) = \zeta_m^{\lambda^2}$ with ζ_m some fixed primitive m-th root of unity.

- For m satisfying these necessary conditions we construct cyclic self-pairings of order m compatible with oriented isogenies, based on generalized Weil and Tate pairings.

- Using these non-trivial cyclic self-pairings, we are the first to identify weak instances of class group action based cryptography. In the best case, we obtain a polynomial time attack on the vectorization problem when $\deg(\phi)$ is known and powersmooth, $\ell^{2r} \mid q-1$, $E(\mathbf{F}_q)[\ell^\infty]$ is cyclic of order at least ℓ^{2r}, and $\ell^{2r} > \deg(\phi)$. This for instance would be the case if one would use a setup like SiGamal [26], but using the group action underlying CRS instead of CSIDH. Note however that our attack does not apply to SiGamal itself for two major reasons: here $\Delta_{\mathcal{O}} = -4p$ and the degree of the secret isogeny is not known.

- We present a more elegant version of existing results [8, 6] on the decisional Diffie–Hellman problem for class group actions. In particular, in Remark 5.5.3 we give a conceptual explanation for a phenomenon observed in [8, App. A]. This also illustrates why the general framework of oriented elliptic curves can be useful even if one is only interested in elliptic curves over $\mathbf{F}_q$ equipped with the natural Frobenius orientation.

5.2 Background

Throughout this paper, k denotes a perfect field (e.g., a finite field $\mathbf{F}_q$) with algebraic closure $\overline{k}$, and K is an imaginary quadratic number field with maximal order $\mathcal{O}_K$.

5.2.1 Oriented elliptic curves

Our main references are Colò–Kohel [12] and Onuki [27], although we present matters in somewhat greater generality (in the sense that we also cover non-supersingular elliptic curves). A *K-orientation* on an elliptic curve E/k is an injective ring homomorphism

$$\iota : K \hookrightarrow \mathrm{End}^0(E) := \mathrm{End}(E) \otimes_{\mathbf{Z}} \mathbf{Q},$$

where $\mathrm{End}(E)$ denotes the full ring of endomorphisms of E (i.e., defined over $\overline{k}$). The couple (E, ι) is called a *K-oriented elliptic curve.*

Example 5.2.1 The standard example to keep in mind is that of an elliptic curve E over a finite field $\mathbf{F}_q$ for which the q-th power Frobenius endomorphism π_q is not a scalar multiplication (that is, we exclude supersingular elliptic curves $E/\mathbf{F}_{p^{2r}}$ on which Frobenius acts as $[\pm p^r]$). In that case we have an orientation

$$\iota : \mathbf{Q}(\sigma) \hookrightarrow \mathrm{End}^0(E) : \sigma \mapsto \pi_q, \qquad \sigma = \frac{t_E + \sqrt{t_E^2 - 4q}}{2} \tag{5.1}$$

with t_E the trace of Frobenius of E over $\mathbf{F}_q$. We call this the *Frobenius orientation.* If (and only if) E is ordinary then ι is an isomorphism. If E is supersingular then the image of ι is the subalgebra $\mathrm{End}_q^0(E) = \mathrm{End}_q(E) \otimes_{\mathbf{Z}} \mathbf{Q}$, with $\mathrm{End}_q(E)$ the ring of $\mathbf{F}_q$-rational endomorphisms of E. By abuse of notation, we will occasionally just identify σ with π_q and refer to ι as a $\mathbf{Q}(\pi_q)$-orientation. ☆

Example 5.2.2 More generally, every endomorphism $\alpha \in \mathrm{End}(E) \setminus \mathbf{Z}$ naturally gives rise to an orientation. Indeed, such an endomorphism necessarily satisfies $\alpha^2 - t\alpha + n = 0$ where the trace $t = \mathrm{Tr}(\alpha)$ and the norm $n = N(\alpha)$ (which we recall is equal to the degree of α) satisfy $t^2 - 4n < 0$. Fixing

$$\sigma = \frac{t + \sqrt{t^2 - 4n}}{2} \in \mathbf{C}$$

we obtain an orientation $\iota : \mathbf{Q}(\sigma) \hookrightarrow \mathrm{End}^0(E)$, which is unique if we impose that $\iota(\sigma) = \alpha$. Every orientation arises in this way. ☆

For an order $\mathcal{O} \subseteq K$, we say that a K-orientation $\iota : K \hookrightarrow \mathrm{End}^0(E)$ is an *$\mathcal{O}$-orientation* if $\iota(\mathcal{O}) \subseteq \mathrm{End}(E)$. If moreover $\iota(\mathcal{O}') \not\subseteq \mathrm{End}(E)$ for every strict superorder $\mathcal{O}' \supsetneq \mathcal{O}$ in K, then we say that it concerns a *primitive* $\mathcal{O}$-orientation. Note that any K-orientation ι is a primitive $\mathcal{O}$-orientation for a unique order $\mathcal{O} \subseteq K$, namely for the order $\iota^{-1}(\mathrm{End}(E))$. We call this order the *primitive order* for the K-orientation. Let us also introduce the following weaker notion:

Definition 5.2.3 An $\mathcal{O}$-orientation on an elliptic curve E/k is said to be *locally primitive* at a positive integer m if the index of $\mathcal{O}$ inside the primitive order is coprime to m. △

The following is a convenient sufficient condition for local primitivity:

Lemma 5.2.4 *Let E/k be an elliptic curve, let $\sigma \in \mathrm{End}(E)$ and let m be a positive* *integer such that*

(i) $\operatorname{char}(k) \nmid m$*,*

(ii) $E[\ell, \sigma] \cong \mathbf{Z}/\ell\mathbf{Z}$ *for every prime divisor* $\ell \mid m$*.*

Then the natural $\mathbf{Z}[\sigma]$*-orientation on* E *is locally primitive at* m*. As a partial converse, we have that this orientation is not locally primitive at* m *as soon as* $E[\ell, \sigma] \cong \mathbf{Z}/\ell\mathbf{Z} \times \mathbf{Z}/\ell\mathbf{Z}$ *for some prime divisor* $\ell \mid m$*.*

Proof. If the orientation is not locally primitive at m, then we must have $(\sigma - a)/\ell \in \operatorname{End}(E)$ for a prime divisor $\ell \mid m$ and some $a \in \mathbf{Z}$. Thus σ would act as multiplication-by-a on $E[\ell]$. By assumption (ii) we necessarily have $a = 0$, but then $E[\ell, \sigma] = E[\ell] \cong \mathbf{Z}/\ell\mathbf{Z} \times \mathbf{Z}/\ell\mathbf{Z}$ in view of assumption (i): a contradiction. Conversely, if $E[\ell, \sigma] \cong \mathbf{Z}/\ell\mathbf{Z} \times \mathbf{Z}/\ell\mathbf{Z}$ then by [36, Cor. III.4.11] we know that there exists an $\alpha \in \operatorname{End}(E)$ such that $\alpha \circ [\ell] = \sigma$, so the primitive order must contain σ/ℓ, hence the $\mathbf{Z}[\sigma]$-orientation is not locally primitive at m. □

Example 5.2.5 The Frobenius orientation on an elliptic curve E over a finite field $\mathbf{F}_q$ is also a $\mathbf{Z}[\pi_q]$-orientation. If $E(\mathbf{F}_q)[\ell] \cong \mathbf{Z}/\ell\mathbf{Z}$ for some prime number $\ell \nmid q$, then by Lemma 5.2.4 applied to $\sigma = \pi_q - 1$ this orientation is locally primitive at ℓ. If $E[\ell] \subseteq E(\mathbf{F}_q)$ then it is not. ☆

If $\phi : E \to E'$ is an isogeny and if ι is a K-orientation on E, then we can define an induced K-orientation $\phi_*(\iota)$ on E' by letting

$$\phi_*(\iota)(\alpha) = \frac{1}{\deg(\phi)} \phi \circ \iota(\alpha) \circ \hat{\phi}, \quad \forall \alpha \in K,$$

where $\hat{\phi}$ denotes the dual isogeny of ϕ. Given two K-oriented elliptic curves (E, ι) and (E', ι'), we say that an isogeny $\phi : E \to E'$ is K-*oriented* if $\iota' = \phi_*(\iota)$; in this case, we write $\phi : (E, \iota) \to (E', \iota')$. The dual of a K-oriented isogeny is automatically K-oriented as well. Two K-oriented elliptic curves (E, ι) and (E', ι') are called *isomorphic* if there exists an isomorphism $\phi : E \to E'$ such that $\phi_*(\iota) = \iota'$.

Example 5.2.6 Let E, E' be elliptic curves over $\mathbf{F}_q$ with the same trace of Frobenius, so that they can both be viewed as K-oriented elliptic curves with $K = \mathbf{Q}(\sigma)$ as in (5.1). Then an isogeny $\phi : E \to E'$ is K-oriented if and only if it is $\mathbf{F}_q$-rational. ☆

5.2.2 Class group actions

The set

$$\mathcal{E}\ell\ell_{\overline{k}}^{\text{all}}(\mathcal{O}) = \{\, (E, \iota) \mid E \text{ ell. curve over } \overline{k}, \iota \text{ primitive } \mathcal{O}\text{-orientation on } E \,\} / \cong$$

of primitively $\mathcal{O}$-oriented elliptic curves over $\overline{k}$ up to isomorphism comes equipped with an action by the ideal class group of $\mathcal{O}$, which we denote by $\operatorname{Cl}(\mathcal{O})$. For elliptic curves over $\mathbf{C}$ with complex multiplication, this is a classical result. The case where k is a finite field and the orientation is by Frobenius is treated in [35, 38]. This group action, which we describe below in more detail, is free, but in general not transitive, see e.g. [35, Thm. 4.5] and [27, Prop. 3.3] for some subtleties. To avoid issues arising from the non-transitivity, we define

$$\mathcal{E}\ell\ell_{\overline{k}}(\mathcal{O}) \subseteq \mathcal{E}\ell\ell_{\overline{k}}^{\mathrm{all}}(\mathcal{O})$$

to be an arbitrary but fixed orbit (in practice, where we want to study a secret relation between two primitively $\mathcal{O}$-oriented elliptic curves, it will concern the orbit containing these two curves.)

The action is defined as follows. Let (E, ι) be a primitively $\mathcal{O}$-oriented elliptic curve and let $[\mathfrak{a}] \in \mathrm{Cl}(\mathcal{O})$ be an ideal class, represented by an invertible ideal $\mathfrak{a} \subseteq \mathcal{O}$ of norm coprime to $\max\{1, \mathrm{char}(k)\}$; every ideal class admits such a representative by [14, Cor. 7.17]. One defines the $\mathfrak{a}$-torsion subgroup as

$$E[\mathfrak{a}] = \bigcap_{\alpha \in \mathfrak{a}} \ker(\iota(\alpha)),$$

which turns out to be finite (of order $N(\mathfrak{a}) = \#(\mathcal{O}/\mathfrak{a})$, to be more precise). Thus there exists an elliptic curve E' and a separable isogeny $\phi_{\mathfrak{a}} : E \to E'$ with $\ker(\phi_{\mathfrak{a}}) = E[\mathfrak{a}]$, which is unique up to post-composition with an isomorphism. The isomorphism class of $(E', \phi_{\mathfrak{a}*}(\iota))$ is independent of the choice of the representing ideal $\mathfrak{a}$. One then lets $[\mathfrak{a}](E, \iota)$ be this isomorphism class, and this turns out to define a free group action.

5.2.3 Horizontal, ascending and descending isogenies

Let $\ell \neq \mathrm{char}(k)$ be a prime number and consider an ℓ-isogeny $\phi : (E_1, \iota_1) \to (E_2, \iota_2)$ of K-oriented elliptic curves. Let $\mathcal{O}_1 \subseteq K$ be the primitive order of ι_1 and let $\mathcal{O}_2 \subseteq K$ be the primitive order of ι_2. Then one of the following is true:

- $\mathcal{O}_1 \subseteq \mathcal{O}_2$ and $[\mathcal{O}_2 : \mathcal{O}_1] = \ell$, in which case ϕ is called *ascending*,
- $\mathcal{O}_1 = \mathcal{O}_2$, in which case ϕ is called *horizontal*,
- $\mathcal{O}_2 \subseteq \mathcal{O}_1$ and $[\mathcal{O}_1 : \mathcal{O}_2] = \ell$, in which case ϕ is called *descending*.

It is clear that the dual of an ascending isogeny is descending and vice versa. All horizontal isogenies are of the form $\phi_{\mathfrak{a}}$ for some invertible ideal $\mathfrak{a} \subseteq \mathcal{O}_1 = \mathcal{O}_2$ of norm ℓ, with dual $\phi_{\overline{\mathfrak{a}}}$. Ascending isogenies are of the form $\phi_{\mathfrak{a}}$ for some *non*-invertible ideal $\mathfrak{a} \subseteq \mathcal{O}_1$ of norm ℓ, while descending isogenies are not of the form $\phi_{\mathfrak{a}}$ at all.

5.3 Generalized Weil and Tate pairings

We review some properties of the generalized Weil and Tate pairings on elliptic curves, with a focus on how the latter can be defined in terms of the former. The main sources of inspiration for this section were papers by Bruin [2] and Garefalakis [20], although now we should highlight the work by Robert [31, §4], which appeared near the submission time of the current article and takes this discussion to a deeper level. Nevertheless, while the following statements may be well-known to some experts, we did not succeed in pinpointing exact references for all of them, so we take the opportunity to fill some apparent gaps in the existing literature.

5.3.1 Weil pairing

Following [20] and [36, Ex. III.3.15], to any elliptic curve isogeny $\psi : E \to E'$ over a perfect field k such that $\operatorname{char}(k) \nmid \deg(\psi)$ one can associate the *ψ-Weil pairing*

$$e_\psi : \ker(\psi) \times \ker(\hat{\psi}) \to \overline{k}^* : (P, Q) \mapsto \frac{g \circ \tau_P}{g},$$

where $\hat{\psi} : E' \to E$ denotes the dual of ψ. Here, $g \in k(E)$ is any function with divisor $\psi^*(Q) - \psi^*(0_{E'})$ and τ_P denotes the translation-by-P map. It can be argued that $(g \circ \tau_P)/g$ is indeed constant. The ψ-Weil pairing takes values in μ_m, with m any positive integer such that $\ker(\psi) \subseteq E[m]$. When applied to the multiplication-by-m map on an elliptic curve E one recovers the classical m-Weil pairing, as it is defined in [36, §III.8].

Lemma 5.3.1 *The ψ-Weil pairing is bilinear, non-degenerate,* $\operatorname{Gal}(\overline{k}, k)$*-invariant and further satisfies:*

1. Skew-symmetry: *for any isogeny $\psi : E \to E'$ we have*

$$e_\psi(P, Q) = e_{\hat{\psi}}(Q, P)^{-1} \qquad \text{for all } P \in \ker(\psi),\ Q \in \ker(\hat{\psi}),$$

2. Compatibility Weil-I: *for any chain of isogenies $E \xrightarrow{\phi} E' \xrightarrow{\psi} E''$ we have*

$$\text{(a)} \quad e_{\psi\circ\phi}(P, Q) = e_\psi(\phi(P), Q) \qquad \text{for all } P \in \ker(\psi \circ \phi),\ Q \in \ker(\hat{\psi}),$$

$$\text{(b)} \quad e_{\psi\circ\phi}(P, Q) = e_\phi(P, \hat{\psi}(Q)) \qquad \text{for all } P \in \ker(\phi),\ Q \in \ker(\hat{\phi} \circ \hat{\psi}),$$

3. Compatibility Weil-II: *for any positive integer m and any isogeny $\phi : E \to E'$ we have*

$$e_m(\phi(P), Q) = e_m(P, \hat{\phi}(Q)) \qquad \text{for all } P \in E[m],\ Q \in E'[m].$$

Proof. We refer to [20, §2] and [36, Ex. III.3.15(c)] for bilinearity, non-degeneracy, Galois invariance and Compatibility Weil-I(a). Compatibility Weil-II is just a restatement of [36, III.Prop. 8.2]. Skew-symmetry is well-known in case $\psi = m$. The general case can be found in [31, §4.1], although this can also been seen as a consequence of the case $\psi = m$. Indeed, write $m = \deg(\psi)$ and pick any point $R \in E'$ such that $\hat{\psi}(R) = P$ and likewise pick any point $S \in E$ such that $\psi(S) = Q$. Observe that R, S

are m-torsion points. Then one checks that

$$e_\psi(P,Q) = e_\psi(\hat{\psi}(R),\psi(S)) = e_m(R,\psi(S)) = e_m(\psi(S),R)^{-1} = \\ e_m(S,\hat{\psi}(R))^{-1} = e_{\hat{\psi}}(\psi(S),\hat{\psi}(R))^{-1} = e_{\hat{\psi}}(Q,P)^{-1}$$

as wanted. Here the first and last equality use Compatibility Weil-I(a), the third equality uses skew-symmetry for the classical m-Weil pairing, and the fourth equality uses Compatibility Weil-II. Compatibility Weil-I(b) is an immediate consequence of Compatibility Weil-I(a) and skew-symmetry. □

For $\psi = m$ there is an equivalent definition of the Weil pairing which is more amenable to computation via Miller's algorithm [24].

Lemma 5.3.2 *Let $P, Q \in E[m]$. Choose divisors*

$$D_P \sim (P) - (0_E) \qquad \textit{and} \qquad D_Q \sim (Q) - (0_E)$$

whose supports are disjoint from $\{(Q), (0_E)\}$ and $\{(P), (0_E)\}$, respectively. Let $f_{m,P}, f_{m,Q} \in k(E)$ be such that

$$\operatorname{div}(f_{m,P}) = m(P) - m(0_E), \qquad \operatorname{div}(f_{m,Q}) = m(Q) - m(0_E).$$

Then $e_m(P,Q) = (-1)^m f_{m,P}(D_Q)/f_{m,Q}(D_P)$.

Proof. See e.g. [25]. □

There is no known analogue of this result for the more general ψ-Weil pairing; see [28, §3.6] for a discussion. Note that it is possible to relax the assumption on the supports of D_P, D_Q by working with normalized functions, along the lines of [25, Def. 4].

5.3.2 Tate pairing

The literature describes a number of related pairings on elliptic curves that are all being referred to as the Tate pairing. We focus on the case $k = \mathbf{F}_q$. Following Bruin [2], to any $\mathbf{F}_q$-rational isogeny $\psi : E \to E'$ such that $\ker(\psi) \subseteq E[m] \subseteq E[q-1]$ we associate the ψ-Tate pairing

$$T_\psi : (\ker(\hat{\psi}))(\mathbf{F}_q) \times \frac{E'(\mathbf{F}_q)}{\psi(E(\mathbf{F}_q))} \to \mu_m \subseteq \mathbf{F}_q^*$$

defined by $T_\psi(P,Q) = e_{\hat{\psi}}(P, \pi_q(R) - R)$, where R is arbitrary such that $\psi(R) = Q$. This is sometimes called the *reduced* Tate pairing in order to distinguish it from the Frey–Rück Tate pairing (see below); this terminology is particularly common in case $\psi = m$.

Remark 5.3.3 Bruin instead writes $e_\psi(\pi_q(R)-R,P)$, so in view of the skew-symmetry we appear to have inverted the pairing value; however, this inversion compensates for the fact that Bruin follows a different convention for the Weil pairing [2, §4]. In particular, our two definitions of the ψ-Tate pairing match. ♢

Lemma 5.3.4 *The ψ-Tate pairing is bilinear, non-degenerate,* $\mathrm{Gal}(\overline{\mathbf{F}}_q,\mathbf{F}_q)$*-invariant and moreover satisfies:*

1. Compatibility Tate-I: *for any chain of $\mathbf{F}_q$-rational isogenies $E \xrightarrow{\phi} E' \xrightarrow{\psi} E''$ we have*

$$T_{\psi\circ\phi}(P,Q)=T_\psi(P,Q) \qquad \textit{for all } P\in(\ker(\hat{\psi}))(\mathbf{F}_q)\ ,\ Q\in E''(\mathbf{F}_q),$$

2. Compatibility Tate-II: *for any positive integer m and any $\mathbf{F}_q$-rational isogeny $\phi : E \to E'$ we have*

$$T_m(\phi(P),Q)=T_m(P,\hat{\phi}(Q)) \qquad \textit{for all } P\in E[m](\mathbf{F}_q),\ Q\in E'(\mathbf{F}_q).$$

Proof. For compatibility Tate-I we note that

$$T_{\psi\circ\phi}(P,Q)=e_{\hat{\phi}\circ\hat{\psi}}(P,\pi_q(R)-R)=e_{\hat{\psi}}(P,\pi_q(\phi(R))-\phi(R))$$

for any R such that $\psi(\phi(R))=Q$; here we used Compatibility Weil-I(b) and the fact that ϕ is defined over $\mathbf{F}_q$. But this is indeed equal to $T_\psi(P,Q)$, because $\psi(\phi(R))=Q$. Compatibility Tate-II is an immediate consequence of Compatibility Weil-II. □

Notice that applying Compatibility Tate-I to $E' \xrightarrow{\phi} E \xrightarrow{\psi} E'$, where ϕ is such that $[m]=\psi\circ\phi$ (e.g., $\phi=\hat{\psi}$ in case ψ is cyclic of degree m), shows that

$$T_\psi(P,Q)=T_m(P,Q) \qquad \text{for all } P\in(\ker(\hat{\psi}))(\mathbf{F}_q)\ ,\ Q\in E'(\mathbf{F}_q)$$

from which one sees that the ψ-Tate pairing is just a restriction of the m-Tate pairing. This is in stark contrast with the ψ-Weil pairing, whose relation to the m-Weil pairing is much more convoluted.

The following is an alternative interpretation of the ψ-Tate pairing in terms of the Weil pairing. This generalizes Garefalakis' main observation [20, §5].

Proposition 5.3.5 *Consider an $\mathbf{F}_q$-rational isogeny $\psi : E \to E'$ between elliptic curves over $\mathbf{F}_q$ and assume that*

$$\ker(\psi)\subseteq E[q-1].$$

Then we obtain a well-defined pairing

$$\frac{E'(\mathbf{F}_q)}{\psi(E(\mathbf{F}_q))}\times(\ker(\hat{\psi}))(\mathbf{F}_q)\to\mathbf{F}_q^*$$

from the $(\pi_q - 1)$*-Weil pairing*

$$e_{\pi_q - 1} : E'(\mathbf{F}_q) \times \ker(\hat{\pi}_q - 1) \to \mathbf{F}_q^*$$

on E'*, by restricting the domain of the second argument to* $\ker(\hat{\pi}_q - 1) \cap \ker(\hat{\psi})$*. Moreover,*

$$T_\psi(P, Q) = e_{\pi_q - 1}(Q, P)^{-1}$$

for all $P \in (\ker(\hat{\psi}))(\mathbf{F}_q)$ *and* $Q \in E'(\mathbf{F}_q)$*.*

Proof. We first show that

$$\ker(\hat{\pi}_q - 1) \cap \ker(\hat{\psi}) = \ker(\pi_q - 1) \cap \ker(\hat{\psi}) = (\ker(\hat{\psi}))(\mathbf{F}_q).$$

Indeed, we have $\ker(\hat{\psi}) \subseteq E'[q-1]$ and $\# \ker(\pi_q - 1) = \# \ker(\hat{\pi}_q - 1) = q - t + 1$, with t the trace of Frobenius. From this it follows that

$$\ker(\pi_q - 1) \cap \ker(\hat{\psi}),\ \ker(\hat{\pi}_q - 1) \cap \ker(\hat{\psi}) \ \subseteq\ E'[t-2].$$

Using that $(\hat{\pi}_q - 1) + (\pi_q - 1) = t - 2$, the desired equality follows.

Next, we observe that any point $Q \in (\ker(\hat{\psi}))(\mathbf{F}_q)$ pairs trivially with $\psi(P)$ for any $P \in E(\mathbf{F}_q)$:

$$e_{\pi_q - 1}(\psi(P), Q) = e_{(\pi_q - 1)\circ\psi}(P, Q) = e_{\psi\circ(\pi_q - 1)}(P, Q) = e_{\pi_q - 1}(P, \hat{\psi}(Q)) = 1,$$

where the first three equalities use Compatibility Weil-I(a), the rationality of ψ, and Compatibility Weil-I(b), respectively. So we indeed end up with a pairing whose domain coincides with that of T_ψ, up to reordering the factors.

Finally, to see that both pairings are each other's inverses, take $P \in (\ker(\hat{\psi}))(\mathbf{F}_q)$ and $Q \in E'(\mathbf{F}_q)$. From Compatibility Tate-I we know that

$$T_\psi(P, Q) = T_{\psi\circ(\pi_q - 1)}(P, Q) = e_{(\hat{\pi}_q - 1)\circ\hat{\psi}}(P, (\pi_q - 1)(R)) = e_{\hat{\psi}\circ(\hat{\pi}_q - 1)}(P, (\pi_q - 1)(R))$$

with R such that $\psi \circ (\pi_q - 1)R = Q$. Compatibility Weil-I(b) allows us to rewrite this as

$$e_{\hat{\pi}_q - 1}(P, \psi((\pi_q - 1)(R))) = e_{\hat{\pi}_q - 1}(P, Q)$$

which indeed equals $e_{\pi_q - 1}(Q, P)^{-1}$ by skew-symmetry. □

We will extend this observation to a wider class of pairings in Section 5.5.

Following [18] and [31, §4.4–4.5] one can also consider the *Frey–Rück* ψ-Tate pairing

$$t_\psi : (\ker(\hat{\psi}))(\mathbf{F}_q) \times \frac{E'(\mathbf{F}_q)}{\psi(E(\mathbf{F}_q))} \to \frac{\mathbf{F}_q^*}{(\mathbf{F}_q^*)^m} : (P, Q) \mapsto f_{m,P}(D_Q)$$

with $f_{m,P}$ and D_Q as in Lemma 5.3.2.[1] It allows for an efficient evaluation through

[1] It may seem suspicious, at first sight, that $f_{m,P}(D_Q)$ does not depend on ψ. However, here too,

Miller's algorithm. The Frey-Rück ψ-Tate pairing relates to the reduced ψ-Tate pairing T_m via the rule

$$T_\psi(P, Q) = t_\psi(P, Q)^{(q-1)/m}, \tag{5.2}$$

see [2, §4] and [31, Rmk. 4.14], which is the reason for calling the former reduced. In particular, also T_ψ can be evaluated efficiently.

Remark 5.3.6 It may be tempting to rephrase Lemma 5.3.2 as

$$e_m(P, Q) = t_m(P, Q)/t_m(Q, P),$$

however one should be careful with this: other representatives of $t_m(P, Q)$ and $t_m(Q, P)$ may fail to quotient to $e_m(P, Q)$. See [19, §IX.6] for a discussion. ◊

5.4 Self-pairings

In this section we analyze self-pairings, which we formally define as follows:

Definition 5.4.1 A *self-pairing* on a finite subgroup G of an elliptic curve E/k is a homogeneous function

$$f : G \to \overline{k}^*$$

of degree 2. In other words, for all $P \in G$ and $\lambda \in \mathbf{Z}$ it holds that $f(\lambda P) = f(P)^{\lambda^2}$. △

As the terminology suggests, our primary examples come from the application of a bilinear pairing to a point and itself. More generally, it is natural to consider

$$f : G \to \overline{k}^* : P \mapsto e(\tau_1(P), \tau_2(P)) \tag{5.3}$$

for endomorphisms $\tau_1, \tau_2 \in \mathrm{End}(E)$ (possibly scalar multiplications), with e a bilinear pairing on a group that contains $\tau_1(G) \times \tau_2(G)$.

Example 5.4.2 Let $m \geq 2$ be an integer. The skew-symmetry of the classical Weil pairing implies that $e_m(P, P) = 1$ for any $P \in E[m]$. More generally, the m-Weil pairing becomes trivial whenever it is evaluated at two points belonging to the same cyclic subgroup $\langle P \rangle \subseteq E[m]$:

$$e_m(\tau_1 P, \tau_2 P) = e_m(P, P)^{\tau_1 \tau_2} = 1 \qquad \text{for any } \tau_1, \tau_2 \in \mathbf{Z}.$$

In particular, if one wants to build non-trivial self-pairings from the classical Weil pairing, then this requires the use of at least one non-scalar τ_i. ☆

Example 5.4.3 The following example is inspired by [19, p. 193]. Consider the elliptic curve $E : y^2 = x^3 + 1$ over a finite field $\mathbf{F}_q$ with $q \equiv 1 \bmod 3$. It comes equipped with

the Frey–Rück ψ-Tate pairing is just a restriction of the Frey–Rück m-Tate pairing.

the $\mathbf{F}_q$-rational automorphism $\tau : (x,y) \mapsto (\omega x, y)$, with ω a primitive 3rd root of unity. Let $\ell \mid \#E(\mathbf{F}_q)$ be a prime satisfying $\ell \equiv 2 \bmod 3$. Then the self-pairing

$$E[\ell] \to \mathbf{F}_q^* : P \mapsto e_\ell(P, \tau(P))$$

takes non-trivial values for any $P \neq 0_E$. Indeed, every non-zero $P \in E[\ell]$ is mapped to an independent point because there are no non-trivial eigenvectors for the action of τ on $E[\ell]$: its characteristic polynomial $x^2 + x + 1$ is irreducible mod ℓ. Since τ is defined over $\mathbf{F}_q$, this reasoning also proves that $E[\ell] \subseteq E(\mathbf{F}_q)$. ☆

Example 5.4.4 As a more interesting example, consider an ordinary elliptic curve $E/\mathbf{F}_q$ with endomorphism ring $\mathbf{Z}[\pi_q]$, and assume $m \mid q-1$. The natural reduction map $E(\mathbf{F}_q) \to E(\mathbf{F}_q)/m(E(\mathbf{F}_q))$ allows us to view the reduced m-Tate pairing as a bilinear map

$$T_m : E(\mathbf{F}_q)[m] \times E(\mathbf{F}_q) \to \mu_m. \tag{5.4}$$

By doing so, we may give up on the right non-degeneracy, but the pairing is still left non-degenerate, that is, for any non-trivial point $P \in E(\mathbf{F}_q)[m]$ there exists a point $Q \in E(\mathbf{F}_q)$ such that $T_m(P,Q) \neq 1$. Since $\mathrm{End}(E) = \mathbf{Z}[\pi_q]$, the group $E(\mathbf{F}_q)$ is cyclic (see [22, Thm. 1] or apply Lemma 5.2.4 to $\sigma = \pi_q - 1$). Thus, in this case, we have an induced self-pairing

$$E(\mathbf{F}_q) \to \mu_m : P \mapsto T_m(\tau P, P), \tag{5.5}$$

where τ denotes scalar multiplication by the index $[E(\mathbf{F}_q) : E(\mathbf{F}_q)[m]]$. This self-pairing is non-trivial as soon as $E(\mathbf{F}_q)[m]$ is non-trivial. Note that we can restrict the domain $E(\mathbf{F}_q)$ to its m-primary part $E(\mathbf{F}_q)[m^\infty]$ without affecting this property. ☆

Remark 5.4.5 By the definition of T_m, the image of (5.5) can be rewritten as

$$e_m\left(\tau P, \frac{\pi_q - 1}{m}(P)\right)$$

which seems to be an instance of (5.3) with e the m-Weil pairing. However, note that $(\pi_q - 1)/m$ is *not* an endomorphism of E. On the other hand, it *does* descend (or rather ascend) to an endomorphism when considered on $E/\langle P \rangle$ and this is enough for the pairing to be defined unambiguously. Recall from Proposition 5.3.5 that (5.5) can also be rewritten as $e_{\pi_q - 1}(P, \tau P)^{-1}$. ◊

Our definition of a self-pairing a priori allows for maps that do *not* come from a bilinear pairing. This is indeed possible and, interestingly, a small example has appeared in the literature. Let E be an elliptic curve over a finite field $\mathbf{F}_q$ with $q \equiv 1 \bmod 4$ and $\#E(\mathbf{F}_q) \equiv 2 \bmod 4$. Then the "semi-reduced Tate pairing"

$$E(\mathbf{F}_q)[2] \to \mu_4 : P \mapsto f_{2,P}(D_R)^{\frac{q^2-1}{4}}, \qquad 2R = P \tag{5.6}$$

from [8, Rmk. 11] maps 0_E to 1 and it sends the point of order 2 to a primitive 4-th root of unity. Such an increase of order is impossible for self-pairings coming from a bilinear pairing along the recipe (5.3). Yet it is easy to check that this does concern a self-pairing.

This is essentially the oddest thing that can happen:

Lemma 5.4.6 *Self-pairings map points of order n to* $\gcd(n,2)n$*-th roots of unity.*

Proof. Let $f : G \to \overline{k}^*$ be a self-pairing on an elliptic curve E. Let $P \in G$ have order n. Then from

$$f(P)^{n^2} = f(nP) = f(0_E) = f(0 \cdot 0_E) = f(0_E)^{0^2} = 1$$

and

$$f(P)^{n^2+2n} = \frac{f(P)^{(n+1)^2}}{f(P)} = \frac{f((n+1)P)}{f(P)} = 1$$

it follows that the order of $f(P)$ divides $\gcd(n^2, n^2+2n) = \gcd(n,2)n$. □

Let us now bring isogenies into the picture. Indeed, as discussed in the introduction, self-pairings are only interesting if they are non-trivial and enjoy compatibility with a natural class of isogenies, in the following sense:

Definition 5.4.7 Consider two elliptic curves E, E' over k equipped with respective self-pairings $f : G \to \overline{k}^*$, $f' : G' \to \overline{k}^*$ for finite subgroups $G \subseteq E$, $G' \subseteq E'$. Let $\phi : E \to E'$ be an isogeny. We say that f and f' are *compatible* with ϕ if

$$\phi(G) \subseteq G', \qquad f'(\phi(P)) = f(P)^{\deg(\phi)}$$

for all $P \in G$. △

The most powerful case is where the domains $G = \langle P \rangle$, $G' = \langle P' \rangle$ are cyclic: then we know that $\phi(P) = \lambda P'$ for some $\lambda \in \mathbf{Z}$ and we can conclude

$$f'(P')^{\lambda^2} = f(P)^{\deg(\phi)}, \tag{5.7}$$

leaking information about λ if $\deg(\phi)$ is known and vice versa. We will sometimes refer to self-pairings with cyclic domains as *cyclic self-pairings.* In the non-cyclic case, extracting such information becomes more intricate, although in certain cases it may still be possible; see Remark 5.6.8. We note that the self-pairing from Example 5.4.4 is cyclic, and it follows from Compatibility Tate-II that it is compatible with horizontal $\mathbf{F}_q$-rational isogenies; more specifically (and more generally), if $m \mid q-1$ and E, E' are elliptic curves over $\mathbf{F}_q$ such that the m-primary parts of $E(\mathbf{F}_q)$, $E'(\mathbf{F}_q)$ are cyclic, then the self-pairings

$$E(\mathbf{F}_q)[m^\infty] \to \mu_m : P \mapsto T_m(\tau P, P), \qquad E'(\mathbf{F}_q)[m^\infty] \to \mu_m : P \mapsto T_m(\tau P, P),$$

with $\tau = [E(\mathbf{F}_q) : E(\mathbf{F}_q)[m]] = [E'(\mathbf{F}_q) : E'(\mathbf{F}_q)[m]]$, are compatible with any $\mathbf{F}_q$-

rational isogeny $\phi : E \to E'$.

The focus of the current paper lies, more generally, on non-trivial cyclic self-pairings on $\mathcal{O}$-oriented elliptic curves, for some arbitrary (but fixed) imaginary quadratic order $\mathcal{O}$. If we merely impose compatibility with endomorphisms coming from $\mathcal{O}$, then this already imposes severe restrictions:

Proposition 5.4.8 *Let $\mathcal{O}$ be an imaginary quadratic order with discriminant $\Delta_{\mathcal{O}}$ and let (E, ι) be an $\mathcal{O}$-oriented elliptic curve over k. Assume that there exists a self-pairing*

$$f : C \to \overline{k}^*$$

on some finite cyclic subgroup $C \subseteq E$ which is compatible with endomorphisms in $\iota(\mathcal{O})$. In other words, for every $\sigma \in \mathcal{O}$ and every $P \in C$ we have

$$\iota(\sigma)(P) \in C, \qquad f(\iota(\sigma)(P)) = f(P)^{N(\sigma)}.$$

Write $m = \#\langle f(C) \rangle$. Then

(i) $\mathrm{char}(k) \nmid m$*,*

(ii) $m \mid \Delta_{\mathcal{O}}$*,*

(iii) with r the 2-valuation of $\Delta_{\mathcal{O}}$, we have:

- *if $r = 2$ then $m \mid \Delta_{\mathcal{O}}/2$,*
- *if $r \geq 3$ then $m \mid \Delta_{\mathcal{O}}/4$.*

Remark 5.4.9 Note that the image of a self-pairing is not necessarily a group, which is why we write $\langle f(C) \rangle$ rather than $f(C)$. ◊

Proof. Statement *(i)* follows immediately from the fact that $\overline{k}^*$ contains no elements of order $\mathrm{char}(k)$.

As for *(ii)* and *(iii)*, let P be a generator of C. Then $f(P)$ has order m. For any $\sigma \in \mathcal{O}$ we have that $\iota(\sigma)(P) = \lambda_\sigma P$ for some $\lambda_\sigma \in \mathbf{Z}$, and via

$$f(P)^{N(\sigma)} = f(\iota(\sigma)(P)) = f(\lambda_\sigma P) = f(P)^{\lambda_\sigma^2}$$

we see that $N(\sigma) \equiv \lambda_\sigma^2 \bmod m$. Writing s for the 2-valuation of m, we make a case distinction:

- If $s \leq 1$ then from Lemma 5.4.6 we see that some multiple R of P must have order m. Let σ be such that $\mathcal{O} = \mathbf{Z}[\sigma]$. From

  $$(\sigma - \hat{\sigma})^2 R = (\sigma^2 + \hat{\sigma}^2 - 2N(\sigma))R = (\lambda_\sigma^2 + \lambda_{\hat{\sigma}}^2 - 2N(\sigma))R = (2N(\sigma) - 2N(\sigma))R = 0$$

  it follows that $m \mid \Delta_{\mathcal{O}}$ as wanted.

- If $s \geq 2$ then Lemma 5.4.6 only shows the existence of a point $R \in C$ of order $m/2$ and we obtain the weaker conclusion $m \mid 2\Delta_{\mathcal{O}}$. But at least this implies that $\Delta_{\mathcal{O}}$ is even, so we must have $r \geq 2$. Write $\Delta_{\mathcal{O}} = -2^r n$ and consider elements in $\mathcal{O}$ of the form
$$\sigma = \frac{\sqrt{\Delta_{\mathcal{O}}}}{2} + 2^t a \qquad a, t \in \mathbf{Z}_{\geq 0},$$
so that $N(\sigma) = 2^{r-2}n + 2^{2t}a^2$ has to be a square modulo 2^s for every choice of a, t. We distinguish further:

  - If r is odd, then also $r-2$ is odd and taking $a = 0$ immediately shows that $s \leq r-2$, as wanted.
  - If r is even, then taking $t = (r-2)/2$ yields that $n + a^2$ must be a square modulo 2^{s-r+2} for all a. If $s \geq r$ then this gives a contradiction both in case $n \equiv 1 \bmod 4$ (take $a = 1$) and in case $n \equiv 3 \bmod 4$ (take $a = 0$). So $s \leq r-1$.

  It remains to show that if $r \geq 4$ is even then in fact $s \leq r-2$. But if $s = r-1$ then taking $t = (r-4)/2$ yields that $4n + a^2$ must be a square modulo 8 for all a, which gives a contradiction (take $a = 0$).

□

We will refer to the quantity $m = \#\langle f(C) \rangle$ as the *order* of the self-pairing f. In the next section, we will show, by explicit construction, that the necessary conditions from Proposition 5.4.8 are in fact *sufficient* for the existence of a family of cyclic self-pairings
$$f_{(E,\iota)} : C_{(E,\iota)} \to \overline{k}^*, \qquad (E,\iota) \in \mathcal{E}\ell\ell_{\overline{k}}(\mathcal{O}),$$
all satisfying $\#\langle \mathrm{im}(f_{(E,\iota)}) \rangle = m$ and compatible with horizontal isogenies (the family will also cover many non-primitively $\mathcal{O}$-oriented elliptic curves and non-horizontal isogenies; more on that in Section 5.5).

Remark 5.4.10 One may want to relax the assumptions from Proposition 5.4.8 and impose compatibility with endomorphisms whose norm is coprime to m only. This is good enough for the applications we have in mind, and the semi-reduced Tate pairing from (5.6) shows that this is a strict relaxation. Indeed, we know from [8, Thm. 10] that it is compatible with $\mathbf{F}_q$-rational isogenies of odd degree, but there exist $\mathbf{F}_q$-rational endomorphisms of even degree for which compatibility fails: denoting the pairing by f, we see from
$$f(P) = \zeta_4 \qquad \text{and} \qquad f((\pi_q - 1)P) = f(0_E) = 1$$
that it cannot be compatible with the endomorphism $\pi_q - 1$, since $N(\pi_q - 1) = \#E(\mathbf{F}_q) \equiv 2 \bmod 4$. This concerns a self-pairing of order 4 on a $\mathbf{Z}[\pi_q]$-oriented elliptic curve, so it would not be allowed for by Proposition 5.4.8 because $\Delta_{\mathbf{Z}[\pi_q]} \equiv 4 \bmod 8$. In Appendix 5.8 we will prove a relaxed version of Proposition 5.4.8, and we will also show (in a non-effective fashion) that the above example is part of a larger class of

self-pairings of 2-power order that are compatible with K-oriented isogenies of odd degree only. $\Diamond$

5.5 Constructing non-trivial self-pairings

Let $\mathcal{O}$ be an order in an imaginary quadratic number field K and let $m \mid \Delta_{\mathcal{O}}$ be a divisor satisfying the necessary conditions from Proposition 5.4.8:

- $\mathrm{char}(k) \nmid m$,
- if $4 \mid \Delta_{\mathcal{O}}$ then $m \mid \Delta_{\mathcal{O}}/2$,
- if $8 \mid \Delta_{\mathcal{O}}$ then $m \mid \Delta_{\mathcal{O}}/4$.

We will construct a family of cyclic self-pairings of order m, one for each $(E, \iota) \in \mathcal{E\ell\ell}_{\overline{k}}(\mathcal{O})$, which is compatible with all horizontal isogenies. More generally, the construction will apply to all $\mathcal{O}$-oriented elliptic curves (E, ι) for which the orientation is locally primitive at m, in the sense of Definition 5.2.3. Compatibility will hold for any K-oriented isogeny between two such curves. Our construction is based on a natural generalization of the ψ-Tate pairing to $\mathcal{O}$-oriented elliptic curves, which we discuss first. We will actually only rely on the cases where ψ is a scalar multiplication, but the discussion is fully general for the sake of analogy with the ψ-Tate pairing.

5.5.1 A generalization of the ψ-Tate pairing

Let $m \geq 2$ be any integer that is invertible in k. Consider two $\mathcal{O}$-oriented elliptic curves (E, ι), (E', ι') and let $\psi : E \to E'$ be a K-oriented isogeny between them. Assume that $\ker(\psi) \subseteq E[m]$ and let $\sigma \in \mathcal{O}$ be such that

$$\mathrm{Tr}(\sigma) \equiv 0 \bmod \gcd(m, N(\sigma)). \tag{5.8}$$

We define

$$T_\psi^\sigma : (\ker(\hat{\psi}))[\sigma] \times \frac{E'[\sigma]}{\psi(E[\sigma])} \to \mu_m \subseteq \overline{k}^* : (P, Q) \mapsto e_{\hat{\psi}}(P, \sigma(R))$$

where $R \in E$ is such that $\psi(R) = Q$ and we abusingly write σ instead of $\iota(\sigma), \iota'(\sigma)$. This is well-defined: indeed,

- we have $(\psi \circ \sigma)(R) = (\sigma \circ \psi)(R) = \sigma(Q) = 0_{E'}$, so $\sigma(R) \in \ker(\psi)$,
- making another choice for R amounts to replacing $R \leftarrow R + T$ for some $T \in \ker(\psi)$, and

$$e_{\hat{\psi}}(P, \sigma T) = e_{\hat{\sigma} \circ \hat{\psi}}(P, T) = e_{\hat{\psi} \circ \hat{\sigma}}(P, T) = e_{\hat{\psi}}(\hat{\sigma}(P), T) = e_{\hat{\psi}}((\mathrm{Tr}(\sigma) - \sigma)(P), T) = 1$$

where the first and third equalities use Compatibility Weil-I and the last equality follows from

$$P \in \ker(\hat{\psi}) \cap \ker(\sigma) \subseteq E'[m] \cap E'[N(\sigma)] = E'[\gcd(m, N(\sigma))].$$

The reader should notice the analogy with the definition of the ψ-Tate pairing from Section 5.3. Indeed, applying the above to elliptic curves over $\mathbf{F}_q$ equipped with the natural Frobenius orientation and to $\sigma = \pi_q - 1$, we exactly recover the ψ-Tate pairing; the assumption $m \mid q-1$ that was made there indeed implies (5.8), i.e. $\mathrm{Tr}(\pi_q - 1) \equiv 0 \bmod \gcd(m, N(\pi_q - 1))$.

The pairing T^σ_ψ is bilinear and non-degenerate. Possibly the easiest way to verify this is by noting that the statement and proof of Proposition 5.3.5 carry over: we have

$$T^\sigma_\psi(P, Q) = e_\sigma(Q, P)^{-1}$$

for all $P \in (\ker(\hat{\psi}))[\sigma]$ and $Q \in E'[\sigma]$, so these properties follow from those of the generalized Weil pairing. Our pairing also satisfies the direct analogues of Compatibilities Tate-I and Tate-II:

1. for any chain of K-oriented isogenies $E \xrightarrow{\phi} E' \xrightarrow{\psi} E''$ between $\mathcal{O}$-oriented elliptic curves we have

$$T^\sigma_{\psi\circ\phi}(P, Q) = T^\sigma_\psi(P, Q) \qquad \text{for all } P \in (\ker(\hat{\psi}))[\sigma],\ Q \in E''[\sigma],$$

2. for any positive integer m and any K-oriented isogeny $\phi : E \to E'$ between $\mathcal{O}$-oriented elliptic curves we have

$$T^\sigma_m(\phi(P), Q) = T^\sigma_m(P, \hat{\phi}(Q)) \qquad \text{for all } P \in E[m, \sigma],\ Q \in E'[\sigma].$$

Again the proofs are copies of the corresponding properties of the ψ-Tate pairing.

5.5.2 Self-pairings from divisors of the discriminant

Now consider $m \in \mathbf{Z}_{\geq 2}$ such that $m \mid \Delta_\mathcal{O}$, unless m is even in which case we make the stronger assumptions that $2m \mid \Delta_\mathcal{O}$ in case $4 \mid \Delta_\mathcal{O}$, and $4m \mid \Delta_\mathcal{O}$ in case $8 \mid \Delta_\mathcal{O}$. Furthermore assume that $\mathrm{char}(k) \nmid m$. Pick any generator $\sigma \in \mathcal{O}$ such that

$$m \mid \mathrm{Tr}(\sigma), \tag{5.9}$$

except in the special case where $v_2(m) = 1$, in which case we want

$$2m \mid \mathrm{Tr}(\sigma) \text{ if } 8 \mid \Delta_\mathcal{O}, \qquad m \mid \mathrm{Tr}(\sigma) \text{ but } 2m \nmid \mathrm{Tr}(\sigma) \text{ if } 8 \nmid \Delta_\mathcal{O}. \tag{5.10}$$

Such a generator always exists. Indeed, if m is odd then we can choose whatever generator $\sigma \in \mathcal{O}$ and replace it by $\sigma - (\mathrm{Tr}(\sigma))/2 \bmod m$ if needed. If m is even and $8 \mid \Delta_\mathcal{O}$ then we can just take $\sigma = \sqrt{\Delta_\mathcal{O}}/2$, whose trace is exactly zero. If m is even and $8 \nmid \Delta_\mathcal{O}$ then we can take $\sigma = \sqrt{\Delta_\mathcal{O}}/2 + m/2$, with trace m.

Conditions (5.9–5.10) trivially imply (5.8), so from the foregoing it follows that to any elliptic curve E equipped with an $\mathcal{O}$-orientation we can associate the non-degenerate bilinear pairing

$$T_m^\sigma : E[m,\sigma] \times \frac{E[\sigma]}{m(E[\sigma])} \to \mu_m \subseteq \overline{k}^*,$$

and we know that this family of pairings is compatible with K-oriented isogenies. As with the standard reduced Tate pairing in Example 5.4.4, we can also view T_m^σ as a left non-degenerate bilinear pairing $E[m,\sigma] \times E[m^\infty,\sigma] \to \mu_m$.

Now assume that the orientation is locally primitive at m. Then the group $E[m^\infty,\sigma]$ is cyclic: if it were not cyclic, we would have $E[m'] \subseteq E[m^\infty,\sigma]$ for some positive divisor $m' \mid m$, but this would mean that $\sigma/m' \in \mathrm{End}(E)$, contradicting that σ is a generator of $\mathcal{O}$ and the orientation is locally primitive. Next, note that our assumptions (5.9–5.10) together with

$$\Delta_{\mathcal{O}} = (\mathrm{Tr}(\sigma))^2 - 4N(\sigma)$$

imply that $m \mid N(\sigma)$. Along with the fact that $E[m^\infty,\sigma]$ is cyclic, this in turn yields that $E[m,\sigma]$ is cyclic of order m. By the left non-degeneracy, we see that T_m^σ is surjective onto μ_m and that, again as in Example 5.4.4, it can be converted into a self-pairing

$$f_{(E,\iota)} : E[m^\infty,\sigma] \to \mu_m : P \mapsto T_m^\sigma(\tau P, P)$$

still satisfying $\#\langle \mathrm{im}(f_{(E,\iota)})\rangle = m$; here τ is the index of $E[m,\sigma]$ in $E[m^\infty,\sigma]$. This proves the claims made at the beginning of this section.

5.5.3 Computing the self-pairings

For the practical applications we have in mind, our base field k will be a finite field $\mathbf{F}_q$, and then a compelling question is: what is the complexity of evaluating the self-pairings constructed above? Concretely, for an $\mathcal{O}$-oriented elliptic curve (E,ι) such that both E and $\iota(\mathcal{O})$ are defined over $\mathbf{F}_q$, and a divisor $m \mid \Delta_{\mathcal{O}}$ at which the orientation is locally primitive, how efficiently can we find an appropriate $\sigma \in \mathcal{O}$ and compute

$$T_m^\sigma(\tau P, P) = e_\sigma(P, \tau P)^{-1}$$

with P a generator of $E[m^\infty,\sigma]$ and τ the index of $E[m,\sigma]$ inside $E[m^\infty,\sigma]$? Here, by "appropriate" we mean that σ should satisfy conditions (5.9–5.10), but it is not necessary that σ is a generator of $\mathcal{O}$, as long as the orientation by $\mathbf{Z}[\sigma]$ remains locally primitive at m.

Example 5.5.1 The situation is particularly nice for the Frobenius orientation in case $m \mid q-1$ and $m \mid \#E(\mathbf{F}_q)$. From the identities $\mathrm{Tr}(\pi_q - 1) = (q-1) - \#E(\mathbf{F}_q)$, $N(\pi_q - 1) = \#E(\mathbf{F}_q)$ and $\Delta_{\mathcal{O}} = \mathrm{Tr}(\pi_q - 1)^2 - 4N(\pi_q - 1)$ it is easy to check that m satisfies our necessary conditions for the existence of an order-m self-pairing. Morover, they show that $\sigma = \pi_q - 1$ meets conditions (5.9–5.10). If the orientation by $\mathbf{Z}[\pi_q]$ is locally primitive at m then the resulting order-m self-pairing

$$E(\mathbf{F}_q)[m^\infty] \to \mathbf{F}_q^* : P \mapsto T_m^{\pi_q - 1}(\tau P, P) = T_m(\tau P, P), \qquad \tau = \frac{\#E(\mathbf{F}_q)[m^\infty]}{m}$$

becomes an instance of the reduced m-Tate pairing, so it can be computed via the Frey–Rück Tate pairing t_m as in (5.2). The latter can be evaluated efficiently using Miller's algorithm, in time $O(\log^2 m \log^{1+\varepsilon} q)$ using fast multiplication. ☆

Example 5.5.2 An interesting case is where $\sigma = \varsigma / b$ for some integer $b \geq 2$, where ς is some easier endomorphism. Then it suffices to compute $T_m^{\varsigma}(\tau P, Q)$ for any $Q \in E$ such that $bQ = P$. Indeed:

$$T_m^{\varsigma}(\tau P, Q) = e_m(\tau P, \varsigma(R)) = e_m(\tau P, \frac{\varsigma}{b}(bR)) = T_m^{\sigma}(\tau P, P),$$

with R such that $mR = Q$, so that $m(bR) = P$. E.g., if $\varsigma = \pi_q - 1$, then this again allows us to resort to the Frey–Rück Tate pairing. ☆

Remark 5.5.3 In the previous example the group $E[m^\infty, \varsigma]$, unlike $E[m^\infty, \sigma]$, may not be cyclic. This sheds a new and more conceptual light on the "not walking to the floor" appendix to [8]. There m was taken to be a prime divisor of $q - 1$; for the sake of exposition, let us ignore the technical (and less interesting) case $m = 2$ in what follows. It was assumed that E is an ordinary elliptic curve over $\mathbf{F}_q$ not located on the crater of its m-isogeny volcano, and that

$$E[m^\infty, \pi_q - 1] = E(\mathbf{F}_q)[m^\infty] \cong \frac{\mathbf{Z}}{m^r \mathbf{Z}} \times \frac{\mathbf{Z}}{m^s \mathbf{Z}}$$

for some $r > s + 1$. For us, the weaker assumptions $r > s$ and $m \mid \Delta_{\mathrm{End}(E)}$ will do. One then simply notes that $\sigma := (\pi_q - 1)/m^s \in \mathrm{End}(E)$ and that, when viewing E as a $\mathbf{Z}[\sigma]$-oriented elliptic curve, the orientation becomes locally primitive at m. By the assumption on $\Delta_{\mathrm{End}(E)}$ we still have

$$m \mid \Delta_{\mathbf{Z}[\sigma]} \qquad \text{and consequently} \qquad \mathrm{Tr}(\sigma) \equiv 0 \bmod m,$$

where the last congruence uses $\Delta_{\mathbf{Z}[\sigma]} = \mathrm{Tr}(\sigma)^2 - 4N(\sigma) = \mathrm{Tr}(\sigma)^2 - 4 \cdot \#E(\mathbf{F}_q)/m^{2s}$. Thus we have a self-pairing

$$E[m^\infty, (\pi_q - 1)/m^s] \to \mu_m : P \mapsto T_m^{(\pi_q - 1)/m^s}(m^{r-s-1}P, P)$$

of order m, with cyclic domain $E[m^\infty, (\pi_q - 1)/m^s] \cong \mathbf{Z}/m^{r-s}\mathbf{Z}$. When computing this self-pairing via the standard m-Tate pairing as in Example 5.5.2, using $\varsigma = \pi_q - 1$ and $b = m^s$, we recover the pairing discussed in [8, App. A]. ◊

Unfortunately, for general σ we do not know of an analogue of the Frey–Rück Tate pairing, nor of an analogue of Lemma 5.3.2 for the generalized Weil pairing. The best methods we can currently think of work by embedding the pairing into a standard Weil pairing, that is, with respect to scalar multiplication. In this way Miller's algorithm becomes available. The embedding is natural via the definition:

$$T_m^\sigma(\tau P, P) = e_m(\tau P, \sigma(R))$$

with $R \in E$ such that $mR = P$. Alternatively, using compatibility Weil-I one can rewrite

$$e_\sigma(P, \tau P)^{-1} = e_{N(\sigma)}(P, \tau R)^{-1}$$

with $R \in E$ a preimage of P under σ. Since m is typically a lot smaller than $N(\sigma)$, and since evaluating σ seems easier than computing a preimage, the first method appears to be preferable in practice.

The complexity then depends heavily on the field of definition of the points in $E[m^\infty, \sigma]$. In the worst case, one may need to unveil the full $N(\sigma)$-torsion to see these points, requiring to switch to $\mathbf{F}_{q^a}$ with a the order of π_q acting on $E[N(\sigma)]$, which is $O(N(\sigma)^2)$. We must also divide P by m to get R, for which we may need to extend further to

$$\mathbb{F}_{q^{aa'}} \quad \text{with } a' = O(m^2).$$

Running Miller's algorithm for the m-Weil pairing over $\mathbf{F}_{q^{aa'}}$ could then cost an atrocious

$$O(\Delta_\mathcal{O}^{2+\varepsilon} m^{2+\varepsilon} \log^{1+\varepsilon} q),$$

where we have approximated $N(\sigma) \approx \Delta_\mathcal{O}$.

However, this is the absolute worst case: one typically expects $E[m^\infty, \sigma] \subseteq E[m^t]$ for some very small constant t, most likely $t = 1$, and then the estimate becomes

$$O(m^{2t+2+\varepsilon} \log^{1+\varepsilon} q).$$

E.g., in Proposition 5.6.5 this will be applied to moduli m of sub-exponential size, leading to a sub-exponential workload. We note that the above estimates ignore the cost of determining $\iota(\sigma)$ and evaluating it on R. This heavily depends on how the orientation is given in practice, which is a separate discussion for which we refer to [39].

5.6 Applications

In this section, we present two applications of the non-trivial self-pairings from Section 5.5. In Section 5.6.1, we show how knowledge of the degree of a secret isogeny together with a non-trivial self-pairing on a large enough subgroup allows us to efficiently attack certain instances of class group action based cryptography. In Section 5.6.2, we use the generalized view of self-pairings to conceptualize previous results on the decisional Diffie–Hellman problem for class group actions [8, 6].

5.6.1 Easy instances of class group action inversion

Using the tools developed in the previous sections, we describe a special family of class group actions on oriented elliptic curves for which the vectorization problem is easy, i.e., the class group action can be efficiently inverted. More precisely, we give a high-level recipe for recovering a secret horizontal isogeny ϕ between two primitively $\mathcal{O}$-oriented elliptic curves (E, ι), (E', ι') whenever $d = \deg(\phi)$ is known and smaller than m^2, where m is a prime power satisfying

$$m^2 \mid \Delta_{\mathcal{O}} \text{ if } m \text{ is odd,} \qquad 4m^2 \mid \Delta_{\mathcal{O}} \text{ if } m \text{ is even.}$$

It is also assumed that $\gcd(m, \mathrm{char}(k), d) = 1$. While it has been previously pointed out that factors dividing the discriminant can cause a decrease of security, see e.g. [3, Rmk. 2] or [8, §5.1], it was unknown that in special cases they allow for a full break of the vectorization problem.

Attack strategy.

Let $\sigma \in \mathcal{O}$ be such that $\mathrm{Tr}(\sigma) \equiv 0 \bmod m^2$ and the orientation by $\mathbf{Z}[\sigma]$ is locally primitive at m. As discussed in Section 5.5.2 such a σ exists and is easy to find; we can even choose σ to be a generator of $\mathcal{O}$, but in certain cases one may want to take a non-generator for reasons of efficiency.[2]

Recall, again from Section 5.5.2, that the groups $E[m^\infty, \sigma]$ and $E'[m^\infty, \sigma]$ are cyclic and we obtain self-pairings

$$f : E[m^\infty, \sigma] \to \mu_{m^2} \qquad \text{and} \qquad f' : E'[m^\infty, \sigma] \to \mu_{m^2}$$

of order m^2 by mapping $P \mapsto T^{\sigma}_{m^2}(\tau P, P)$, where

$$\tau = [E[m^\infty, \sigma] : E[m^2, \sigma]] = [E'[m^\infty, \sigma] : E'[m^2, \sigma]].$$

Now, pick respective generators P, P' of $E[m^\infty, \sigma]$, $E'[m^\infty, \sigma]$. Because ϕ is K-oriented and its degree is coprime to m, we know that $P' = \mu\phi(P)$ for some unit $\mu \in \mathbf{Z}/m^2\mathbf{Z}$. The compatibility of f and f' with K-oriented isogenies then implies

$$f'(P') = f(P)^{d\mu^2}.$$

Knowing d, we can determine $\mu^2 \bmod m^2$ using a discrete logarithm computation in μ_{m^2}, which leaves at most four options for $\mu \bmod m^2$: two options if m is odd and four options if m is a power of 2. Given a correct guess for $\mu \bmod m^2$, we obtain knowledge of pair of points

$$Q = \mu\tau P \qquad \text{and} \qquad Q' = \tau P'$$

of order m^2 that are connected via ϕ.

[2] For instance, to allow for σ of the form $(\pi_q - 1)/b$ as in Example 5.5.2.

Remark 5.6.1 Guessing $-\mu$ is in fact equally fine, because it is of course good enough to recover $-\phi = [-1] \circ \phi$. Therefore, only in the case where m is a power of 2 there is an actual need for guessing between $\pm\mu$ and $\pm(1 + m^2/2)\mu$, where we have to repeat the procedure below in case of a wrong guess. ♢

Using a reduction by De Feo et al.,[3] the problem of recovering ϕ given its images on the cyclic subgroup $\langle Q \rangle$ of order m^2 can be reduced to the problem of recovering a related degree-d isogeny $\phi_0 : E_0 \to E_0'$ given its images on $E_0[m]$. The idea is to compute the isogenies $\psi : E \to E_0$, $\psi' : E' \to E_0'$ with kernels generated by mQ and $m\phi(Q)$, respectively, and complete the diagram:

$$\begin{tikzcd}
E_0 \arrow[r, "\phi_0"] & E_0' \\
E \arrow[u, "\psi"] \arrow[r, "\phi"] & E' \arrow[u, "\psi'"']
\end{tikzcd}$$

The points $Q_0 := \psi(Q)$ and $Q_0' := \psi'(Q') = \psi'(\phi(Q))$ are of order m and we have $\phi_0(Q_0) = Q_0'$. Further, by picking any generator R_0 of $\ker(\hat{\psi})$ we obtain a basis $\{Q_0, R_0\}$ of $E_0[m]$. If we choose a generator R_0' of $\ker(\hat{\psi}')$ then it is easy to argue that $R_0' = \lambda\phi_0(R_0)$ for some $\lambda \in \mathbf{Z}$ that is coprime to m. The exact value of $\lambda \bmod m$ can be recovered via a discrete logarithm computation by comparing

$$e_m(Q_0', R_0') = e_m(\phi_0(Q_0), \lambda\phi_0(R_0)) = e_m(Q_0, R_0)^{\lambda d} \quad \text{with} \quad e_m(Q_0, R_0),$$

hence we can assume that $\lambda = 1$. Thus, we are given the images of ϕ_0 on a basis of $E_0[m]$. Since $m^2 > d$, we can use Robert's method from [30, §2], together with the refinement discussed in [30, §6.4], to evaluate ϕ_0 on arbitrary inputs. In particular, we can evaluate ϕ_0 on a basis of $E_0[d]$ in order to determine the kernel of ϕ_0 explicitly; this kernel can then be pushed through $\hat{\psi}$ to obtain the kernel of ϕ.

Remark 5.6.2 In our main use cases, namely attacking special instances of CRS, rather than evaluating ϕ_0 on a basis of $E_0[d]$ (which may be defined over a huge field extension only) we want to proceed as follows. For simplicity, let us focus on the dummy-free set-up with $e = 1$ (see Section 5.1). Then we have $d = \ell_1\ell_2 \cdots \ell_r$ for distinct small primes ℓ_i that split in $\mathcal{O}$. In this context, recovering ϕ amounts to finding for each $i = 1, 2, \ldots, r$ the prime ideal $\mathfrak{l}_i$ above ℓ_i (one out of two options) for which $E[\mathfrak{l}_i]$ is annihilated by ϕ. Then ϕ is the isogeny corresponding to the invertible ideal $\mathfrak{l}_1\mathfrak{l}_2 \cdots \mathfrak{l}_r \subseteq \mathcal{O}$. Since $\gcd(m, d) = 1$ this can be tested directly on E_0 by evaluating ϕ_0 in a generator of $\psi(E[\mathfrak{l}_i])$. ♢

[3] The reduction was presented at the KU Leuven isogeny days in 2022 and an article about this is in preparation [17].

Weak instances over $\mathbf{F}_q$.

Whether or not the above strategy turns into an efficient algorithm depends amongst others on the field arithmetic involved, the cost of evaluating $\iota(\sigma)$, $\iota'(\sigma)$, and the cost of computing discrete logarithms in μ_{m^2}. The following proposition gives instances where it indeed leads to a polynomial-time attack:

Proposition 5.6.3 *Let E, E' be elliptic curves defined over a finite field $\mathbf{F}_q$, equipped with their Frobenius orientations and connected by an unknown horizontal isogeny ϕ of known degree d, assumed B-powersmooth and coprime to q. Let $\mathcal{O} \subseteq \mathbf{Q}(\pi_q)$ be their joint primitive order. Assume that there exists a prime power $m = \ell^r$ satisfying $\ell \leq B$, $\ell \nmid qd$, $\ell^{2r} > d$, and*

$$\ell^{2r} \mid \Delta_{\mathcal{O}} \text{ if } \ell \text{ is odd}, \qquad \ell^{2r+2} \mid \Delta_{\mathcal{O}} \text{ if } \ell = 2.$$

Further, assume that there exists a positive integer b coprime to q such that $\sigma = (\pi_q - 1)/b \in \mathcal{O}$, $\mathrm{Tr}(\sigma) \equiv 0 \bmod \ell^{2r}$ and $\ell \nmid [\mathcal{O} : \mathbf{Z}[\sigma]]$. Then the invertible ideal $\mathfrak{a} \subseteq \mathcal{O}$ for which $\phi = \phi_{\mathfrak{a}}$ can be computed in time $\mathrm{poly}(\log q, B)$.

Proof. First note that

$$d = O(m^2) = O(|\Delta_{\mathcal{O}}|) \quad \text{and} \quad |\Delta_{\mathcal{O}}| = (4q - \mathrm{Tr}(\pi_q)^2)/[\mathcal{O} : \mathbf{Z}[\pi_q]]^2 = O(q)$$

so any subroutine which runs in time $\mathrm{poly}(d, m)$ also runs in time $\mathrm{poly}(q)$. The orientation by $\mathbf{Z}[\sigma]$ being locally primitive at ℓ, we know that

$$E(\mathbf{F}_q) \cong E'(\mathbf{F}_q) \cong \frac{\mathbf{Z}}{bb'\mathbf{Z}} \times \frac{\mathbf{Z}}{bb'c\mathbf{Z}}$$

for positive integers b', c, where $\ell \nmid b'$, that can be determined in time $\mathrm{poly}(\log q)$ using a point-counting algorithm [34]. Define $\kappa = \gcd(\ell^\infty, c)$, where we note that our assumptions imply that $\ell^{2r} \mid \kappa$: indeed recall from Section 5.5.2 that $E[\ell^{2r}, \sigma] \subseteq E[\sigma] \cong \mathbf{Z}/b'\mathbf{Z} \times \mathbf{Z}/b'c\mathbf{Z}$ has order ℓ^{2r}. A generator $P \in E[\ell^\infty, \sigma]$ is found by repeatedly sampling $X \leftarrow E(\mathbf{F}_q)$ until $P = \frac{bb'c}{\kappa} X$ has order κ. Following Example 5.5.2, the self-pairing

$$f(P) = T^{\sigma}_{\ell^{2r}}(\tau P, P) = T^{\frac{\pi_q - 1}{b}}_{\ell^{2r}}(\tau P, P) = T_{\ell^{2r}}(\tau P, \frac{b'c}{\kappa} X), \qquad \tau = \frac{\kappa}{\ell^{2r}}$$

can then be computed in time $\mathrm{poly}(\log q)$ via the Frey–Rück Tate pairing. Likewise, we can efficiently evaluate f' at a generator $P' \in E'[\ell^\infty, \sigma]$, necessarily satisfying $P' = \mu\phi(P)$ for some μ. As outlined above, via a discrete logarithm computation in $\mu_{\ell^{2r}}$, which can be done in time $\mathrm{poly}(\log q, B)$, we obtain $\mu^2 \bmod \ell^{2r}$. Assuming a correct guess for μ, from this we obtain our order-ℓ^{2r} points Q, $Q' = \phi(Q)$ and we are all set for the torsion-point attack. Note that the points Q, Q' are defined over $\mathbf{F}_q$, hence so are the curves E_0, E_0' and evaluating ϕ_0 at a point in $E_0(\mathbf{F}_{q^a})$ only involves arithmetic over $\mathbf{F}_{q^a}$. We then proceed as outlined in Remark 5.6.2, with the difference that d need not be square-free: we only require it to be powersmooth. This means that for each prime power $\ell_i^{e_i}$ dividing d, we have to test up to $2^{e_i} - 1 = O(B)$ ideals of norm $\ell_i^{e_i}$ for annihilation by ϕ_0. All arithmetic can be done in an extension of degree $a = \mathrm{poly}(B)$, from which the proposition follows. □

Example 5.6.4 An example application of Proposition 5.6.3 is where $\ell^{2r} \mid q - 1$ for a small prime ℓ and $r \geq 1$ and $E(\mathbf{F}_q)[\ell^\infty]$ is cyclic of order at least ℓ^{2r}. Then $m := \ell^r$ and $\sigma := \pi_q - 1$ meet the above requirements. Indeed:

- the orientation by $\mathbf{Z}[\pi_q - 1]$ is locally primitive at ℓ by Lemma 5.2.4,
- $\mathrm{Tr}(\pi_q - 1) = q - 1 - \#E(\mathbf{F}_q) \equiv 0 \bmod \ell^{2r}$,
- $\Delta_{\mathbf{Z}[\pi_q - 1]} = \mathrm{Tr}(\pi_q - 1)^2 - 4\#E(\mathbf{F}_q)$ is divisible by ℓ^{2r}, and by ℓ^{2r+2} if $\ell = 2$.

Here is a baby example with $\ell = 2$. Let E be the ordinary elliptic curve defined by

$$y^2 = x^3 + 106960359001385152381x + 100704579394236675333$$

over $\mathbf{F}_p$ with $p := 2^{30} \cdot 167133741769 + 1$. So here we take $\sigma := \pi_p - 1$ and $m := 2^{15}$. One checks that $E[\sigma] = E(\mathbf{F}_p)$ is a cyclic group of order

$$2^{30} \cdot 5^2 \cdot 7 \cdot 11 \cdot 13 \cdot 17 \cdot 19 \cdot 23 \cdot 29 \cdot 31,$$

in particular its subgroup $E(\mathbf{F}_p)[2^\infty]$ is cyclic of order 2^{30} as wanted. In this case it is easy to check that the $\mathbf{Z}[\sigma]$-orientation is primitive overall, i.e., not just locally at 2. This is a minimal example for a curve one would construct for a SiGamal-type encryption scheme [26] using the group action underlying CRS instead of the CSIDH group action; see below. By Proposition 5.6.3, one can recover horizontal isogenies of known powersmooth degree $d < 2^{30}$. We implemented the attack in the Magma computer algebra system [1],[4] only skipping the final step, i.e. computing the actual evaluation algorithm as described in [30]. ☆

A generalization.

The above recipe can be generalized to the case where multiple squared prime powers $m_1^2, \ldots, m_r^2$ divide $\Delta_{\mathcal{O}}$ and the degree d of our secret isogeny ϕ is known and smaller than $m_1^2 \cdots m_r^2$. This time we use a cyclic self-pairing of order $m_1^2 \cdots m_r^2$ to recover $\mu^2 \bmod m_1^2 \cdots m_r^2$, with μ as before. Thus, we have 2^r or 2^{r+1} options for μ depending on whether one of the m_i is even (or in fact 2^{r-1} or 2^r options in case we do not care about a global sign). The rest of the recipe follows mutatis mutandis.

Proposition 5.6.5 (informal) *Let E, E' be elliptic curves defined over a finite field $\mathbf{F}_q$, equipped with their Frobenius orientations and connected by an unknown horizontal isogeny ϕ of known degree d, assumed B-powersmooth and coprime to q. Let $\mathcal{O} \subseteq$*

[4] See `https://github.com/KULeuven-COSIC/Weak-Class-Group-Actions` for the code.

$\mathbf{Q}(\pi_q)$ be their joint primitive order. Assume that there exist $r \approx \sqrt{\log q}$ prime powers $m_1, \ldots, m_r \in L_q(1/2)$ coprime to qd such that $m_1^2 \cdots m_r^2 > d$ and

$$m_1^2 \cdots m_r^2 \mid \Delta_{\mathcal{O}} \qquad \textit{and} \qquad 4m_1^2 \cdots m_r^2 \mid \Delta_{\mathcal{O}} \textit{ if some } m_i \textit{ is even.}$$

Then it is expected that the invertible ideal $\mathfrak{a} \subseteq \mathcal{O}$ for which $\phi = \phi_{\mathfrak{a}}$ can be computed in time $\operatorname{poly}(B) \cdot L_q(1/2)$.

Proof sketch. Let $\sigma \in \mathcal{O}$ be such that $\operatorname{Tr}(\sigma) \equiv 0 \bmod m_1^2 \cdots m_r^2$ and the orientation by $\mathbf{Z}[\sigma]$ is locally primitive at $m_1 \cdots m_r$. If it so happens that $\sigma = (\pi_q - 1)/b$ for some b coprime to q then we can just mimic the previous proof: the main difference is that, this time, there are about $2^r \approx 2^{\sqrt{\log q}} = L_q(1/2)$ possible guesses for the secret scalar μ, from which the stated runtime follows.

In general however, it may not be possible to pick σ of the said form, and then the domains $E[(m_1 \cdots m_r)^\infty, \sigma]$ and $E'[(m_1 \cdots m_r)^\infty, \sigma]$ of our self-pairings may be defined over a field extension of degree $L_q(1)$ only, in which case there is no hope for a sub-exponential runtime. For this reason, the attack should be broken up in pieces. Writing $m_1^{t_1} \cdots m_r^{t_r}$ for the order of $E[(m_1 \cdots m_r)^\infty, \sigma] \cong E'[(m_1 \cdots m_r)^\infty, \sigma]$, as discussed in Section 5.5.3 we heuristically expect that $t_i = O(1)$ for all $i = 1, \ldots, r$. If this is indeed the case, then for each i we can find generators $P_i \in E[m_i^\infty, \sigma]$, $P_i' \in E'[m_i^\infty, \sigma]$ over an extension of degree $L_q(1/2)$. The cyclic self-pairings

$$T^{\sigma}_{m_i^2}(\tau P, P) \quad \text{and} \quad T^{\sigma}_{m_i^2}(\tau P', P'), \qquad \tau = m_i^{t_i - 2}$$

can thus be computed in time $L_q(1/2)$ and this also accounts for the subsequent discrete logarithm computation. Assuming a correct guess for the scalar μ_i such that $P_i' = \mu_i \phi(P_i)$, we obtain a pair of order-m_i^2 points Q_i, $Q_i' = \phi(Q_i)$. Note that, while these points are defined over an extension of degree $L_q(1/2)$, the groups they generate are $\mathbf{F}_q$-rational because our orientation is by Frobenius. In particular, the isogenies ψ_1, ψ_1' and codomains $E_{0,1}$, $E_{0,1}'$ corresponding to Q_1, Q_1' are defined over $\mathbf{F}_q$. The idea is now to push the points Q_2, Q_2' through ψ_1, ψ_1' and repeat the argument, leading to a diagram

$$\begin{tikzcd}
E_{0,r} \arrow[r, "\phi_0"] & E'_{0,r} \\
\vdots \arrow[u, "\psi_r"] & \vdots \arrow[u, "\psi'_r"'] \\
E \arrow[u, "\psi_1"] \arrow[r, "\phi"] & E' \arrow[u, "\psi'_1"']
\end{tikzcd}$$

The map ϕ_0 on top comes equipped with its images on a basis of $E_{0,r}[m_i]$ for each $i = 1, \ldots, r$. For the evaluation of ϕ_0 on arbitrary inputs, we can then proceed as in [29, Prop. 2.9] and conclude as before. □

Unaffected schemes.

From the above propositions it follows that a CRS-instantiation using curves whose discriminants are divisible by (large) powers of smallish primes may be vulnerable to a sub-exponential attack. In particular, from a security point of view, walking down the volcano to instantiate CRS is worse than CRS close to the crater. Each descending step on the ℓ-volcano adds a factor ℓ^2 to our discriminant and thus we can recover isogenies of degree ℓ^2 times larger than a level above, using the attack outlined in this section. We examine how some proposed constructions avoid this problem already.

Schemes that use the maximal order as their orientation are not vulnerable to our attack. We need that a prime power, not a prime, divides the discriminant, because the De Feo et al. reduction works only for points of square order. The maximal order has a discriminant that is square-free, at worst after dividing by 4, so the above does not apply. The CSIDH variant CSURF is an example of a scheme that uses the maximal order [3], where the discriminant is not merely square-free but even prime. Similarly, in the original CSIDH proposal the discriminant is four times a large prime and thus there is no factor of the discriminant large enough to enable our attack.

Schemes that are close to the crater are also secure. For instance, the SCALLOP scheme [16] uses curves one level underneath the crater in the f-volcano, where f is a large prime. Thus the discriminant is of the form $f^2 \cdot d$, where d is square-free away from 4. Theoretically, we can still use a point of order f^2 to recover an isogeny of degree at most f^2. However, to actually see the f-torsion we would need to pass to an extension of degree $O(f)$, which is infeasible for large enough f.

Another scheme worth mentioning is the higher-degree supersingular group actions [11]. Here the order used is $\mathbf{Z}[\sqrt{-dp}]$ for some square-free d, which has discriminant $-dp$ or $-4dp$. Even if d was a square, d is chosen small relative to p, and as such applying the attack above to these orientations, we could recover an isogeny of degree $2d$ at best.

Pairing-based attack strategy on SiGamal.

We end by commenting on a strategy, proposed to us by Luca De Feo and involving self-pairings, to break the IND-CPA security of the SiGamal public-key encryption scheme [26]. In SiGamal, the hardness of the IND-CPA game – i.e., given the encryption of one out of two known plaintexts, guessing which one has been encrypted – relies [26, Thm. 8] on an *ad hoc* assumption called the *P-CSSDDH assumption.*

More precisely, let p be a prime of the form $2^r\ell_1 \cdots \ell_n - 1$, where $r \geq 2$ and $\ell_1, \ldots, \ell_n$ are distinct odd primes. Moreover, let E_0 be the supersingular elliptic curve over $\mathbf{F}_p$ of equation $y^2 = x^3 + x$, P_0 a random generator of $E_0(\mathbf{F}_p)[2^r]$ and $\mathfrak{a}, \mathfrak{b}$ random elements of odd norm in $\mathrm{Cl}(\mathbf{Z}[\pi_p])$. Then the P-CSSDDH assumption is as follows: given the curves $E_0, [\mathfrak{a}]E_0, [\mathfrak{b}]E_0, [\mathfrak{ab}]E_0$ and the points $P_0, P_1 = \phi_{\mathfrak{a}}(P_0)$ and $P_2 = \phi_{\mathfrak{b}}(P_0)$, no efficient algorithm can distinguish $P_3 = \phi_{\mathfrak{ab}}(P_0)$ from a uniformly random 2^r-torsion point $P_3' \in [\mathfrak{a}][\mathfrak{b}]E_0(\mathbb{F}_p)$. Schematically:

$$\begin{tikzcd}
(E_0, P_0) \arrow[r, "\mathfrak{a}"] \arrow[d, "\mathfrak{b}"'] & ([\mathfrak{a}]E_0, P_1 = \phi_{\mathfrak{a}}(P_0)) \arrow[d, "\mathfrak{b}"] \\
([\mathfrak{b}]E_0, P_2 = \phi_{\mathfrak{b}}(P_0)) \arrow[r, "\mathfrak{a}"'] & ([\mathfrak{a}\mathfrak{b}]E_0, P_3 = \phi_{\mathfrak{a}\mathfrak{b}}(P_0), P_3')
\end{tikzcd}$$

If there existed (efficiently computable) non-trivial self-pairings f_i on the subgroups $\langle P_i \rangle$, say of order 2^s, compatible with $\mathbf{F}_p$-rational isogenies of odd degree, then

$$\begin{aligned}
f_1(P_1) &= f_1(\phi_{\mathfrak{a}}(P_0)) = f_0(P_0)^{N(\mathfrak{a})} \\
f_2(P_2) &= f_2(\phi_{\mathfrak{b}}(P_0)) = f_0(P_0)^{N(\mathfrak{b})} \\
f_3(P_3) &= f_3(\phi_{\mathfrak{a}\mathfrak{b}}(P_0)) = f_0(P_0)^{N(\mathfrak{a})N(\mathfrak{b})}.
\end{aligned}$$

Thus, the P-CSSDDH challenge could then be reduced to a decisional Diffie–Hellman problem on μ_{2^s}. However, the existence of such self-pairings f_i is ruled out by Propositions 5.4.8 and 5.8.1. Since $\Delta_{\mathcal{O}} = -4p$ and $p \equiv 3 \bmod 4$ by construction, we are condemned to $s = 2$. This is of no use since $\mathfrak{a}$ and $\mathfrak{b}$ are assumed to have odd norm.

5.6.2 Decisional Diffie–Hellman revisited

Genus theory [14, Ch. I§3B] attaches to every imaginary quadratic order $\mathcal{O}$ a list of *assigned characters*, which form a set of generators for the group of quadratic characters $\chi : \mathrm{Cl}(\mathcal{O}) \to \{\pm 1\}$. In detail: if

$$\Delta_{\mathcal{O}} = -2^r m_1^{r_1} m_2^{r_2} \cdots m_n^{r_n}$$

denotes the factorization of $\Delta_{\mathcal{O}}$ into prime powers, then the assigned characters include

$$\chi_{m_i} : [\mathfrak{a}] \mapsto \left(\frac{N(\mathfrak{a})}{m_i}\right), \qquad i = 1, \ldots, n, \tag{5.11}$$

and this list is extended with a subset of

$$\delta : [\mathfrak{a}] \mapsto \left(\frac{-1}{N(\mathfrak{a})}\right), \qquad \epsilon : [\mathfrak{a}] \mapsto \left(\frac{2}{N(\mathfrak{a})}\right), \qquad \delta\epsilon : [\mathfrak{a}] \mapsto \left(\frac{-2}{N(\mathfrak{a})}\right).$$

Concretely, the character δ is included if $r = 2$ and $-\Delta_{\mathcal{O}}/4 \equiv 1 \bmod 4$, or if $r \geq 4$. The character ϵ is included if $r = 3$ and $-\Delta_{\mathcal{O}}/8 \equiv 3 \bmod 4$, or if $r \geq 5$. The character $\delta\epsilon$ is included if $r = 3$ and $-\Delta_{\mathcal{O}}/8 \equiv 1 \bmod 4$, or if $r \geq 5$. In all this, $\left(\frac{\cdot}{\cdot}\right)$ denotes the Legendre/Jacobi symbol and it is assumed that $[\mathfrak{a}]$ is represented by an invertible ideal $\mathfrak{a} \subseteq \mathcal{O}$ of norm coprime with $\Delta_{\mathcal{O}}$.

In the context of breaking the decisional Diffie–Hellman problem for ideal class group actions, it was observed in [8, 6] that, given two primitively $\mathcal{O}$-oriented elliptic curves

$$(E,\iota),\ (E',\iota') = [\mathfrak{a}](E,\iota)\ \in \mathcal{E}\ell\ell_{\overline{k}}(\mathcal{O})$$

that are connected by an unknown ideal class $[\mathfrak{a}]$, it is possible to compute $\chi([\mathfrak{a}])$ for any assigned character χ, purely from the knowledge of (E,ι), (E',ι'), and at the cost of essentially one discrete logarithm computation (e.g., in the group μ_m in case $\chi = \chi_m$ for an odd prime divisor $m \mid \Delta_{\mathcal{O}}$).

Even though we have not much to add over [8, 6] in terms of efficiency or generality, in this section we want to make the nearly obvious remark that cyclic self-pairings are excellently suited for accomplishing this task. Indeed, if m is an odd prime divisor of $\Delta_{\mathcal{O}}$, then we can consider the cyclic self-pairings

$$f : C \to \mu_m \subseteq \overline{k}^*, \qquad f' : C' \to \mu_m \subseteq \overline{k}^*$$

of order m from Section 5.5. Taking any generators $P \in C$, $P' \in C'$, we know that $P' = \lambda\phi_{\mathfrak{a}}(P)$ for some $\lambda \in \mathbf{Z}$ that is invertible mod m and then

$$f'(P') = f(P)^{\lambda^2 N(\mathfrak{a})} \qquad \text{so that} \qquad \chi_m([\mathfrak{a}]) = \left(\frac{\log_{f(P)} f'(P')}{m}\right).$$

None of the methods from [8, 6] are literal applications of this simple strategy. Indeed, in the case of [8], which focuses on ordinary elliptic curves over finite fields, the self-pairing step is preceded by a walk to the floor of the m-isogeny volcano truncated at $\mathbf{Z}[\pi_q]$, in order to ensure cyclic rational m^∞-torsion, at which point the usual reduced m-Tate pairing can be used. The method from [6] applies to arbitrary orientations and avoids such walks, but it does not use cyclic self-pairings; rather, it uses self-pairings with non-cyclic domains and, as a result, the argumentation becomes more intricate; see Remark 5.6.8 for a discussion. So we hope to have convinced the reader that, at least conceptually, this new method is simpler. It is also helpful in understanding and generalizing the "not walking to the floor" phenomenon from [8, App. A], as was already discussed in Remark 5.5.3.

Remark 5.6.6 If $r \geq 4$ then we can use the cyclic self-pairings of order 2^{r-2} from Section 5.5 for determining $N(\mathfrak{a})$ mod 2^{r-2}, and this is enough for evaluating $\delta, \epsilon, \delta\epsilon$ in case they exist. The situation is more subtle if

- $r = 2$ and $-\Delta_{\mathcal{O}}/4 \equiv 1 \bmod 4$ (to evaluate δ),
- $r = 3$ (to evaluate one of ϵ, $\delta\epsilon$).

Both cases can be handled by descending to elliptic curves that are primitively $(\mathbf{Z} + 2\mathcal{O})$-oriented, similar to the approach from [8, §3.1]. In the former case this may not be needed: according to Proposition 5.8.1, there may exist cyclic self-pairings that allow us to compute $N(\mathfrak{a})$ mod 4 directly. Indeed, for $k = \mathbf{F}_p$ and $\mathcal{O} = \mathbf{Z}[\sqrt{-p}]$ this is handled by the semi-reduced Tate pairing from [8, Rmk. 11], which was studied precisely for this purpose. But for arbitrary orientations we are currently missing such a pairing. ◊

Remark 5.6.7 If $m = \text{char}(k)$ then our order-m cyclic self-pairing is not available. However, in view of the character relation [8, Eq. (1)] it is always possible to discard one assigned character, so this concern is usually void.[5] This is in complete analogy with [8, 6]. $\Diamond$

Remark 5.6.8 In [6] an alternative attack to the DDH problem for oriented curves, that applies to arbitrary orientations, is described, using the Weil pairing rather than the Tate pairing. Here, the situation is slightly more intricate, in the sense that the domain of the self-pairing is no longer cyclic. More specifically, the self-pairing associated to [6, Thm. 1] may be constructed as follows. Let $\mathcal{O}$ be an imaginary quadratic order, let E be an $\mathcal{O}$-oriented elliptic curve, and suppose that $m \mid \Delta_{\mathcal{O}}$ for some odd prime number m. Then we can write $\mathcal{O} = \mathbb{Z}[\sigma]$, for some σ of norm coprime to m [6, Lem. 1]. We define $f : E[m] \to \mu_m,\ f(P) := e_m(P, \sigma(P))$. One easily checks that this is indeed a non-trivial self-pairing compatible with horizontal isogenies. Interestingly, the proof of [6, Thm. 1] shows that f can still be employed to recover the norm of a connecting ideal up to squares modulo m. A similar phenomenon occurs in [6, Prop. 1 & 2], where the associated self-pairings are maps $E[2] \to \mu_4$ and $E[4] \to \mu_8$ respectively. $\Diamond$

5.7 Conclusions and open problems

In this paper we have derived necessary and sufficient conditions for non-trivial cyclic self-pairings that are compatible with oriented isogenies, to exist. We have given examples of such pairings based on the generalized Weil and Tate pairings.

As an application, we have identified weak instances of class group actions assuming the degree of the secret isogeny is known and sufficiently small; some of these instances succumb to a polynomial time attack. We note that these cases are rare, but exist nonetheless; this situation is somewhat reminiscent of anomalous curves for which the ECDLP can be solved in polynomial time [33, 37]. These instances can be easily identified in that they require (large) square factors of $\Delta_{\mathcal{O}}$. This also shows that protocols that operate on or close to the crater are immune to this attack. To err on the side of caution it is probably best to limit oneself to (nearly) prime $\Delta_{\mathcal{O}}$.

The following problems remain open:

- In our attack we require square factors m^2 of $\Delta_{\mathcal{O}}$ to be able to derive the action of the secret isogeny on the full $E[m]$, which is required as input to the algorithm from [30]. However, it is well known that a degree d isogeny is uniquely determined if it is specified on more than $4d$ points, so knowing the image of a single point of order $m > 4d$ should suffice. The problem remains to find a method akin to [30] that can handle such one-dimensional input.

- Is it possible to exploit partial information, e.g. how valuable is it to know the action of a secret isogeny on a single point of order $m < 4d$?

[5] If $\text{char}(k) = 2$ then it seems like we may be missing more than one assigned character, but see [6, Footnote 1] for why this is not the case.

- At the moment we have only used the generalized Weil and Tate pairings for endomorphisms, whereas the definition also allows for more general isogenies ψ. Can this somehow be exploited in a more powerful attack?

- Our definition of a self-pairing on cyclic groups of even order allows for instances not derived from a bilinear pairing, e.g. the semi-reduced Tate pairing given in [8, Rmk. 11]. Proposition 5.8.1 below shows that such self-pairings indeed exist more generally, but unfortunately the proof does not give a method to efficiently compute them. Regardless of computational considerations, it would be interesting to find a more direct mathematical construction of these self-pairings and thereby genuinely complete the classification from Sections 5.4 and 5.5.

- Are there efficient Miller-type algorithms for computing the generalized Weil and Tate pairings? If not, do they exist for a larger class of endomorphisms than just $\sigma = \pi_q - 1$? At least, can these pairings be computed without needlessly extending the base field?

5.8 Relaxing the compatibility assumption

Proposition 5.8.1 *We inherit the notation/assumptions from Proposition 5.4.8, but now we only require that our cyclic self-pairing*

$$f : C \to \overline{k}^*$$

of order m is compatible with endomorphisms $\iota(\sigma)$ for which $\gcd(N(\sigma), m) = 1$. Then $\operatorname{char}(k) \nmid m$, and writing $\Delta_{\mathcal{O}} = -2^r n$ with n odd, we have:

(a) if $r = 0$ and $n \equiv 3 \bmod 8$ then $m \mid \Delta_{\mathcal{O}}$,

(b) if $r = 0$ and $n \equiv 7 \bmod 8$ then $m \mid 2\Delta_{\mathcal{O}}$,

(c) if $r = 2$ and $n \equiv 1 \bmod 4$ then $m \mid \Delta_{\mathcal{O}}$,

(d) if $r = 2$ and $n \equiv 3 \bmod 4$ then $m \mid \Delta_{\mathcal{O}}/2$,

(e) if $r = 3, 4$ then $m \mid \Delta_{\mathcal{O}}/4$,

(f) if $r \geq 5$ then $m \mid \Delta_{\mathcal{O}}/2$.

Conversely, if m satisfies these necessary conditions, then we can equip every $\mathcal{O}$-oriented elliptic curve (E, ι) over k for which the orientation is locally primitive at m with a cyclic self-pairing

$$f_{(E,\iota)} : C_{(E,\iota)} \to \overline{k}^*$$

of order m, such that these self-pairings are compatible with all K-oriented isogenies of degree coprime with m (as usual, K denotes the imaginary quadratic number field containing $\mathcal{O}$).

Proof. Write $m = 2^s m'$ with m' odd. Note that the statement $\mathrm{char}(k) \nmid m$ is again immediate.

In order to prove the other divisibility conditions, it is easy to see that one can always find a generator $\sigma \in \mathcal{O}$ of norm coprime with m', and by mimicking the proof of Proposition 5.4.8 (see the part "If $s \leq 1$ then ...") we find that $m' \mid \Delta_{\mathcal{O}}$. Since the self-pairing

$$C \to \overline{k}^* : P \mapsto f(P)^{m'} \tag{5.12}$$

has order 2^s, the remaining divisibility conditions just follow from the case $m = 2^s$ which is discussed below. This ignores a subtlety, namely that (5.12) may be incompatible with endomorphisms σ for which $\gcd(N(\sigma), 2^s m') \neq 1$, rather than just $\gcd(N(\sigma), 2^s) \neq 1$. However, it is easy to check that the proof below does not suffer from this.

As for the converse statement, the cyclic self-pairings

$$f_{(E,\iota),m'} : C_{(E,\iota),m'} \to \overline{k}^*$$

of order m' that were constructed in Section 5.5 are compatible with K-oriented isogenies of *any* degree. So, here too, if we manage to find cyclic self-pairings

$$f_{(E,\iota),2^s} : C_{(E,\iota),2^s} \to \overline{k}^*$$

of order 2^s that are compatible with K-oriented isogenies of odd degree, then

$$C_{(E,\iota),2^s} \times C_{(E,\iota),m'} \to \overline{k}^* : P \mapsto f_{(E,\iota),2^s}(P) f_{(E,\iota),m'}(P)$$

is a family of cyclic self-pairings of the desired kind (we can assume that $C_{(E,\iota),2^s}$ is 2-primary, so that the domain is indeed cyclic).

Therefore, from now on we concentrate on the case $m = 2^s$, i.e., $m' = 1$. We proceed by the case distinction from the proposition statement:

(a) If $s \geq 1$ then by Lemma 5.4.6 we know that $C[2] \cong \mathbf{Z}/2\mathbf{Z}$. The generator $\sigma = (1+\sqrt{\Delta_{\mathcal{O}}})/2$ satisfies $\mathrm{Tr}(\sigma) \equiv N(\sigma) \equiv 1 \bmod 2$, so when acting on $E[2]$ it has characteristic polynomial $x^2 + x + 1$, which is irreducible. But by compatibility with σ we know that $C[2]$ is an eigenspace: a contradiction.

(b) If $s \geq 2$ then as in the proof of Proposition 5.4.8 we find that $n = N(\sqrt{\Delta_{\mathcal{O}}})$ must be a square modulo 4: a contradiction. If $s = 1$ then we can construct the desired family of self-pairings as follows. Let $C_{(E,\iota)}$ be the subgroup of $E[2]$ that is fixed by $\sigma = (1 + \sqrt{\Delta_{\mathcal{O}}})/2$. This is a cyclic group of order 2 because the characteristic polynomial is $x^2 + x$ in this case. We then simply define

$$f_{(E,\iota)} : C_{(E,\iota)} \to \{\pm 1\} : P \mapsto -1, 0_E \mapsto 1$$

It is trivial that this family is compatible with K-oriented isogenies of odd degree (but note, as a sanity check for Proposition 5.4.8, that it is not compatible with the even-degree endomorphism σ).

We now discuss the cases $r \geq 2$. Note that the existence part is completely covered by Section 5.5, so it suffices to prove the necessary conditions, except in cases (c) and (f). We will use the notation

$$\sigma_a := a + \sqrt{\Delta_{\mathcal{O}}}/2$$

for any $a \in \mathbf{Z}$. This is an element of $\mathcal{O}$ with norm $a^2 + 2^{r-2}n$.

(c) If $s \geq 3$ then we arrive at a contradiction because $\{n, n+4\} = \{N(\sigma_0), N(\sigma_2)\}$ must both be squares modulo 8.

For existence when $s = 2$, fix an $\mathcal{O}$-oriented elliptic curve (E, ι) and consider the non-zero point $P \in E[2]$ annihilated by σ_1. This point exists because the characteristic polynomial of σ_1 mod 2 is x^2, and it is unique because otherwise $E[2] \subseteq \ker(\sigma_1)$ would imply that 4 divides $1 + n$, a contradiction. Consider the self-pairing

$$f_{(E,\iota)} \colon C_{(E,\iota)} \to \mu_4 \colon P \mapsto \zeta_4, 0_E \mapsto 1$$

where $C_{(E,\iota)} = \langle P \rangle$ and ζ_4 is some fixed primitive 4-th root of unity. This is indeed a self-pairing of order 4: we have

$$f_{(E,\iota)}(\lambda P) = f_{(E,\iota)}(P)^{\lambda^2}$$

for any $\lambda \in \mathbf{Z}$ because odd squares are congruent to 1 modulo 4. It is easy to see that $f_{(E,\iota)}$ is compatible with oriented endomorphisms of odd degree. Indeed, every such endomorphism σ can be written as $a + b\sigma_0$ for some integers a and b, where exactly one among a and b is even since $N(\sigma) = a^2 + b^2 n$ is odd. Thus

$$f_{(E,\iota)}(\sigma(P)) = f_{(E,\iota)}((a+b)P) = f_{(E,\iota)}(P)^{a^2+b^2+2ab} = f_{(E,\iota)}(P)^{N(\sigma)}.$$

To turn this into a family of self-pairings compatible with odd-degree K-oriented isogenies, with every $\mathcal{O}$-oriented elliptic curve (E', ι') that is connected to (E, ι) via a K-oriented isogeny of degree 1 mod 4, we associate a self-pairing as above. If (E', ι') is connected via a K-oriented isogeny of degree 3 mod 4, then we do the same, except we map P to $-\zeta_4$ instead of ζ_4. This is unambiguous because if (E', ι') was connected to (E, ι) via K-oriented isogenies of degrees 1 and 3 mod 4, then (E, ι) would have an oriented endomorphism of degree 3 mod 4: a contradiction since we have shown above that all oriented endomorphisms have norm of the form $a^2 + b^2 n$. By construction, this family of self-pairings is then indeed compatible with K-oriented isogenies of odd degree.[6]

Finally, if $s = 1$, then we can just resort to our family of self-pairings from Section 5.5.

(d) If $s \geq 2$ then we find that $n = N(\sigma_0)$ must be a square modulo 4: a contradiction.

[6] The construction may not reach every $\mathcal{O}$-oriented elliptic curve (E', ι'), because there may not exist an oriented isogeny to (E, ι), e.g. in view of [27, Prop. 3.3], but we can simply repeat the procedure inside every connected component.

(e) If $r = 3$ and $s \geq 2$ then $1 + 2n = N(\sigma_1)$ is a square mod 4, while if $r = 4$ and $s \geq 3$ then $1 + 4n = N(\sigma_1)$ is a square mod 8: contradictions.

(f) Assume $s \geq r$. By Lemma 5.4.6 we know that $C[2^{s-1}] \cong \mathbf{Z}/2^{s-1}\mathbf{Z}$. Since f is compatible with σ_a for every odd integer a, each of these endomorphisms acts on C by scalar multiplication. But then the same must be true for σ_0: let $\lambda \in \mathbf{Z}$ be a corresponding scalar. Since $\mathrm{Tr}(\sigma_0) = 0$ the eigenvalues of σ_0 acting on $E[2^{s-1}]$ are then given by $\pm\lambda$ and therefore

$$-\lambda^2 \equiv N(\sigma_0) = 2^{r-2}n \bmod 2^{s-1}. \tag{5.13}$$

On the other hand, the compatibility implies that $N(\sigma_a) \equiv (\lambda + a)^2 \bmod 2^s$ for all odd integers a. Along with the above congruence this yields $a^2 - \lambda^2 \equiv (\lambda+a)^2 \bmod 2^{s-1}$. Plugging in $a = \pm 1$ we find that $(\lambda+1)^2 \equiv (\lambda-1)^2 \bmod 2^{s-1}$, so that $\lambda \equiv 0 \bmod 2^{s-3}$. This means that the left-hand side of (5.13) vanishes mod 2^{s-1}, leaving us with $2^{r-2}n \equiv 0 \bmod 2^{s-1}$: a contradiction.

For existence when $s < r$, it suffices to assume that $s = r - 1$. Fix an $\mathcal{O}$-oriented elliptic curve (E, ι) such that the orientation is locally primitive at 2. Note that $2^{r-2} \mid N(\sigma_{2^{r-3}})$, so from Lemma 5.2.4 we see that $E[2^{r-2}, \sigma_{2^{r-3}}]$ is cyclic of order 2^{r-2}. Fix a generator P and define the self-pairing

$$f_{(E,\iota)} : C_{(E,\iota)} \to \mu_{2^{r-1}} : \lambda P \mapsto \zeta_{2^{r-1}}^{\lambda^2},$$

where $\zeta_{2^{r-1}}$ is some generator of $\mu_{2^{r-1}}$. As in (c), this is a well-defined self-pairing of order 2^{r-1}. Indeed, for any λ and t we have

$$f_{(E,\iota)}((\lambda + 2^{r-2}t)P) = f_{(E,\iota)}(P)^{\lambda^2 + 2^{r-1}t\lambda + 2^{2(r-2)}t^2} = f_{(E,\iota)}(\lambda P).$$

To see compatibility with odd-degree endomorphisms, similar to in (c), we remark that every oriented endomorphism σ can be written as $a + b\sigma_0$ for some integers a and b. In particular, $N(\sigma) = a^2 + 2^{r-2}b^2$, which is odd if and only if a is. Then

$$f_{(E,\iota)}(\sigma(P)) = f_{(E,\iota)}((a - 2^{r-3}b)P) = f_{(E,\iota)}(P)^{a^2 + 2^{r-2}ab} = f_{(E,\iota)}(P)^{N(\sigma)},$$

where the last equality follows from the fact that $ab \equiv b^2 \bmod 2$ because a is odd, hence $2^{r-2}ab \equiv 2^{r-2}b^2 \bmod 2^{r-1}$. To turn this into a family of self-pairings compatible with odd-degree K-oriented isogenies, we proceed as in (c): if (E', ι') is a primitively $\mathcal{O}$-oriented elliptic curve (locally at 2) connected to (E, ι) via a K-oriented isogeny $\phi : E \to E'$ of odd degree, then we equip (E', ι') with the above self-pairing, except that we use

$$\zeta_{2^{r-1}}^{\deg(\phi)} \quad \text{instead of} \quad \zeta_{2^{r-1}}$$

as our primitive 2^{r-1}-th root of unity, and we choose the specific generator

$P' = \phi(P)$ of $E'[2^{r-2}, \sigma_{2^{r-3}}]$.[7] To see that this self-pairing is independent of the choice of ϕ, let

$$\phi_1, \phi_2 \colon E \to E'$$

be two K-oriented isogenies of odd degree, and write P'_i for $\phi_i(P)$. Then $P'_1 = \lambda P'_2$ for some odd λ, and we need to check that $\deg(\phi_1) \equiv \lambda^2 \deg(\phi_2) \bmod 2^{r-1}$. Notice that $\hat{\phi}_2 \circ \phi_1$ is an oriented endomorphism of E sending P to $\lambda \deg(\phi_2) P$. By compatibility of $f_{(E,\iota)}$ with oriented endomorphisms of odd degree we have $(\lambda \deg(\phi_2))^2 \equiv \deg(\phi_1) \deg(\phi_2) \bmod 2^{r-1}$. The thesis immediately follows from the fact that $\deg(\phi_2)$ is a unit modulo 2^{r-1}.

□

Remark 5.8.2 The above proof naturally raises the question whether the self-pairings in the boundary cases

- $s = r = 2$, $n \equiv 1 \bmod 4$,
- $s = r - 1 \geq 4$,

whose existence was shown in a non-effective way, admit a more direct description. Such a description would be needed for these self-pairings to be of any practical use. In the former case, we know that the answer is yes for the Frobenius orientation, thanks to the semi-reduced Tate pairing from (5.6); see also Remark 5.4.10. Unfortunately, this construction is of Frey–Rück type, i.e., involving Miller functions, and we do not know if/how it generalizes to arbitary orientations. ◊

[7] Here again, as in Footnote 6, the construction may not reach every instance of (E', ι'), but we can repeat the procedure in every connected component.

5.9 Bibliography

[1] Wieb Bosma, John Cannon, and Catherine Playoust. The Magma algebra system. I. The user language. *J. Symbolic Comput.*, 24(3-4):235–265, 1997. Computational algebra and number theory (London, 1993).

[2] Peter Bruin. The Tate pairing for Abelian varieties over finite fields. *J. Théor. Nr. Bordx.*, 23:323–328, 2011.

[3] Wouter Castryck and Thomas Decru. CSIDH on the surface. In Jintai Ding and Jean-Pierre Tillich, editors, *PQCrypto 2020*, volume 12100 of *Lecture Notes in Computer Science*, pages 111–129. Springer, 2020.

[4] Wouter Castryck and Thomas Decru. An efficient key recovery attack on SIDH. In *Eurocrypt 2023 Pt. 5*, volume 14008 of *Lecture Notes in Computer Science*, pages 423–447. Springer, 2023.

[5] Wouter Castryck, Marc Houben, Simon-Philipp Merz, Marzio Mula, Sam van Buuren, and Frederik Vercauteren. Weak instances of class group action based cryptography via self-pairings. Full version on ePrint Archive available at `https://eprint.iacr.org/2023/549`, 2023.

[6] Wouter Castryck, Marc Houben, Frederik Vercauteren, and Benjamin Wesolowski. On the decisional Diffie-Hellman problem for class group actions on oriented elliptic curves. *Res. Number Theory*, 8(4):Paper No. 99, 18, 2022.

[7] Wouter Castryck, Tanja Lange, Chloe Martindale, Lorenz Panny, and Joost Renes. CSIDH: An efficient post-quantum commutative group action. In *Asiacrypt 2018 Pt. 3*, volume 11274 of *Lecture Notes in Computer Science*, pages 395–427. Springer, 2018.

[8] Wouter Castryck, Jana Sotáková, and Frederik Vercauteren. Breaking the decisional Diffie-Hellman problem for class group actions using genus theory. In *Crypto 2020 Pt. 2*, volume 12171 of *Lecture Notes in Computer Science*, pages 92–120. Springer, 2020.

[9] Daniel Cervantes-Vázquez, Mathilde Chenu, Jesús-Javier Chi-Dominguez, Luca De Feo, Francisco Rodriguez-Henriquez, and Benjamin Smith. Stronger and faster side-channel protections for CSIDH. In *Latincrypt 2019*, volume 11774 of *Lecture Notes in Computer Science*, pages 173–193. Springer, 2019.

[10] Jorge Chávez-Saab, Jesús-Javier Chi-Dominguez, Samuel Jaques, and Francisco Rodriguez-Henriquez. The SQALE of CSIDH: sublinear Vélu quantum-resistant isogeny action with low exponents. *Journal of Cryptographic Engineering*, 12(3):349–368, 2022.

[11] Mathilde Chenu and Benjamin Smith. Higher-degree supersingular group actions. *Mathematical Cryptology*, 1(2):85–101, 2021.

[12] Leonardo Colò and David Kohel. Orienting supersingular isogeny graphs. *Journal of Mathematical Cryptolology*, 14(1):414–437, 2020.

[13] Jean-Marc Couveignes. Hard homogeneous spaces, 2006. Unpublished article, available at `https://eprint.iacr.org/2006/291`.

[14] David A. Cox. *Primes of the form $x^2 + ny^2$: Fermat, class field theory, and complex multiplication.* Pure and Applied Mathematics. Wiley, second edition, 2013.

[15] Pierrick Dartois and Luca De Feo. On the security of OSIDH. In *PKC 2022 Pt. 1*, volume 13177 of *Lecture Notes in Computer Science*, pages 52–81. Springer, 2022.

[16] Luca De Feo, Tako Boris Fouotsa, Péter Kutas, Antonin Leroux, Simon-Philipp Merz, Lorenz Panny, and Benjamin Wesolowski. SCALLOP: scaling the CSI-FiSh. In *PKC 2023 Pt. 1*, volume 13940 of *Lecture Notes in Computer Science*, pages 345–375, 2023.

[17] Luca De Feo et al. Modular isogeny problems. Private communication.

[18] Gerhard Frey and Hans-Georg Rück. A remark concerning m-divisibility and the discrete logarithm in the divisor class group of curves. *Mathematics of Computation*, 62(206):865–874, 1994.

[19] Steven Galbraith. Pairings. In Ian F. Blake, Gadiel Seroussi, and Nigel P. Smart, editors, *Advances in Elliptic Curve Cryptography*, volume 317 of *LMS Lecture Note Series*, chapter 9, pages 183–213. Cambridge University Press, 2005.

[20] Theodoulos Garefalakis. The generalized Weil pairing and the discrete logarithm problem on elliptic curves. In *Latin 2002: Theoretical Informatics*, volume 2286 of *Lecture Notes in Computer Science*, pages 118–130. Springer, 2002.

[21] 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.

[22] Hendrik W. Lenstra. Complex multiplication structure of elliptic curves. *J. Number Theory*, 56:227–241, 1996.

[23] Luciano Maino, Chloe Martindale, Lorenz Panny, Giacomo Pope, and Benjamin Wesolowski. A direct key recovery attack on SIDH. In *Eurocrypt 2023 Pt. 5*, volume 14008 of *Lecture Notes in Computer Science*, pages 448–471. Springer, 2023.

[24] Victor S. Miller. Short programs for functions on curves, 1986. Unpublished note, available at `https://crypto.stanford.edu/miller/miller.pdf`.

[25] Victor S. Miller. The Weil pairing, and its efficient calculation. *Journal of Cryptology*, 17(4):235–261, 2004.

[26] Tomoki Moriya, Hiroshi Onuki, and Tsuyoshi Takagi. SiGamal: a supersingular isogeny-based PKE and its application to a PRF. In *Asiacrypt 2020 Pt. 2*, volume 12492 of *Lecture Notes in Computer Science*, pages 551–580. Springer, 2020.

[27] Hiroshi Onuki. On oriented supersingular elliptic curves. *Finite Fields Appl.*, 69:Paper No. 101777, 18, 2021.

[28] Damien Robert. Efficient algorithms for abelian varieties and their moduli spaces, 2021. Habilitation à Diriger des Recherches.

[29] Damien Robert. Some applications of higher dimensional isogenies to elliptic curves (overview of results), 2022. Preprint available at https://eprint.iacr.org/2022/1704.

[30] Damien Robert. Breaking SIDH in polynomial time. In *Eurocrypt 2023 Pt. 5*, volume 14008 of *Lecture Notes in Computer Science*, pages 472–503. Springer, 2023.

[31] Damien Robert. The geometric interpretation of the Tate pairing and its applications, 2023. Preprint available at https://eprint.iacr.org/2023/177.

[32] Alexander Rostovtsev and Anton Stolbunov. Public-key cryptosystem based on isogenies, 2006. Unpublished article, available at https://eprint.iacr.org/2006/145.

[33] Takakazu Satoh and Kiyomichi Araki. Fermat quotients and the polynomial time discrete log algorithm for anomalous elliptic curves. *Commentarii Mathematici Universitatis Sancti Pauli*, 47:81–92, 1998.

[34] René Schoof. Elliptic curves over finite fields and the computation of square roots mod p. *Mathematics of Computation*, 44(170):483—494, 1985.

[35] René Schoof. Nonsingular plane cubic curves over finite fields. *J. Combin. Theory Ser. A*, 46(2):183–211, 1987.

[36] Joseph H. Silverman. *The arithmetic of elliptic curves*, volume 106 of *Graduate Texts in Mathematics.* Springer, second edition, 2009.

[37] Nigel P. Smart. The discrete logarithm problem on elliptic curves of trace one. *Journal of Cryptology*, 12:193–196, 1999.

[38] William C. Waterhouse. Abelian varieties over finite fields. *Ann. Sci. École Norm. Sup.*, 2:521–560, 1969.

[39] Benjamin Wesolowski. The supersingular isogeny path and endomorphism ring problems are equivalent. In *IEEE FOCS 2021*, pages 1100–1111, 2022.

Chapter 6

Generalized class polynomials

This chapter consists of a paper written together with Marco Streng. It has been published as

Marc Houben and Marco Streng. Generalized class polynomials. *Res. Number Theory*, 8(4):Paper No. 103, 26, 2022. `https://doi.org/10.1007/s40993-022-00400-2`.

Both authors of this paper contributed equally to the work.

Compared to the published version, we corrected minor typos. The numbering (of e.g. theorems and definitions) in the published version is different.

Acknowledgements

We thank Karim Belabas, Peter Bruin, Andreas Enge, David Kohel, Filip Najman, and Andrew Sutherland for helpful discussions and Edgar Costa for helping us use the LMFDB. We would further like to thank the organisers Samuele Anni, Valentijn Karemaker, and Elisa Lorenzo Garcia of AGC^{2}T 2021 for the inspiring setting in which the first ideas for this project came to be. Finally, we would like to thank the anonymous referees for their many comments that led to improvements of the exposition.

Abstract

The Hilbert class polynomial has as roots the j-invariants of elliptic curves whose endomorphism ring is a given imaginary quadratic order. It can be used to compute elliptic curves over finite fields with a prescribed number of points. Since its coefficients are typically rather large, there has been continued interest in finding alternative modular functions whose corresponding class polynomials are smaller. Best known are Weber's functions, which reduce the size by a factor of 72 for a positive density subset of imaginary quadratic discriminants. On the other hand, Bröker and Stevenhagen showed that no modular function will ever do better than a factor of 100.83. We introduce a generalization of class polynomials, with reduction factors that are not limited by the Bröker-Stevenhagen bound. We provide examples matching Weber's reduction factor. For an infinite family of discriminants, their reduction factors surpass those of all previously known modular functions by a factor at least 2.

6.1 Introduction

The *Hilbert class polynomial* $H_D[j]$ of the imaginary quadratic order $\mathcal{O}$ of discriminant D is the minimal polynomial of the j-invariant of an elliptic curve with endomorphism ring $\mathcal{O}$. It is a defining polynomial of the ring class field of $\mathcal{O}$ and can be used for constructing elliptic curves over a finite field with a given number of points. Its coefficients are however rather large, which limits its practical usefulness. Already in 1908, Weber [37] therefore introduced alternative *class invariants* to be used instead of j, which resulted in *class polynomials* with coefficients that have roughly 1/72 of the digits of the coefficients of the Hilbert class polynomial for certain discriminants.

There has been continued interest in alternative class invariants ever since (e.g. [2, 30, 18, 17, 31, 8, 10, 11, 4, 14, 12, 9]). None however matched, let alone surpassed, the factor 72 of Weber's functions. Moreover, Bröker and Stevenhagen [4] showed that no class invariant will ever do better than a factor 100.83. Under Selberg's eigenvalue conjecture [32, Conjecture 1], this bound reduces to 96.

We introduce *generalized (multivariate) class polynomials*, define an appropriate notion of their *reduction factor*, and show that this notion indeed gives a measure of their "size" compared to the Hilbert class polynomial (Section 6.3). Contrary to classical class polynomials, the reduction factors of generalized class polynomials are not limited by the Bröker-Stevenhagen bound.

We give a family of generalized class polynomials for which we prove that the reduction factor matches Weber's 72 for a large range of values of D, including infinitely many values of D where no reduction of 36 or better was previously known (Section 6.4). We also give an example that possibly achieves the factor 120 (Remark 6.7.6).

Though the focus of this paper is on introducing the generalized class invariants and studying their height, we also give a preliminary analysis indicating that the height reduction leads to a speed-up in their computation (Section 6.6), and we show how to use them for constructing elliptic curves over finite fields (Section 6.5).

6.2 Generalized class polynomials

Definition 6.2.1 By a *modular curve over* $\mathbf{Q}$ we mean a smooth, projective, geometrically irreducible curve C over $\mathbf{Q}$ together with a map $\psi : \mathbf{H} \to C(\mathbf{C})$ from the upper half space $\mathbf{H} \subset \mathbf{C}$ with the following property. There exists a positive integer N such that for every function $f \in \mathbf{Q}(C)$, the function $f \circ \psi$ is a modular function for $\Gamma(N)$ with all q-expansion coefficients in $\mathbf{Q}^{\mathrm{ab}}$.

We identify f with $f \circ \psi$ and we identify ψ with the induced morphism of curves $X(N) \to C$. △

For an order $\mathcal{O}$ in an imaginary quadratic number field K, we denote by $K_{\mathcal{O}}$ the associated ring class field. Let f be a modular function and $\tau \in \mathbf{H}$ imaginary quadratic, say a root of $aX^2 + bX + c$ for coprime integers a, b, c. The pair (f, τ) is called a *class invariant* for the imaginary quadratic order $\mathcal{O} = \mathbf{Z}[a\tau]$ if $f(\tau)$ lies in the ring class field $K_{\mathcal{O}}$. The *discriminant* D of the class invariant is the discriminant of

$\mathcal{O}$. The Galois group G of $K(f(\tau))/K$ is isomorphic via the Artin map to a quotient of the Picard group $\mathrm{Cl}(\mathcal{O})$. Associated to a class invariant is its minimal polynomial over K, also known as the *class polynomial*,

$$H_\tau[f] := \prod_{\sigma \in G} \big(X - \sigma(f(\tau))\big) \quad \in K[X].$$

Under additional restrictions, class polynomials can sometimes be shown to have coefficients in $\mathbf{Q}$ (cf. [9, Thm. 4.4], [13, Thm. 5.4]); in that case we call the class polynomials *real*. Oftentimes, a modular function admits class invariants for an infinite family of discriminants, determined by a certain congruence condition ([31], [9, Thm. 4.3]). Sometimes the discriminant uniquely determines the class polynomial for a given modular function.

Example 6.2.2 The modular j-function admits a unique class polynomial for any discriminant $D < 0$, called the *Hilbert class polynomial* $H_D[j] := H_\tau[j]$. It can be seen as a function on $\mathbf{P}^1$ whose zeros are the j-invariants of elliptic curves with CM by the imaginary quadratic order of discriminant D and whose poles are restricted to the point at infinity. ☆

We propose a generalization of class polynomials, seen as functions on modular curves of higher genus, for which the classical class polynomials can be viewed as the genus zero case. We will mostly restrict ourselves to the case of genus one, as this will make notation considerably less complicated. We discuss the arbitrary genus case in Section 6.7. Let C be a modular curve over $\mathbf{Q}$ with a smooth Weierstrass model $y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$, and suppose that $(x, \tau), (y, \tau)$ are class invariants for some imaginary quadratic $\tau \in \mathbf{H}$. Consider $G = \mathrm{Gal}(K(x(\tau), y(\tau))/K)$ and $m = \#G$. If we denote by $\mathcal{D}$ the divisor of the unique point at infinity of C, then $\mathcal{L}(\infty\mathcal{D})$ has a basis $b_0 = 1, b_1 = x, b_2 = y, b_3 = x^2, b_4 = xy, b_5 = x^3, b_6 = x^2y, \ldots$ (ordered by ascending degree). There exist $a_i \in K$, not all zero, such that

$$\sum_{i=0}^{m} a_i b_i(\tau) = 0. \tag{6.1}$$

In fact, up to scaling by an element of $K^\times$, there exists a unique function $F_\tau[C] = \sum_{i=0}^{m} a_i b_i \in K(C)$ such that

$$\mathrm{div} F_\tau[C] = \left[\sum_{\sigma \in G} (\sigma(\psi(\tau)))\right] + \left(-\sum_{\sigma \in G} \sigma(\psi(\tau))\right) - (m+1)\mathcal{D}. \tag{6.2}$$

Definition 6.2.3 We call $F_\tau[C]$ as in (6.2) a *generalized class function* for τ. The associated *generalized class polynomial* is the unique $H_\tau[C] \in K[X, Y]$ of degree ≤ 1 in Y such that $H_\tau[C](x, y) = F_\tau[C]$. △

We note that the polynomial $H_\tau[C]$ depends on the choice of x and y, but we leave this out of the notation. In Section 6.7 (and in particular Definition 6.7.3) we will allow more general divisors $\mathcal{D}$ and bases $\mathcal{B}$, leading to more general functions $F_\tau[C,\mathcal{B}]$ and polynomials $H_\tau[C,\mathcal{B}]$.

Definition 6.2.4 We call the point $P=\sum_{\sigma\in G}\sigma(\psi(\tau))\in C(K)$ the *Heegner point* of the class function F. △

If the Heegner point P is the point at infinity, then $a_m=0$. Otherwise, the point $-P$ is a zero of F. In particular, if $P=-(0,0)$, then $a_0=0$.

For $N\in\mathbf{Z}_{>0}$, we denote by $X^0(N)$ the smooth, projective, geometrically irreducible curve over $\mathbf{Q}$ with function field consisting of the modular functions for the modular group $\Gamma^0(N)=\{\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)\in\mathrm{SL}_2(\mathbf{Z})\mid b\equiv 0 \pmod N\}$ that have rational q-expansion. We denote by $X^0_+(N)$ the quotient of $X^0(N)$ by the Fricke-Atkin-Lehner involution $z\mapsto -N/z$, and write $\eta(z)$ for the Dedekind η-function

$$\eta(z)=q^{1/24}\prod_{n=1}^{\infty}(1-q^n),\quad\text{where}\quad q=\exp(2\pi i z).$$

Example 6.2.5 Consider the genus one modular curve $C:=X^0_+(119)$. Its conductor as an elliptic curve is 17 (Cremona label 17a4)[1]. A Weierstrass model for E is given by[2]

$$y^2+3xy-y=x^3-3x^2+x, \tag{6.3}$$

where $x,y\in\mathbf{Q}(C)$ have respective q-expansions

$$\begin{aligned} x &= q^{-2}+q^{-1}+1+q+2q^2+2q^3+3q^4+3q^5+4q^6+5q^7+\ldots,\\ y &= q^{-3}+1+2q+2q^2+4q^3+4q^4+7q^5+9q^6+12q^7+\ldots,\\ &\quad\text{where this time } q=\exp(2\pi i z/119).\end{aligned}$$

The "double eta quotient" $\mathfrak{w}_{7,17}$ given by

$$\mathfrak{w}_{7,17}(z)=\frac{\eta(z/7)\eta(z/17)}{\eta(z)\eta(z/119)} \tag{6.4}$$

[1] One way to deduce this is as follows. Using the command `J0(119).decomposition()` in SageMath [36] one finds that C has conductor 17. For each of the Weierstrass models of the now finitely many possible curves [23], there are finitely many options for the divisor of the function $\mathfrak{w}_{7,17}$ given by (6.4). The curve C has two rational CM points (both of discriminant -19), so given a possible Weierstrass model together with a possible divisor for $\mathfrak{w}_{7,17}$, one can first determine $\mathfrak{w}_{7,17}$ as a function of the Weierstrass coordinates x,y by evaluating in one CM point, and then determine whether it has the expected value in the other CM point. This process excludes all but one of the options, and we at once in fact deduce both the Weierstrass model (6.3) and the relation between $\mathfrak{w}_{7,17}$ and x and y (6.5).

[2] We note that a slightly "simpler" Weierstrass model $v^2+uv+v=u^3-u^2-u$ exists by taking $u=x$ and $v=-y-2x$, but the given model (6.3) turns out to yield slightly better practical reduction factors (see Section 6.4.5).

is invariant under the action of $\Gamma^0(N)$ [27, Thm. 1] and the Fricke-Atkin-Lehner involution [11, Thm. 2], hence also forms an element of the (rational) function field of C. It is related to x and y by

$$\mathfrak{w}_{7,17} = -y + x^2 - x. \tag{6.5}$$

The curve $X_+^0(119)$ has two cusps, and they are both rational. In the given Weierstrass model, these correspond to the point $(0,0)$ and the point at infinity. Numerical examples of generalized class polynomials specifically for $X_+^0(119)$ are given in Section 6.4.5. We will treat this curve as our main test case in the rest of the paper. ☆

6.3 Estimates and reduction factors

6.3.1 Reduction factors

We define the *reduction factor* of a modular curve C to be

$$r(C) = \frac{\deg(j : X(N) \to \mathbf{P}^1)}{\deg(\psi : X(N) \to C)}. \tag{6.6}$$

In the case $C = \mathbf{P}^1$, we denote this number also by $r(\psi)$ and our notation and terminology coincide with that of [4]. The number $r(\psi)^{-1}$ is denoted by $\widehat{c}(\psi)$ in [8] and by $c(\psi)$ in [9]. Bröker and Stevenhagen [4, Theorem 4.1][3] show $r(\psi) \leq 32768/325 \leq 100.83$. Under Selberg's eigenvalue conjecture, one can even prove $r(\psi) \leq 96$. The best known ψ achieves $r(\psi) = 72$. This result does not however apply directly to $r(C)$. For example, we have

$$r(X^0(N)) = N \prod_{p|N} (1 + \frac{1}{p}) \quad \text{and} \quad r(X_+^0(N)) = \frac{1}{2} r(X^0(N)) \quad \text{if} \quad N > 1. \tag{6.7}$$

Our main example $C = X_+^0(119)$ therefore achieves $r(C) = \frac{1}{2}(7+1)(17+1) = 72$. For (hyper)elliptic modular curves C we get $r(C) \leq 201.65$ (or $r(C) \leq 192$ under Selberg's eigenvalue conjecture), by applying the bounds to the x-function. Surprisingly, all elliptic curve quotients of $X^0(N)$ we found so far have $r \leq 72$ (Section 6.4.7). In Section 6.7 we will discuss higher-genus curves, which allow for unbounded $r(C)$.

Remark 6.3.1 In the applications we have in mind, the reduction factor is the main source of improvement in computational efficiency. It is important to note, however, that this number $r(C)$ does not tell the complete story, even in the "classical" setting ($C \cong \mathbf{P}^1$), for example for the following reasons.

1. There are many challenges when computing class polynomials, and even more with generalized class polynomials. See Section 6.6.

[3] The arXiv version v1 of [4] has weaker bounds than the final publication and needs to be combined with [21, Appendix 2] to get the same result.

2. In the CM method (Section 6.5), we will want to find a j-invariant in $\mathbf{F}_p$ from a point in $C(\mathbf{F}_p)$. This is done using the minimal polynomial of the j-function over $\mathbf{Q}(C)$, known as the *modular polynomial* (Lemma 6.5.1). This works best if the degree of j over $\mathbf{Q}(C)$ is small. For example, this degree is 1 for $C = X^0(N)$, is 2 for $C = X^0_+(N)$, and ranges from 1 to 20 in [9, Table 7.1], making $X^0_+(119)$ a good choice in this respect.

3. If the (generalized) class polynomial is not real, then its coefficients lie in an imaginary quadratic extension of $\mathbf{Q}$; roughly doubling its bit size. This issue can be avoided by imposing additional restrictions on C or τ, see Sections 6.4.2 and 6.4.3.

On the other hand, there are two important tricks that may be used in complementary directions, providing computational improvements beyond the reduction factor $r(C)$:

1. Under some constraints, typically when all primes dividing the level of the modular curve ramify in the CM field, both the degree and height of the class polynomial are cut in half. This happens for example in the record-computation of [14] for the Atkin invariant A_{71} when 71 divides the discriminant, leading to class polynomials that are $2^2 \cdot 36 = 144$ times smaller than the Hilbert class polynomial (note that the reduction factor $r(A_{71})$ is 36 in this case). The same trick also applies to generalized class polynomials, see Section 6.4.4, which in the case of $X^0_+(119)$ leads to a factor $2^2 \cdot 72 = 288$ in size reduction.

2. When the class number is composite, one can decompose the ring class field into a tower of fields whose defining polynomials have smaller degrees, also leading to a significant speed-up in the CM method [35].

These last two tricks only work when the class number is composite. We expect both of them to work well for generalized class polynomials, so will mainly restrict to the case of prime class number in our examples, as this more clearly illustrates the role of the parameter $r(C)$. ◊

The goal of the rest of this section is to show under some hypotheses that the reduction factor $r(C)$ is indeed an asymptotic reduction factor of the size of the polynomials involved. For that, we will first introduce the appropriate notions of "size".

6.3.2 Measures of polynomials and heights of their roots

For a polynomial $A \in \mathbf{C}[X]$, let $|A|_1$ (resp. $|A|_\infty$) be the sum (resp. maximum) of the absolute values of the coefficients of A. The *Mahler measure* of a polynomial $A = a\prod_{i=1}^{n}(X - \alpha_i) \in \mathbf{C}[X]$ is

$$\mathcal{M}(A) = |a| \prod_i \max\{1, |\alpha_i|\}.$$

Lemma 6.3.2 *We have*

$$\begin{array}{rcccl} |A|_\infty & \leq & |A|_1 & \leq & (n+1)|A|_\infty, \\ \mathcal{M}(A) & \leq & |A|_1 & \leq & 2^n\mathcal{M}(A), \end{array}$$

$$\begin{array}{rcl} |\log|A|_1 - \log|A|_\infty| & \leq & \log(n+1), \\ |\log|A|_\infty - \log(\mathcal{M}(A))| & \leq & n\log(2). \end{array}$$

Proof. The first two inequalities are by definition and the third is Equation (6) of [24]. For its converse, observe that we have $|AB|_1 \leq |A|_1|B|_1$, and hence also $|A|_1 \leq |a|\prod_i \max\{2, 2|\alpha_i|\} \leq 2^n\mathcal{M}(A)$. Then take logarithms. □

For an element α in a number field L of degree n, we define its *(absolute logarithmic) height* to be

$$h(\alpha) = \frac{1}{n}\sum_v \max\{0, \log|\alpha|_v\},$$

where the sum ranges over the Archimedean and non-Archimedean absolute values, suitably normalized (that is, those denoted $||\cdot||_v$ in [19, §B.1]). If α is a root of an irreducible $A \in \mathbf{Z}[X]$ of degree n, then we have

$$\log(\mathcal{M}(A)) = nh(\alpha). \tag{6.8}$$

Remark 6.3.3 Another measure for the complicatedness of A would be its total bit size, or the sum s of the logarithms of the absolute values of the nonzero coefficients. We will instead focus on $|A|_\infty$ for the following reasons.

First of all, for computational purposes, it is more useful to look at $p = \deg(A)\cdot\log|A|_\infty$, as the required precision (or number of primes with the CRT approach) is proportional to $\log|A|_\infty$ and the number of computations to do with that precision is proportional to $\deg(A)$.

Secondly, we get the impression from numerical computations that s is close to p. For example, the value of s/p is spread out over the interval $(0.75, 0.9)$ for the larger discriminants in both Section 6.4.5 and Example 6.7.4.

Finally, it is hard to prove lower bounds on s other than $s \geq \log|A|_\infty$, as it seems to already be hard to show that a sufficient proportion of coefficients is nonzero. ♢

6.3.3 Proof of the height reduction

Theorem 6.3.4 *Let C be a modular curve over $\mathbf{Q}$ and suppose that C is an elliptic curve of rank 0 with Weierstrass coordinates x and y. Suppose that $\tau \in \mathbf{H}$ ranges over a sequence of imaginary quadratic points for which C yields real generalized class polynomials $H_\tau[C]$, and with*

$$\frac{h(j(\tau))}{\log(\log(\#\operatorname{Cl}(\mathcal{O})))} \to \infty. \tag{6.9}$$

Scale each $H_\tau[C]$ such that it has coprime coefficients in $\mathbf{Z}$. *Then*

$$d \cdot \frac{\log |\mathbf{F}|_\infty}{\log |H_\tau[j]|_\infty} \to \frac{1}{r(C)},$$

where d is the degree of $K_\mathcal{O}$ over $K(\psi(\tau))$.

Remark 6.3.5 We argue that the hypothesis (6.9) is very reasonable. Under GRH, we have

$$\#\,\mathrm{Cl}(\mathcal{O}) = O(\sqrt{|D|}\log(\log|D|)), \tag{6.10}$$

where D is the discriminant of $\mathcal{O}$ (see [22, 9.Theorem 1 and 11. on page 371], suitably extended to arbitrary D.) Moreover, [8, §6.2] gives the approximation $\log|H_\tau[j]|_\infty \approx \pi\sqrt{|D|}S(D)$, with $S(D) = \sum_Q a^{-1}$, where the sum ranges over reduced primitive quadratic forms $Q = ax^2 + bxz + cz^2$ of discriminant D. We now give a heuristic lower bound of this sum on average over all $|D| \leq X$. We have $\sum_D S(D) \approx \sum_Q a^{-1}$, where this time the sum is taken over all reduced quadratic forms of negative discriminant $> -X$ (using the heuristic that imprimitive forms have a negligible contribution). As we are only computing a lower bound, we may restrict to $a \leq \sqrt{X/8}$. Then b ranges from $-a$ to a, and c ranges from a or $a+1$ to $\lfloor (X+b^2)/(4a)\rfloor$; a range that contains at least $\lfloor X/(8a)\rfloor$ integers. This yields at least roughly $X/4$ values of b and c for each a, hence $\sum_D S(D)$ is roughly at least $(X/4)\sum_{a^2 \leq X/8} a^{-1} \geq \frac{1}{8}X\log(X)$. It follows that the average $S(D)$ is at least proportional to $\log|D|$. Thus, for "average" $S(D)$, we have that $\log|H_\tau[j]|_\infty$ is at least proportional to $\sqrt{|D|}\log|D|$. Combined with (6.10), (6.8), and Lemma 6.3.2, we find for such D that $h(j(\tau))/\log(\log(\#\,\mathrm{Cl}(\mathcal{O})))$ is at least proportional to $\log|D|/(\log(\log|D|))^2$. We thus see that (6.9) indeed holds for "average" $S(D)$. ◊

Theorem 6.3.4 is the analogue of the following result.

Theorem 6.3.6 (cf. Enge-Morain [8]) *Let f be a modular function and suppose that $\tau \in \mathbf{H}$ ranges over a sequence of imaginary quadratic points for which (f, τ) is a class invariant with $h(j(\tau)) \to \infty$. Then $d \cdot \frac{\log|H_\tau[f]|_\infty}{\log|H_\tau[j]|_\infty} \to \frac{1}{r(f)}$, where d is the degree of $K_\mathcal{O}$ over $K(f(\tau))$.*

The goal of the remainder of Section 6.3 is to prove Theorem 6.3.4. We start with a proof of Theorem 6.3.6.

Proof. Let m be the degree of $K(f(\tau))$ over K and let $n = dm$ be the degree of $K_\mathcal{O}$ over K. By Lemma 6.3.2 and (6.8), we get $|\frac{1}{n}\log|H_\tau[j]|_\infty - h(j(\tau))| \leq \log(2)$ and $|\frac{d}{n}\log|H_\tau[f]|_\infty - h(f(\tau))| \leq \log(2)$.

As $h(j(\tau)) \to \infty$, we also get

$$\frac{h(f(\tau))}{h(j(\tau))} \to \frac{1}{r(f)} \tag{6.11}$$

by [19, Proposition B.3.5(b)]. Altogether, this gives the result. □

Proposition 6.3.7 *Let C be a modular curve over $\mathbf{Q}$ and suppose that C is an elliptic curve of rank 0 with Weierstrass coordinates x and y. For every imaginary quadratic $\tau \in \mathbf{H}$ for which C yields a real generalized class polynomial $H_\tau[C]$, let m be the degree of $K(\psi(\tau))$ over K and let $d' \in \{1,2\}$ be the degree of $K(\psi(\tau))/K(x(\tau))$. Scale each $H_\tau[C]$ such that it has coprime coefficients in $\mathbf{Z}$. Then we have*

$$\left|\log|H_\tau[C]|_\infty - \frac{d'}{2}\log|H_\tau[x]|_\infty\right| < B\max\{1, m\log(\log(m))\},$$

for some constant B that only depends on C and the choice of Weierstrass model.

Proof. **We first put the equation for C in a nice form.** We have $C : y^2 + g(x)y = f(x)$. Without loss of generality we have $g = 0$ and $f \in \mathbf{Z}[X]$ monic of odd degree such that $f(z) \leq -1$ for all real $z \leq 0$. Indeed, we obtain $g = 0$ by the substitution $y' = y + \frac{1}{2}g(x)$, then do scalings $x' = vx$ and $y' = wy$ to make f integral and (thanks to its odd degree) monic, and then do a substitution $x' = x + c$ to make $f(z) \leq -1$ for all $z \leq 0$. This affects $H_\tau[C] = A + BY$ and $H_\tau[x]$ as follows. The first substitution changes A into $A + \frac{1}{2}g(X)B$, the second changes A into $A(vX)$ and B into $wB(vX)$, and the third changes A into $A(X + c)$. Each of these substitutions change $\log(\max\{|A|_1, |B|_1\})$ at most by $O(m)$, as does clearing the denominators afterwards.

Next, we relate a norm of $H_\tau[C]$ to $H_\tau[x]$. The extra elliptic curve point $(a/b^2, c/b^3) := \sum_{\sigma\in G}\sigma(\psi(\tau)) \in C(\mathbf{Q})$ from (6.2) (which is minus the Heegner point) is torsion by our assumption that C has rank 0. There are finitely many torsion points in $C(\mathbf{Q})$, hence finitely many possibilities for the polynomial $T = b^2X - a$. Writing $H_\tau[C] = A(X) + B(X)Y$, we get that $N(H_\tau[C]) = A(X)^2 + (-f(X))B(X)^2$ has the same divisor as the primitive polynomial $H_\tau[x]^{d'} \cdot T$, hence there is a constant $s \in \mathbf{Z} \setminus \{0\}$ with $N(H_\tau[C]) = sH_\tau[x]^{d'} \cdot T$.

We claim that $s = \pm 1$. If not, take a prime $p \mid s$ and consider the highest-weight term of $(H_\tau[C] \bmod p)$, where X has weight 2 and Y has weight $\deg(f)$. This gives rise to the highest-degree term of $(N(H_\tau[C]) \bmod p)$, which is therefore nonzero, a contradiction.

Now we use interpolation to bound $H_\tau[C]$ in terms of $H_\tau[x]$. We will choose interpolation points $z = g(i) \leq 0$. Note that for $z \leq 0$ we have

$$A(z)^2, B(z)^2 \leq A(z)^2 + (-f(z))B(z)^2 = N(H_\tau[C]) \leq \max\{1, |z|\}^m |H_\tau[x]|_1^e |T|_1,$$

and since there are finitely many polynomials T, we get

$$\log|A(z)|, \log|B(z)| \leq \frac{m}{2}\max\{0, \log|z|\} + \frac{d'}{2}\log|H_\tau[x]|_1 + O(1).$$

Interpolation then gives, for $P \in \{A, B\}$:

$$P(X) = \sum_{i=1}^{k} P(g(i)) \prod_{j \neq i} \frac{X - g(j)}{g(i) - g(j)}, \tag{6.12}$$

where $k = \deg(P) + 1 = O(m)$.

Taking $g(u) = -\log(eu)^2$, we find $|g(i) - g(j)| \geq |i - j| \min_{z \in [1,k]} |g'(u)| = |i - j| \min_{u \in [1,k]} 2\frac{\log(eu)}{u} = 2|i - j|\frac{\log(ek)}{k}$. So for each i there are at most $k/\log(k)$ values of $j \neq i$ with $|g(i) - g(j)| < 1$ and each of them has $|g(i) - g(j)| \geq 1/k$. We get

$$\log \prod_{j \neq i} \frac{1}{|g(i) - g(j)|} \leq (k/\log(k)) \log(k) = k = O(m).$$

For the other factors in (6.12), we have that $\log |X - g(j)|_1 \leq \log(1 + \log(em)^2) = O(\log(\log(m)))$, so $\log \prod_j |X - g(j)|_1 = O(m \log(\log(m)))$, as well as $\log |P(g(i))| \leq \frac{d'}{2} \log |H_\tau[x]|_1 + O(m \log(\log(m)))$. Taking the sum in (6.12) gives another $+\log(k)$, so that the end result is $\log |P(X)|_1 \leq \frac{d'}{2} \log |H_\tau[x]|_1 + O(m \log(\log(m)))$. By Lemma 6.3.2, this also holds with $|\cdot|_\infty$, which proves the upper bound on $\log |H_\tau[C]|_\infty$.

For the lower bound, note that $H_\tau[x]^{d'}$ is a factor of $Q = A^2 - f(X) \cdot B^2$, and we have $|Q|_1 \leq |A|_1^2 + |f|_1 |B|_1^2 \leq |f|_1 (m+1)^2 |H_\tau[C]|_\infty^2$. Using the fact that $\mathcal{M}$ is multiplicative by definition and is related to $|\cdot|_1$ and $|\cdot|_\infty$ by Lemma 6.3.2, we get exactly what we need: $d' \log |H_\tau[x]|_\infty \leq d' \log \mathcal{M}(H_\tau[x]) + O(m) \leq \log \mathcal{M}(Q) + O(m) \leq \log |Q|_1 + O(m) \leq 2 \log(|H_\tau[C]|_\infty) + O(m)$. □

Proof of Theorem 6.3.4. Denote again by $n = \#\mathrm{Cl}(\mathcal{O})$ the degree of $K_{\mathcal{O}}$ over K. First we apply Theorem 6.3.6 to x and get $dd' \frac{\log |H_\tau[x]|_\infty}{\log |H_\tau[j]|_\infty} \to \frac{2}{r(C)}$. Proposition 6.3.7, together with the hypothesis $h(j(\tau))/(n \log(\log(n))) \to \infty$, gives $\frac{1}{d'} \frac{\log |H_\tau[C]|_\infty}{\log |H_\tau[x]|} \to \frac{1}{2}$ (as in the proof of Theorem 6.3.6). The product of these two limits gives the result. □

Remark 6.3.8 Theorem 6.3.4 states that asymptotically the effect of the choice of a model of the curve C is negligible, as is the effect of replacing f by $2f$ or $f + 1$ or any other element of $\mathbf{Q}(f)$ in Theorem 6.3.6.

However, in practice the error terms can be quite large and depend on these choices. For example, if f is integral over $\mathbf{Z}[j]$ then $H_\tau[f]$ is monic, and if f^{-1} is integral over $\mathbf{Z}[j]$, then f has zero constant coefficient. This can make a difference in practical examples as it forces the coefficients at the beginning and end to be small, though this improvement is negligible asymptotically by the theorems. See also Remark 6.3.3. ◊

6.4 Class invariants for $X^0(N)$ and $X_+^0(N)$

In this section we assume that C is a quotient over $\mathbf{Q}$ of $X^0(N)$; in other words, C is a smooth, projective, geometrically irreducible curve over $\mathbf{Q}$ with function field consisting only of modular functions for $\Gamma^0(N)$ that have rational q-expansion. We will show how to obtain generalized class functions for every discriminant $D < 0$ that is square modulo $4N$ (Section 6.4.1).

In some cases we get further reductions from class invariants generating subfields of $K_{\mathcal{O}}$ or from real class polynomials (Sections 6.4.2–6.4.4).

In Sections 6.4.5–6.4.6 we study what this means for $X^0_+(119)$ and in Section 6.4.7 we look for more examples of elliptic curve quotients of $X^0(N)$.

6.4.1 Class invariants for $X^0(N)$

The following result does not require C to be an elliptic curve, except that (unless C is an elliptic curve) one needs to read the definitions in Section 6.7 for the parts about generalized class polynomials.

Proposition 6.4.1 (based on Schertz [31]) *Let $C = (C, \psi)$ be a quotient over $\mathbf{Q}$ of $X^0(N)$ and let $D < 0$ be a square modulo $4N$.*

There exist $a, b, c \in \mathbf{Z}$ with $a, c > 0$, $b^2 - 4ac = D$, $N \mid c$, and $\gcd(a, N) = \gcd(a, b, c) = 1$. Choose such a, b, c, let $\tau \in \mathbf{H}$ be a root of $aX^2 + bX + c$, with order $\mathcal{O} = \mathbf{Z}[a\tau]$, which has discriminant D. Then we have

$$\psi(\tau) \in C(K_{\mathcal{O}}),$$

thus giving rise to a generalized class polynomial $H_\tau[C]$.

*The Galois orbit of $\psi(\tau)$ can be computed as follows. There exists an N-*system*, that is, there exist $\tau_1, \ldots, \tau_n \in \mathbf{H}$ such that $(\tau_i\mathbf{Z} + \mathbf{Z})_i$ is a system of representatives of $\mathrm{Cl}(\mathcal{O})$ and such that τ_i is a root of $a_iX^2 + b_iX + c_i$ with $\gcd(a_i, N) = \gcd(a_i, b_i, c_i) = 1$ and $b_i \equiv b \bmod 2N$. Moreover, for any such choice, we have*

$$\mathrm{Gal}(K_{\mathcal{O}}/K) \cdot \psi(\tau) = \{\psi(\tau_i) : i = 1, \ldots, n\}.$$

Proof. For the existence of a, b, c, take an arbitrary square root b of D modulo $4N$, let $a = 1$, and $c = (b^2 - D)/4$. Then the existence of an N-system is [31, Proposition 3].

For any $f \in \mathbf{Q}(C)$, Theorem 4 of Schertz [31] states $f(\tau) \in K_{\mathcal{O}} \cup \{\infty\}$ and gives the $\mathrm{Gal}(K_{\mathcal{O}}/K)$-orbit as $\{g(N\tau_i) : i\}$, under an additional condition on the function $f(1/z)$. However, the condition on $f(1/z)$ is not needed, as stated in Theorems 3.9 and 4.4 of [13]. This proves the result. □

6.4.2 Real class polynomials from ramification

There are some situations in which we can actually get real class polynomials, cutting the total required bit size in half. The first such situation is when all primes dividing N ramify.

Proposition 6.4.2 (based on Enge-Morain [9]) *Let $C = (C, \psi)$ be a quotient over $\mathbf{Q}$ of $X^0(N)$ and let $D < 0$ be a discriminant divisible by N if N is odd and by $4N$ if N is even.*

There exist $a, b, c \in \mathbf{Z}$ with $a, c > 0$, $N \mid b, c$, $\gcd(a, N) = 1$, and $b^2 - 4ac = D$. Choose such a, b, c, let $\tau \in \mathbf{H}$ be a root of $aX^2 + bX + c$, with order $\mathcal{O} = \mathbf{Z}[a\tau]$, which has discriminant D.

Then the $\mathrm{Gal}(K_{\mathcal{O}}/K)$-orbit of $\psi(\tau)$ is stable under complex conjugation, and hence we may take $H_\tau[C] \in \mathbf{Q}[X, Y]$.

Proof. If D is odd, take $b = N$, and if D is even, take $b = 0$. If N is even, then we find $4N \mid b^2 - D$. If N is odd, then we find both $4 \mid b^2 - D$ and $N \mid b^2 - D$, hence also $4N \mid b^2 - D$. Let $a = 1$ and $c = (b^2 - D)/4$.

The complex conjugate of $\psi(\tau)$ is $\psi(-\overline{\tau})$ by the fact that the q-expansion coefficients are real. Here $-\overline{\tau}$ is a root of $aX^2 - bX + c$, and as $N \mid b$, we can choose the N-system in Proposition 6.4.1 in such a way that $-\overline{\tau} = \tau_i$ for some i. This proves the result. □

6.4.3 Real class polynomials from $X_+^0(N)$

The second situation in which we get real class polynomials is when working with quotients of $X_+^0(N)$.

Proposition 6.4.3 (based on Theorem 3.4 of Enge-Schertz [10]) *In the situation of Proposition 6.4.1, suppose furthermore that C is a quotient of $X_+^0(N)$, and that $\gcd(c/N, N) = 1$.*

Then the $\mathrm{Gal}(K_{\mathcal{O}}/K)$-orbit of $\psi(\tau)$ is stable under complex conjugation, and hence we may take $H_\tau[C] \in \mathbf{Q}[X, Y]$.

Proof. The complex conjugate of $\psi(\tau)$ is $\psi(-\overline{\tau})$ by the fact that the q-expansion coefficients are real. As ψ is invariant under the Fricke-Atkin-Lehner involution, this in turn is $\psi(\tau')$ with $\tau' = N/\overline{\tau}$, a root of $(c/N)X^2 + bX + Na$. As c/N is coprime to N, we can choose the N-system in Proposition 6.4.1 in such a way that $\tau' = \tau_i$ for some i. This proves the result. □

To use this result, we will need $\gcd(c/N, N) = 1$, which can be achieved most of the time, as follows.

Lemma 6.4.4 *If D is a square modulo $4N$ and $D = F^2 D_0$ for a negative fundamental discriminant D_0 and a positive integer F coprime to N, then there exist a, b, c as in Proposition 6.4.1 with $\gcd(c/N, N) = 1$.*

More generally, let $D < 0$ be a square modulo $4N$. Then there exist a, b, c as in Proposition 6.4.1 with $\gcd(c/N, N) = 1$ if and only if all of the following do not hold.

1. *there exists a prime $p \mid N$ with $\mathrm{ord}_p(N)$ odd and $\mathrm{ord}_p(D) > \mathrm{ord}_p(4N)$,*

2. *$m := \mathrm{ord}_2(N) > 0$ and D is of the form $2^{m+1}d$ with $d \equiv 1 \pmod 4$,*

3. *$m := \mathrm{ord}_2(N) > 0$ and D is of the form $2^m d$ with $d \equiv 1 \pmod 8$.*

Proof. The triple (a,b,c) exists if and only if there exists $b \in \mathbf{Z}$ such that for all $p \mid N$: $\mathrm{ord}_p(b^2 - D) = \mathrm{ord}_p(4N)$.

Suppose that we are not in case (1), (2), or (3). By the Chinese remainder theorem, it suffices to find one $b \in \mathbf{Z}$ for each $p \mid N$. So let $p \mid N$ be prime and let $k = \mathrm{ord}_p(4N)$ and $l = \mathrm{ord}_p(D)$. If $k < l$, then as we are not in case (2), we find that k is even, and we can take $b = p^{(k/2)}$. If $k = l$, then we can take $b = p^e$ with $e > k/2$. Now the case $k > l$ remains. As D is a square modulo $4N$, there exists $b_0 \in \mathbf{Z}$ be such that $D \equiv b_0^2 \pmod{4N}$. If $\mathrm{ord}_p(b_0^2 - D) = \mathrm{ord}_p(4N)$, then we are done, so suppose $\mathrm{ord}_p(b_0^2 - D) > k$.

Note that $2\,\mathrm{ord}_p(b_0) = l$, hence l is even. Let $b = b_0 + p^e$ with e to be determined later. We get $b^2 - D = (b_0^2 - D) + 2p^e b_0 + p^{2e}$, and the terms have valuation $> k$, $e + (l/2) + \mathrm{ord}_p(2)$, $2e$ respectively.

If $p \neq 2$, then we choose $e = k-(l/2)$, so $2e = k+(k-l) > k$, hence $\mathrm{ord}_p(b^2-D) = k$. If $p = 2$ and $k > l+2$, then we choose $e = k - (l/2) - 1$, so $2e = k + (k - l - 2) > k$, hence $\mathrm{ord}_p(b^2 - D) = k$.

Now only the case $p = 2$ with $k - l \in \{1, 2\}$ remains. Write $d = 2^{-l}D$ and $b_1 = 2^{-(l/2)}b_0$, so b_1 is odd and $b_1^2 - d$ is divisible by 2^{k-l}.

In the case $k-l = 1$, we get $b_1^2 - d \equiv 0 \pmod 2$, and we claim that this is nonzero modulo 4. Indeed, b_1^2 is 1 modulo 4 and d is not (as we are not in case (2)). Therefore $\mathrm{ord}_2(b_1^2 - d) = 1$ and $\mathrm{ord}_2(b_0^2 - D) = 1 + l = k$, so we take $b = b_0$.

In the case $k-l = 2$, we get $b_1^2 - d \equiv 0 \pmod 4$, and we claim that this is nonzero modulo 8. Indeed, b_1^2 is 1 modulo 8, and d is not (as we are not in case (3)). Therefore $\mathrm{ord}_2(b_1^2 - d) = 2$ and $\mathrm{ord}_2(b_0^2 - D) = 2 + l = k$, so we take $b = b_0$.

Conversely, suppose that b exists.

In case (1), we have $\mathrm{ord}_p(D) > \mathrm{ord}_p(4N)$, hence $2\,\mathrm{ord}_p(b) = \mathrm{ord}_p(4N)$ is odd, contradiction.

In case (2), we have $\mathrm{ord}_2(b^2 - 2^{m+1}d) = m + 2$, hence $m + 1 = 2\,\mathrm{ord}_2(b) =: 2e$. Write $b = 2^e b_1$ and note $\mathrm{ord}_2(b_1^2 - d) = 1$, but $b_1^2 - d$ is 0 modulo 4.

In case (3), we have $\mathrm{ord}_2(b^2 - 2^m d) = m + 2$, hence $m = 2\,\mathrm{ord}_2(b) =: 2e$. Write $b = 2^e b_1$ and note $\mathrm{ord}_2(b_1^2 - d) = 2$, but $b_1^2 - d$ is 0 modulo 8.

It remains only to prove the first statement, for which it suffices to show that the exceptions (1), (2), and (3) all imply $\gcd(N, F) > 1$. In case (1), we see that $p^2 \mid D$ and if $p = 2$, then $p^4 \mid D$, hence $p \mid F$. In cases (2) and (3), write $D = 2^v d$ with $v \in \{m, m+1\}$. As D is a square modulo 2^{m+2}, we find that v is even, and hence $D = (2^{v/2})^2 d$ for a discriminant d, so $2 \mid F$. □

Lemma 6.4.5 *Let N be the product of distinct odd primes $p_1, \ldots, p_k$. The negative discriminants that are a square modulo $4N$ and not in one of the exceptions of Lemma 6.4.4 have density*

$$\prod_{i=1}^{k} \frac{p_i^2 + p_i - 2}{2p_i^2}$$

in the set of all negative discriminants.

The negative fundamental discriminants that are a square modulo $4N$ (which are *not in one of the exceptions of Lemma 6.4.4) have density*

$$\prod_{i=1}^{k} \frac{p_i^2 + p_i - 2}{2(p_i^2 - 1)}$$

in the set of all fundamental negative discriminants.

Proof. Being a discriminant is the condition of being 0 or 1 modulo 4. It is equivalent to being a square modulo 4. This is independent of being a square modulo p_i that does not suffer from (1), which is happens for the $(p_i - 1)/2$ residue classes modulo p_i that are nonzero squares modulo p_i, and the $p_i - 1$ nonzero residue classes modulo p_i^2 that are zero modulo p_i. As $p_i(p_i - 1)/2 + p_i - 1 = (p_i^2 + p_i - 2)/2$, we get the first statement.

Being a fundamental discriminant means being nonzero modulo the squares of all odd primes and being $1, 5, 8, 9, 12, 13$ modulo 16. This happens for $\zeta(2)^{-1}(1-1/4)^{-1}\frac{6}{16}$ of all negative integers. In order to restrict this to products that satisfy the conditions of Lemma 6.4.4, we have to adjust the Euler product exactly by the given factor. □

For example, if $N = 119 = 7 \cdot 17$, then the numbers in Lemma 6.4.5 are > 0.2898 and $19/64 > 0.2968$.

6.4.4 Lower-degree class polynomials from ramification

In the case where all primes dividing N ramify, we get an even greater size reduction. The point $\psi(\tau)$ will then be defined over a subfield, cutting the degree of its minimal polynomial in half. This in turn also cuts the height of the coefficients of this polynomial in half, as we get $d \geq 2$ in Theorem 6.3.4. The amount of work required for computing the class polynomial, as well as the bit size of the polynomial (Remark 6.3.3), is related to the degree times the logarithm of the largest coefficient, and this product is reduced by a factor $\geq 2 \times 2 \times r(C) = 4r(C)$.

Proposition 6.4.6 (based on Enge-Schertz [12]) *Let $C = (C, \psi)$ be a quotient over $\mathbf{Q}$ of $X^0(N)$ and let $D = F^2 D_0 < 0$ be such that $N \mid D$, $\gcd(F, N) = 1$, and $D \notin \{N, 4N\}$.*

There exist $a, b, c \in \mathbf{Z}$ with $a > 0$, $N \mid b$, $c = N$, $b^2 - 4ac = D$, and $\gcd(a, b, c) = 1$. Choose such a, b, c, let $\tau \in \mathbf{H}$ be a root of $aX^2 + bX + c$, with order $\mathcal{O} = \mathbf{Z}[a\tau]$, which has discriminant D.

Let $\mathfrak{n} = ((-b + \sqrt{D})/2, a)$, and let $K_{\mathcal{O}}^{[\mathfrak{n}]}$ be the subfield of $K_{\mathcal{O}}$ fixed by the image of $\mathfrak{n}$ under the Artin map. Then $[\mathfrak{n}]$ has order 2 in $\mathrm{Cl}(\mathcal{O})$ and $\psi(\tau) \in C(K_{\mathcal{O}}^{[\mathfrak{n}]})$, where $K_{\mathcal{O}}$ has degree 2 over $K_{\mathcal{O}}^{[\mathfrak{n}]}$.

We get $m \leq \#\mathrm{Cl}(\mathcal{O})/2$ in the definition of $H_\tau[C]$, we get $H_\tau[C] \in \mathbf{Q}[X, Y]$, and we get and $d \geq 2$ in Theorem 6.3.4.

If $\mathfrak{a}_i$ are the ideals $\tau_i\mathbf{Z} + \mathbf{Z}$ of an N-system, then $\mathfrak{a}_i$ and $\mathfrak{a}_i\mathfrak{n}$ yield the same point $\psi(\tau_i)$, while $\mathfrak{a}_i^{-1}$ and $\mathfrak{a}_i^{-1}\mathfrak{n}$ yield $\overline{\psi(\tau_i)}$.

Proof. This is exactly what we get when applying [12, Theorem 9] to the coordinate functions f of C. □

6.4.5 Numerical results for $X^0_+(119)$

For the rest of this section we will return to our main Example 6.2.5, so set $N = 119 = 7 \cdot 17$. For any τ as in Proposition 6.4.3, we have $H_\tau[C] \in \mathbf{Q}[X, Y]$. By scaling, we may assume that the coefficients of $H_\tau[C]$ are integral and coprime, and that the leading coefficient (i.e. the coefficient of the monomial of highest degree as a function on C) is positive, and this uniquely determines $H_\tau[C] \in \mathbf{Z}[X, Y]$.

For any discriminant $D < 0$ coprime to N such that D is a square modulo N, there are two generalized class polynomials (depending on the choice of τ). We experimentally computed both of these for all fundamental discriminants of prime class number < 100. The main reason for restricting to prime class number is to exclude the two tricks of Remark 6.3.1; for these discriminants, the reduction factor thus provides a fair comparison with the Hilbert class polynomial. The method we employ numerically evaluates class invariants by their q-expansions, and finds a minimal polynomial relation (6.1) using lattice basis reduction (LLL). We leave faster methods for future research, but see Section 6.6 for the first ideas. Since the q-expansions can only be evaluated up to finite precision, this does not result in provably correct polynomials, although – based on heuristic estimates – they are highly unlikely to be incorrect.

A few examples of computed polynomials are listed in Table 6.1. Here, for the given discriminant D, we consistently chose τ such that its primitive equation is $X^2 + bX + (b^2 - D)/4$ with $b \in \mathbf{Z}_{>0}$ *minimal* satisfying $b^2 \equiv D \pmod{4N}$ and $\gcd((b^2 - D)/(4N), N) = 1$.

D	n	$F_\tau[C]$
-52	2	$y + 1$
-523	5	$x^3 + x^2 - 2xy - 3x - 2y$
-5347	13	$x^7 + 58x^6 - 13x^5y - 39x^5 - 143x^4y - 85x^4 - 135x^3y$ $-19x^3 - 51x^2y + 47x^2 + 7xy - 12x - y + 1$
-15139	29	$x^{15} + 1028x^{14} - 40x^{13}y + 37342x^{13} - 10557x^{12}y + 79865x^{12}$ $-167759x^{11}y - 385199x^{11} - 474165x^{10}y - 425857x^{10} - 69261x^9y$ $+345059x^9 + 493309x^8y + 309689x^8 + 168403x^7y - 132377x^7$ $-145439x^6y - 22165x^6 - 16029x^5y + 16139x^5 + 15225x^4y - 4867x^4$ $-7127x^3y - 456x^3 + 623x^2y + 423x^2 + 337xy - 65x - 64y$

Table 6.1: Some conjecturally correct generalized class functions for $C = X^0_+(119)$. The second column lists the class number n of the discriminant D.

Still assuming that $H_\tau[C]$ is scaled such that it has coprime coefficients in $\mathbf{Z}$, we denote by

$$r_A(\tau) := \frac{\log |H_\tau[j]|_\infty}{\log |H_\tau[C]|_\infty}$$

the *practical reduction factor* of τ. Under the assumption $h(j(\tau))/\log(\log(n)) \to \infty$ for $n = \#\operatorname{Cl}(\mathcal{O})$ (cf. Theorem 6.3.4) we have $d^{-1}r_A(\tau) \to r(C)$. Experimentally obtained practical reduction factors, plotted against both the class number n and $\log(|H_\tau[j]|_\infty)/\log(\log(n))$, can be seen in Figure 6.1. To visualize the role of the class number and the hypothesis $h(j(\tau))/\log(\log(n)) \to \infty$, the points of higher class number are given a darker color in the second figure.

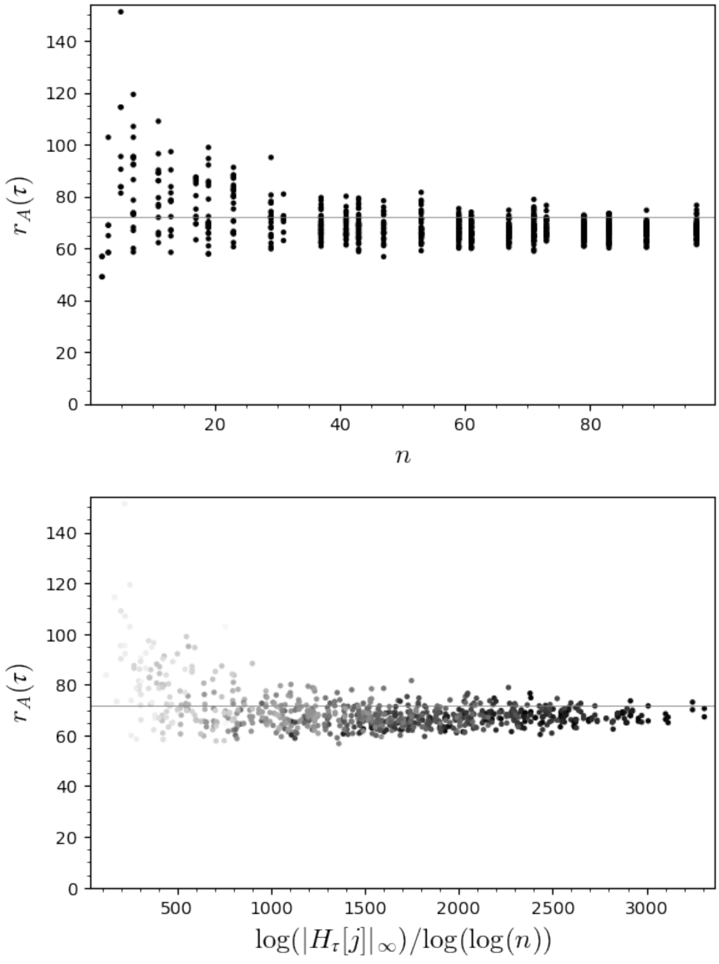


Figure 6.1: Practical reduction factors for $H_\tau[X_+^0(119)]$ for fundamental discriminants D with $\gcd(D, N) = 1$ and prime class number $n < 100$.

The values of the practical reduction factor $r_A(\tau)$ seem to be around their expected asymptotic value $r(C) = 72$ (represented by the horizontal grey line), though the convergence is not apparent; especially compared to, e.g. some classical class polynomials [8, Fig. 1]. However, in practical applications (see Section 6.5), the class numbers employed are typically several orders of magnitude higher (cf. e.g. [35]), so here we expect the speed of convergence not to cause major deviations in expected running times (cf. Section 6.6). For small class numbers, one can in practice even take advantage of this phenomenon by constructing generalized class polynomial with surprisingly good practical reduction factors by selecting a basis of $\mathcal{L}(\infty\mathcal{D})$ different from $1, x, y, x^2, xy, \ldots$ (see Example 6.7.4).

6.4.6 Comparison with existing class invariants

Real class invariants typically arise subject to congruence conditions on the discriminant. For example, Weber's functions with reduction factor 72 are not known to give class invariants for discriminants $\equiv 5 \pmod 8$. The reduction factors obtained by class invariants coming from the family of (double) eta quotients $\mathfrak{w}_n$ and $\mathfrak{w}_{p,q}$ (such as the Weber function $\mathfrak{w}_2$, as well as the function $\mathfrak{w}_{7,17}$ of Example 6.2.5) have been extensively studied; cf. most notably [9]. These modular functions are not known to yield class invariants if D is not a square modulo $4n$ or $4pq$. Hence, to the best of our knowledge, they also are not applicable to discriminants $\equiv 5 \pmod 8$ as soon as n, p or q is even. Excluding these cases, the (double) eta quotient with highest known reduction factor is $\mathfrak{w}_9$, with a reduction factor of 36 [9, Table 7.1].

A less-studied generalization are *multiple eta quotients* [12], which are quotients of products of 2^k eta functions. As far as we know these do not yield reduction factors better than 36 for $k > 1$.

The only other known family of "good" class invariants (in the sense that they have large reduction factors) are the Atkin functions A_p for prime numbers p, defined to be the smallest-degree functions in $\mathcal{L}(\infty\mathcal{D})$, where $\mathcal{D}$ is the unique cusp of $X^0_+(p)$. The "best" known one here is A_{71}, again with a reduction factor of 36, owing to the fact that $X^0_+(71)$ has genus zero [14, §3]).

The curve $C = X^0_+(119)$ has a reduction factor $r(C) = 72$ and yields real class invariants whenever D is a square modulo $4 \cdot 7 \cdot 17$ and not divisible by 7^2 or 17^2. The set of such D has density $> 28.98\%$ among the set of all negative discriminants (by Lemma 6.4.5). Out of these discriminants, one-fourth are $\equiv 5 \pmod 8$. Hence, for at least $28.98\% \cdot \frac{1}{4} > 7.24\%$ of imaginary quadratic discriminants, the reduction factor exceeds the previously best known reduction factors by a factor of at least two.

Remark 6.4.7 One should note that the above comparison does not take into account the discussion of Remark 6.3.1. Most importantly, the reduction factor is *not* synonymous with the true size reduction of the class polynomials. Indeed, as noted in that remark, the record-breaking CM construction [35] uses the Atkin invariant A_{71} of reduction factor 36, because the effective size reduction of class polynomials is by a factor of roughly $2^2 \cdot 36 = 144$ for certain discriminants. However, by Section 6.4.4, the same trick applies to generalized class polynomials, leading for $X^0_+(119)$ to a size reduction of $2^2 \cdot 72 = 288$, again for a positive density subset of discriminants. In Figure 6.2 we plot the practical reductions in bit size we found compared to the Hilbert class polynomial using this trick. $\Diamond$

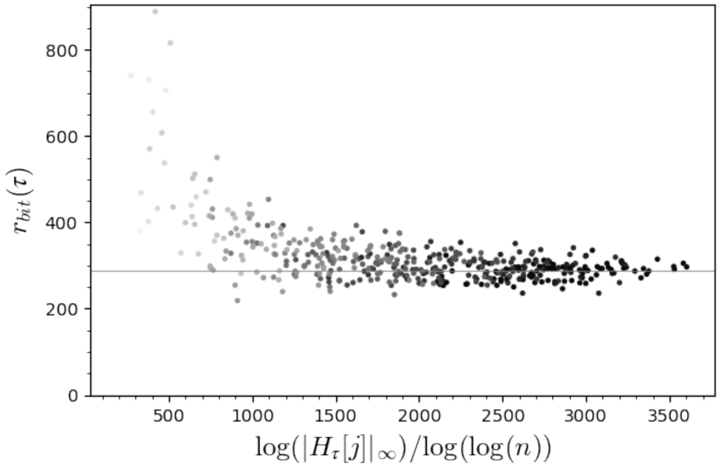


Figure 6.2: Bit-length reduction for $H_\tau[X_+^0(119)]$ for discriminants $D \equiv 0 \pmod{119}$ of class number $n < 100$.

Remark 6.4.8 Note that the "classical" class polynomial $H_\tau[x]$, arising from the function x on $X_+^0(119)$ by itself attains a reduction factor of 36 for the same 28.98% of discriminants. This beats all previously-known class invariants for a smaller subset ($\approx$ 1.2%) of discriminants: those that additionally are non-square modulo both 3 and 71. This x can be viewed as a generalisation of the Atkin functions to non-prime levels: it is the function of minimal degree in $\mathcal{L}(\infty\mathcal{D})$ for one of the cusps $\mathcal{D}$ of $X_+^0(119)$.

Similarly, the degree-two map of the hyperelliptic curve $X_+^0(191)$ (not to be confused with 119) has reduction factor 48, as observed by David Kohel in the AGC^{2}T 2021 Zulip group chat. This beats the reduction factor 32 of the Atkin function A_{191} of degree 3 on the same curve (see Example 6.7.2).

This shows that the search for generalized class invariants can even uncover new "classical" class invariants. $\Diamond$

6.4.7 More modular curves of genus one

We searched for more elliptic curves that could be used, and the results are in Tables 6.2, 6.3, and 6.4. In our search, we used the fact that $X_0(N)$ is well-studied and that there is an isomorphism $X_0(N) \to X^0(N) : z \mapsto Nz$. Surpisingly, we found lots of elliptic curves with reduction factor 72 and no elliptic curves with a greater reduction factor.

N		$g(X)$	$r(X)$	$\deg(\phi)$	$g(C)$	$r(C)$
119	$= 7 \cdot 17$	1	72	1	1	72
120	$= 2^3 \cdot 3 \cdot 5$	7	144	2	1	72
144	$= 2^4 \cdot 3^2$	5	144	2	1	72
176	$= 2^4 \cdot 11$	7	144	2	1	72
188	$= 2^2 \cdot 47$	9	144	2	1	72
131	$= 131$	1	66	1	1	66
75	$= 3 \cdot 5^2$	1	60	1	1	60
95	$= 5 \cdot 19$	1	60	1	1	60
171	$= 3^2 \cdot 19$	5	120	2	1	60
54	$= 2 \cdot 3^3$	1	54	1	1	54
81	$= 3^4$	1	54	1	1	54
90	$= 2 \cdot 3^2 \cdot 5$	4	108	2	1	54
108	$= 2^2 \cdot 3^3$	4	108	2	1	54
110	$= 2 \cdot 5 \cdot 11$	5	108	2	1	54
135	$= 3^3 \cdot 5$	4	108	2	1	54
136	$= 2^3 \cdot 17$	6	108	2	1	54
142	$= 2 \cdot 71$	8	108	2	1	54
159	$= 3 \cdot 53$	4	108	2	1	54
101	$= 101$	1	51	1	1	51
48	$= 2^4 \cdot 3$	1	48	1	1	48
56	$= 2^3 \cdot 7$	1	48	1	1	48
63	$= 3^2 \cdot 7$	1	48	1	1	48
64	$= 2^6$	1	48	1	1	48
84	$= 2^2 \cdot 3 \cdot 7$	4	96	2	1	48
96	$= 2^5 \cdot 3$	3	96	2	1	48
105	$= 3 \cdot 5 \cdot 7$	5	96	2	1	48
124	$= 2^2 \cdot 31$	6	96	2	1	48
128	$= 2^7$	3	96	2	1	48
141	$= 3 \cdot 47$	6	96	2	1	48
155	$= 5 \cdot 31$	4	96	2	1	48
191	$= 191$	2	96	2	0	48

Table 6.2: The curves $X = X^0_+(N)$ for which there exists a map $\phi : X \to C$ of degree ≤ 2 with $g(C) \leq 1$ and $r(C) \geq 48$. We used Furumoto-Hasegawa [15] and Jeon [20] to get a complete list.

In Section 6.7, we will allow curves of higher genus, which do achieve arbitrarily high values of $r(C)$. Moreover, our search is by no means exhaustive, as Tables 6.2 and 6.3 restrict to maps $\phi : X \to C$ of degree ≤ 2 and Table 6.4 only looks at one curve $X = X^0(N)$ per isomorphism class of curves C. For example, the curve $C = X^0_+(119)$ has $r(C) = 72$. However, in the Cremona database, it is listed as 17a4, and comes with a modular parametrization $\phi_{17} : X_0(17) \to C$ of degree 1, which has $r(\phi_{17}) = 18$. This is why C does not appear in Table 6.4.

N		$g(X)$	$r(X)$	$\deg(\phi)$	$g(C)$	$r(C)$
36	$= 2^2 \cdot 3^2$	1	72	1	1	72
60	$= 2^2 \cdot 3 \cdot 5$	7	144	2	1	72
72	$= 2^3 \cdot 3^2$	5	144	2	1	72
92	$= 2^2 \cdot 23$	10	144	2	1	72
94	$= 2 \cdot 47$	11	144	2	1	72
49	$= 7^2$	1	56	1	1	56
24	$= 2^3 \cdot 3$	1	48	1	1	48
32	$= 2^5$	1	48	1	1	48
42	$= 2 \cdot 3 \cdot 7$	5	96	2	1	48
48	$= 2^4 \cdot 3$	3	96	2	0	48
62	$= 2 \cdot 31$	7	96	2	1	48
69	$= 3 \cdot 23$	7	96	2	1	48

Table 6.3: The curves $X = X^0(N)$ for which there exists a map $\phi : X \to C$ of degree ≤ 2 with $g(C) \leq 1$ and $r(C) \geq 48$ and N is not already in Table 6.2. We used Ogg [28] and Bars [1] to get a complete list.

Finally, the tables are restricted to quotients of $X^0(N)$. Letting go of $X^0(N)$, we find that the genus-one modular curves $7C^1$, $8K^1$, $9H^1$, $12V^1$, $15I^1 = X_1(15)$, $16M^1$, $24J^1$, $27C^1$, $32E^1$ in the Pauli-Cummins database [6] all achieve $r(C) \in \{84, 96, 108\}$. We have not pursued these curves yet, as Proposition 6.4.1 does not apply to them.

6.5 Application: the CM method

Class polynomials are used in the *CM method* for constructing elliptic curves over finite fields with a specified characteristic polynomial of Frobenius.

The input to the CM method is a monic quadratic polynomial $P = x^2 - tx + q \in \mathbf{Z}[x]$, where q is a prime power coprime to t, and the discriminant $d = t^2 - 4q$ is negative. The output is an elliptic curve $E/\mathbf{F}_q$ with $q + 1 - t$ rational points, which has P as its characteristic polynomial of Frobenius.

The algorithm of the classical CM method (without using class invariants for now) is as follows. Let $K = \mathbf{Q}(\sqrt{d})$.

1. Compute the Hilbert class polynomial H_K of $\mathcal{O}_K$.
2. Find a root $j_0 \in \mathbf{F}_q$ of H_K (which is known to split into linear factors in $\mathbf{F}_q$).
3. Construct an elliptic curve $E/\mathbf{F}_q$ with $j(E) = j_0$. Compute all twists of E and return the one with $q + 1 - t$ rational points.

In practice, one can discard the curves for which $(q+1-t)Q \neq O$ for some random point Q, although there are also more straightforward methods to select the correct twist [29].

E	N	$r(X)$	$\deg(\phi)$	$\mathrm{rank}(E)$	$r(C)$
36a1	$2^2 \cdot 3^2$	72	1	0	72
92a1	$2^2 \cdot 23$	144	2	0	72
94a1	$2 \cdot 47$	144	2	0	72
144a1	$2^4 \cdot 3^2$	288	4	0	72
368e1	$2^4 \cdot 23$	576	8	1	72
558a1	$2 \cdot 3^2 \cdot 31$	1152	16	1	72
704a1	$2^6 \cdot 11$	1152	16	1	72
704k1	$2^6 \cdot 11$	1152	16	1	72
1728a1	$2^6 \cdot 3^3$	3456	48	1	72
1728v1	$2^6 \cdot 3^3$	3456	48	1	72
3456a1	$2^7 \cdot 3^3$	6912	96	1	72
3456e1	$2^7 \cdot 3^3$	6912	96	0	72
131a1	131	132	2	1	66
575a1	$5^2 \cdot 23$	720	12	1	60
711a1	$3^2 \cdot 79$	960	16	1	60
755b1	$5 \cdot 151$	912	16	1	57
999b1	$3^3 \cdot 37$	1368	24	1	57
49a1	7^2	56	1	0	56
1323m1	$3^3 \cdot 7^2$	2016	36	1	56
243a1	3^5	324	6	1	54
405c1	$3^4 \cdot 5$	648	12	1	54
459a1	$3^3 \cdot 17$	648	12	1	54
101a1	101	102	2	1	51
335a1	$5 \cdot 67$	408	8	1	51
591a1	$3 \cdot 197$	792	16	1	99/2
485b1	$5 \cdot 97$	588	12	1	49
723b1	$3 \cdot 241$	968	20	1	242/5
69a1	$3 \cdot 23$	96	2	0	48
105a1	$3 \cdot 5 \cdot 7$	192	4	0	48
141d1	$3 \cdot 47$	192	4	1	48
155c1	$5 \cdot 31$	192	4	1	48
213a1	$3 \cdot 71$	288	6	0	48

Table 6.4: The elliptic curves $E/\mathbf{Q}$ of conductor $<$ 500.000 such that the modular parametrization $\phi : X \to E$ according to the LMFDB [23, 5, 36] gives $r(C) \geq 66$ or gives $r(C) \geq 48$ and odd N.

As the degree and height of the Hilbert class polynomial grow quickly with the absolute value of the discriminant Δ_K of K, the CM method is only feasible for small values of $|\Delta_K|$. The record computation of [35] uses class invariants, specifically arising from the Atkin function A_{71}. Combined with the tricks listed in Remark 6.3.1 this allows to handle a case where $|\Delta_K| > 10^{16}$.

We will now describe how to apply the CM method using generalized class polynomials. Hence let C be an elliptic modular curve. Since we are working with alternative class invariants instead of the usual j-invariant, we will relate the two using *modular polynomials* as follows.

Lemma 6.5.1 *Let $d_j := [\mathbf{Q}(C, j) : \mathbf{Q}(C)]$. Then there exists a polynomial $\Psi_C = \sum_{i=0}^{d_j} f_i Z^i \in \mathbf{Z}[X, Y][Z]$ of degree d_j in Z such that*

(i) $\Psi_C(j) = 0$;

(ii) $\deg_Y(f_i) \leq 1$ *for each* i;

(iii) the coefficients (in $\mathbf{Z}$*) of* Ψ_C *viewed as an element of* $\mathbf{Z}[X, Y, Z]$ *are coprime;*

(iv) viewed as elements of $\mathbf{Q}(C)$, *the* f_i *have at most one common zero in* $C(\overline{\mathbf{Q}})$.

Furthermore, Ψ_C *is unique up to sign.*

Proof. Consider the minimal polynomial $\Psi_C^0 = \sum_{i=0}^{d_j} g_i Z^i \in \mathbf{Q}(C)[Z]$ of j over $\mathbf{Q}(C)$. Let

$$\mathcal{E} := \sum_{P \in C \setminus \{O\}} \min_i(\mathrm{ord}_P(g_i))(P).$$

Then $\mathcal{E} - \left(\sum_{P \in C} \mathrm{ord}_P(\mathcal{E})P\right) - (\deg(\mathcal{E}) - 1)(O)$ is a $\mathbf{Q}$-rational principal divisor. There is a unique function g up to $\mathbf{Q}^\times$-scaling such that $\mathrm{div}(g) = \mathcal{E}$. Dividing each g_i by g gives $g_i \in \mathcal{L}(\infty(O)) = \mathbf{Q}[x, y]$ satisfying ((iv)) and unique up to $\mathbf{Q}^\times$. Now take representatives f_i satisfying ((ii)) and scale to get ((iii)), which makes Ψ_C unique up to sign. □

For each curve C with which we would like to apply the generalized CM method, the polynomial $\Psi_C \in \mathbf{Z}[X, Y, Z]$ can be precomputed and stored. Next we need a criterion for which discriminants D yields class invariants. For example, if $C = X_+^0(N)$ then this is given by Proposition 6.4.1. Now, given a desired characteristic polynomial of Frobenius $x^2 - tx + q$ such that $D = t^2 - 4q$ satisfies this criterion, we have the following algorithm for sufficiently large $|D|$.

(1) Compute a generalized class function F of discriminant D as well as its Heegner point Q.

(2a) Find a zero $P = (x, y) \in C(\mathbf{F}_q)$ of F that is neither $-Q$ nor a common root of the polynomials $f_1, \ldots, f_{d_j}$ of Lemma 6.5.1.

(2b) Find all roots $j_0 \in \mathbf{F}_q$ of the polynomial $\Psi_C(x, y, Z) \in \mathbf{F}_q[Z]$.

(3) For each j_0, construct an elliptic curve $E/\mathbf{F}_q$ with $j(E) = j_0$ and all of its twists up to isomorphism over $\mathbf{F}_q$. Return one with $q + 1 - t$ rational points.

The main advantage compared to the classical CM method, both in terms of memory and speed, is expected to be in the (dominant) first step (1) (see Section 6.6). Out of the computationally non-dominant steps, only (2a) is less straightforward. One way to proceed would be as follows.

(i) Compute $F_x := N_{\mathbf{F}_q(C)/\mathbf{F}_q(x)}(F)$.

(ii) Find a root $x \in \mathbf{F}_q$ of F_x.

(iii) Solve for the corresponding value of y using the linear polynomial $H_\tau[C](x, Y)$, or continue with both solutions y coming from the Weierstrass equation.

Remark 6.5.2 The polynomial F_x is very close to the classical class polynomial $H_\tau[x]$; indeed, it has the same roots, together with one additional root at the x-coordinate of the Heegner point of F. The norm computation in step ((i)) is however computationally asymptotically dominated by the computation of F. ♢

6.6 The computational benefits of our invariants

6.6.1 Space complexity of the functions

The advantage of using generalized class functions lies in their size. This already gives a serious advantage when storing one or more class polynomials for later use, e.g. for various values of q in the CM method. Additionally, one would expect the smaller size to make the generalized class polynomials less expensive to compute. Again for C a modular elliptic curve with a given Weierstrass model, we present a preliminary analysis of the cost of computing a generalized class polynomial $H_\tau[C]$ when compared to the "classical" class polynomial $H_\tau[x]$ (though recall that the latter already dominates all previously-known class invariants along a positive density subset of discriminants for $C = X_+^0(119)$, cf. Section 6.4.6).

6.6.2 Speed of complex analytic computation

We now explain how to adapt the complex analytic approximation algorithm to generalized class polynomials.

To compute the classical class polynomial $H_\tau[x]$ one first evaluates $x(\tau)$ and all its conjugates, which are of the form $x_i(\tau_i)$, where x_i and τ_i can be obtained using Shimura's reciprocity law [18] or N-systems [31]. Then one multiplies the linear polynomials $X - x_i(\tau_i)$ together in a binary tree using fast multiplication algorithms.

As $H_\tau[C]$ has roughly half the height, we only need half the precision at each step. This gives a great speed-up when evaluating $x_i(\tau_i)$, but then we also need to compute $y_i(\tau_i)$. Fortunately that should only take a fraction of the time required for computing

$x_i(\tau_i)$, as we can first compute it to low precision and then obtain as many digits as desired quickly using

$$y = \frac{-g(x) + \sqrt{g(x)^2 + 4f(x)}}{2}$$

for $C : y^2 + g(x)y = f(x)$.

The **binary tree** step is harder to analyze. Instead of having polynomials $A(X) = \prod_{i \in S}(X - x_i(\tau))$ to multiply for various subsets $S \subset \{1, 2, \ldots, n\}$, we will have pairs (F, Q) with $F = A(X) + B(X)Y$ and

$$\mathrm{div}(F) = \sum_{i \in S}(P_i) + (Q) - (\#S + 1)\mathcal{D}.$$

Instead of a single multiplication A_1A_2 to go from S_1 and S_2 to $S_3 = S_1 \sqcup S_2$, we now need to compute the point $Q_3 = Q_1 + Q_2$ (with the elliptic curve group law) and a function F_3 with

$$\mathrm{div}(F_3) = \sum_{i \in S}(P_i) + (Q_3) - (\#S_3 + 1)\mathcal{D} = \mathrm{div}(F_1) + \mathrm{div}(F_2) + (Q_3) + (O) - (Q_1) - (Q_2).$$

The following formula can be used:

$$F_3 = \frac{F_1\ F_2\ R \mod (Y^2 - f(X))}{(X - x(Q_1))(X - x(Q_2))}, \quad \text{where} \tag{6.13}$$

$$R = (x(Q_1) - x(Q_2))\ Y\ +\ (y(-Q_2) - y(-Q_1))\ X \tag{6.14}$$

$$+\, x(Q_2)y(-Q_1) - x(Q_1)y(-Q_2), \tag{6.15}$$

and where the reduction modulo $Y^2 - f(X)$ keeps the outcome of degree ≤ 1 in Y.

We can multiply F_1 with F_2 using three multiplications of half the degree, by the same trick that is used in Karatsuba multiplication. Indeed, let

$$C = A_1A_2, \quad D = B_1B_2, \quad \text{and} \quad E = (A_1 + B_2)(A_2 + B_1)$$

to get $F_1F_2 = (C + Df) + (E - C - D)Y$. So computing F_3 involves three polynomial multiplications of half the degree of F_1 and F_2, as well as various multiplications and long divisions by fixed-degree polynomials and various additions and subtractions. The most serious computations in the binary tree are now done with half the degree *and* half the number of digits, but three times as often, which takes 3/16th of the time with naive multiplication and still less than 3/4 of the time with quasi-linear-time multiplication. The impact of the extra additions and subtractions, as well as the extra multiplications by a linear polynomial in X and Y and long division by the denominator of (6.13) requires further analysis, but we expect this to be minor. Regardless, for large discriminants, the main bottleneck is in memory complexity (as noted in [7, Section 7]), and here we obtain an improvement of a factor of 1/2 when passing from $H_\tau[x]$ to $H_\tau[C]$.

6.6.3 Adapting the CRT method

Overview of CRT class polynomial computation

We now heuristically estimate the expected speed-up when computing $H_\tau[C]$ instead of $H_\tau[x]$ using the (currently state-of-the-art) CRT method for class polynomial computation [14, 34, 35]. We restrict to the case of C such that all q-expansion coefficients of x and y are rational, and will analyse some steps only in the main case where C is a quotient of $X_+^0(N)$. To keep our exposition simple, we will not treat the main improvement of [35], even though we do expect it to combine well with our generalized class invariants. We plan to give a more detailed account and an implementation in future work.

For the CM method, it is more efficient to directly compute class polynomials modulo q using the *online CRT* as in [34, Section 2]. In other words, we never write down $H_\tau[C] \in \mathbf{Z}[X,Y]$, but instead compute $(H_\tau[C] \bmod q) \in \mathbf{F}_q[X,Y]$ directly from $(H_\tau[C] \bmod p)$ for p in a set S of small primes. The space complexity of the CM method is then $n\log(q)$, which is independent of our choice of class function. The set S must be chosen in such a way that $\prod_{p\in S} p$ is larger than 4 times the largest coefficient.

By cutting the number of digits in half when switching from x to C, we essentially cut $\#S$ in half. If the amount of work that we do for each prime p does not grow too much, then our class function $H_\tau[C]$ yields a speed-up over the classical class polynomial $H_\tau[x]$.

What needs to be done for each p is the following.

1. Enumerate all E'' with endomorphism ring $\mathcal{O}$ and compute the appropriate points in $C(\mathbf{F}_p)$.
2. Compute $(F \bmod p)$ by putting together the information from Step 1.

In practice, for "typical" discriminants D with 9 to 14 digits, Sutherland [34, Sections 8.3 and 8.4] finds that performing Steps 1 and 2 together $\#S$ times is the dominant part of the CRT method.

We will now argue why we expect each of these steps to take (much) less than twice as long with the generalized class polynomial for suitable C. Together with the fact that our set S is only half the original size due to the reduction factor, this means that computing $H_\tau[C]$ takes less time than computing $H_\tau[x]$.

Enumerating via the Fricke involution

Step 1 is already very subtle in the case of a single class invariant f. Indeed, there could be multiple Galois orbits of values $f(\tau)$ for the same order $\mathcal{O}$, and hence multiple irreducible class polynomials $H_{\tau_i}[f] \in K[X]$. In the CRT method, one has to make sure to compute the polynomials $(H_{\tau_i}[f] \bmod p)_p$ for the same value of i, and only for τ_i for which this is a class invariant. This issue is addressed in detail in [14, Section 4].

We will first explain how to adapt one solution to our main case of quotients C of $X_+^0(N)$ where N is coprime to the conductor of $\mathcal{O}$ and $D = \operatorname{disc}(\mathcal{O})$ is a square modulo $4N$.

We adapt the method of Section 4.3 of [14] as follows. We have $\mathbf{Q}(X^0(N)) = \mathbf{Q}(j, j_N)$, where $j_N(z) = j(z/N) = j(W_N z)$ for the Fricke-Atkin-Lehner involution $W_N : z \mapsto -N/z$ (this follows for example from [33, Proposition 6.9]). In particular, every function $f \in \mathbf{Q}(C)$ for a quotient C of $X^0(N)$ can be expressed as a rational function in j and j_N. In practice, these expressions can be quite large, but (analogously to [14, Lemma 2]) we can also obtain the value $f(z)$ as a root of $\gcd(\Psi_f(X, j(z)), \Psi_{f\circ W_N}(X, j_N(z)))$ instead.

In the particular case where C is a quotient of $X^0_+(N)$, we even have $\mathbf{Q}(C) \subset \mathbf{Q}(X^0_+(N)) = \mathbf{Q}(j + j_N, j \cdot j_N)$, and we can use Ψ_f instead of $\Psi_{f\circ W_N}$.

So instead of enumerating just the j-values, we wish to link them with the corresponding j_N-values, and we do that as follows.

Suppose that N is coprime to the conductor of $\mathcal{O}$ and that D is a square modulo $4N$. Then by Lemma 6.4.4 we get $a, b, c \in \mathbf{Z}$ with $a, c > 0$, $b^2 - 4ac = D$, $N \mid c$, and $\gcd(ac/N, N) = \gcd(a, b, c) = 1$. In line with Lemma 2 of [14] we could even take $c = N$ by replacing a by ac/N. We take $z = \frac{-b+\sqrt{D}}{2a}$, $\mathfrak{n} = a\overline{z}\mathbf{Z} + N\mathbf{Z}$, and $\mathfrak{a} = z\mathbf{Z} + \mathbf{Z}$. Then we have $\mathcal{O} = az\mathbf{Z} + \mathbf{Z}$, and we find that $\mathfrak{n}$ is an invertible $\mathcal{O}$-ideal with $\mathcal{O}/\mathfrak{n} \cong \mathbf{Z}/N\mathbf{Z}$. In fact, we find $\overline{\mathfrak{n}}\mathfrak{a} = z\mathbf{Z} + N\mathbf{Z}$ and hence

$$\sigma_{[\mathfrak{n}]} j(z) = j(\mathfrak{n}^{-1}\mathfrak{a}) = j(\overline{\mathfrak{n}}\mathfrak{a}) = j_N(z).$$

Exactly as in Section 4.3 of [14], we list the j-values of elliptic curves over $\mathbf{F}_p$ with endomorphism ring $\mathcal{O}$, and arrange them into unoriented $[\mathfrak{n}]$-isogeny cycles. If C is a quotient of $X^0_+(N)$ over $\mathbf{Q}$, then for each edge of this graph, we find the f-value from the two j-values of the end points. (In the case where the $[\mathfrak{n}]$-isogeny cycles are 2-cycles, we only get one f-value per 2-cycle and we get a lower-degree class polynomial $H_\tau[f]$.)

In practice, we could do this for $f = x$ exactly as in [14], and then solve for y using $\Psi_C(x, y, j) = 0$, which is linear in y. The only additional work compared to what is done in [14] is computing and solving the linear equation to get y, which is much faster than all the other steps.

In particular, Step 1 takes much less than twice as long with C than with x, while we need to do it only half as often, which leads to a speed-up. Further research into these modular polynomials is needed in order to determine the exact gain.

To also make this work for quotients of $X^0(N)$ that are not quotients of $X^0_+(N)$, one would need to compute oriented $[\mathfrak{n}]$-isogeny cycles.

Other tricks for enumerating

The methods from [14, Sections 4.1 and 4.2] also seem amenable.

The main computational tool at the beginning of Section 4.1 is the modular polynomial $\Phi_{\ell,f}$, which we generalize from f to C as follows.

Let $\Phi_{\ell,C}$ be a Gröbner basis of the ideal in $\mathbf{Q}[X_1, Y_1, X_2, Y_2]$ of polynomials that vanish on $\{(\psi(z), \psi(\ell z)) : z \in \mathbf{H}\}$, with respect to the lexicographic ordering with $X_1 > Y_1 > X_2 > Y_2$. To get from $\psi(z)$ to all possible values of $\psi(\ell z)$, one substitutes $\psi(z)$ for (X_1, Y_1), and then solves first for X_2 and then for Y_2. For each C and ℓ this works for all but a finite set of primes p. Such multivariate modular polynomials would need to be precomputed. One possible starting point for computing these would be [25, 26], which compute multivariate (Hilbert) modular polynomials, each with a different method. For yet another approach to computing modular polynomials, see [3].

We expect the reduction factor to also give a reduction of the size of these multivariate modular polynomials, but on the other hand, we need two of them: one to solve for x of an isogenous curve, and one to evaluate in x and get y. As evaluating is faster than solving, we expect the use of these modular polynomials to take much less than twice as long (and we need to do it only half as often, because we have half as many primes).

The 'Trace Trick' of [14, Section 4.2] enables the use of the Weber function $\mathfrak{f}$ in the CRT method. In case we would also need this trick, for some more exotic curves C, we could consider applying it with arbitrary functions $f \in \mathbf{Q}(C)$ such as $f = ax + by$ for small integers a and b. In loc. cit. the relevant trace is computed with much fewer primes, so it is ok to apply this with the lower reduction factor of f.

We did not yet consider the general algorithm of [14, Section 4.4]. It is the method that works for all class invariants, but is only practical under additional restrictions. We do not have examples of generalized class invariants where this trick is needed. The challenging step to generalize is factoring a large-degree function in $\mathbf{Q}(C)$ in order to obtain the small class functions.

Constructing a function from its roots

In the CRT setting the multiplications and long-divisions by small-degree polynomials of Section 6.6.2 only take time $O(nM(\log(p)))$ per level, which is asymptotically dominated by the $O(M(n\log(p))$ time of the multiplications of large-degree polynomials. Therefore, Step 2 seems to take about 1.5 times as long per prime p for $H_\mathcal{T}[C]$ when compared to $H_\mathcal{T}[x]$.

The total running time

Concluding this preliminary analysis, we estimate the cost of computing $H_\mathcal{T}[C]$ to be significantly lower compared to $H_\mathcal{T}[x]$, though further research, in particular into (the implementation of) modular polynomials for C is required to determine the exact gain. This is beyond the scope of the current paper, which focuses on introducing the generalized class functions and their height reduction. We plan to give a more detailed account and an implementation in future work.

6.7 General curves and bases

Now suppose that our modular curve C is not necessarily an elliptic curve. Let $\mathcal{D}$ be an effective divisor over $\mathbf{Q}$ on C and let $\mathcal{B} = \{b_0, b_1, \ldots\}$ be a $\mathbf{Q}$-basis of $\mathcal{L}(\infty\mathcal{D})$ ordered by ascending degree.

The classical case is the case where we have one modular function f and we take $C = \mathbf{P}^1$, $\psi = f = (f : 1)$, $\mathcal{D} = ((1 : 0)) = (\infty)$, and $\mathcal{B} = \{1, f, f^2, \ldots\}$. The case of all previous sections of this paper is the case where C is an elliptic curve given by a Weierstrass equation, $\mathcal{D} = ((0:1:0))$, and $\mathcal{B} = \{1, x, y, x^2, xy, x^3, x^2y, \ldots\}$.

Example 6.7.1 One systematic way to choose a $\mathbf{Q}$-basis of $\mathcal{L}(\infty\mathcal{D})$ is as follows. First choose $x \in \mathcal{L}(\infty\mathcal{D}) \setminus \mathbf{Q}$ of some degree d. (For example, one can take $x = f$ with $d = 1$ in the classical case, and $x = x$ with $d = 2$ in the elliptic case.) Now, let $y_0 = 1$ and choose y_j for $j = 1, 2, \ldots, d-1$ in such a way that

$$y_j \in \mathcal{L}(m_j\mathcal{D}) \setminus \langle y_k x^e : k < j, e \in \mathbf{Z}\rangle,$$

where m_j is minimal such that this set is non-empty. This way we obtain a vector $\vec{y} = (y_0, \ldots, y_{d-1})$ of d functions. (For example, in the classical case we have $\vec{y} = 1$, and in the elliptic case we chose $\vec{y} = (1, y)$.) Then $\mathcal{B} = \{x^e y_j : e \in \mathbf{Z}_{\geq 0}, j \in \{0, 1, 2, \ldots, d-1\}\}$ is a basis of $\mathcal{L}(\infty\mathcal{D})$. We order this basis by ascending degree $de + m_j$, and if two elements have the same degree, then we put the one with lowest j first. ☆

Example 6.7.2 Consider the modular curve $X_+^0(191)$ (not to be confused with 119), which is hyperelliptic with model $t^2 = s^6 + 2s^4 + 2s^3 + 5s^2 - 6s + 1$ [16, Table 3], and the unique cusp is at $\mathcal{D} = ((1:1:0))$. One of the possible bases of $\mathcal{L}(\infty\mathcal{D})$ obtained by the recipe above is $\mathcal{B} = \{1, x, y_1, y_2, x^2, x^2y_1, x^2y_2, \ldots\}$, where $x = (t + s^3 + s + 1)/2$, $y_1 = sx$, and $y_2 = s(y_1 + 1)$. The degrees of these functions are respectively 3, 5, and 7.

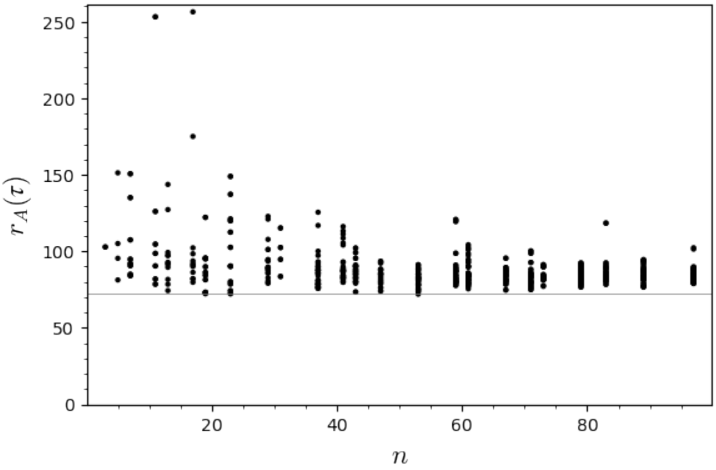
The function x is, up to multiplicative and additive constants, equal to the Atkin function A_{191}. The reduction factors are $r(C) = 96$, $r(s) = 48$, and $r(A_{191}) = 32$. ☆

As in Section 6.2, let $\tau \in \mathbf{H}$ imaginary quadratic and assume that (b_i, τ) is a class invariant for every $b_i \in \mathcal{B}$. Then, again unique up to scaling, we obtain a non-zero function $F_\tau[C, \mathcal{B}] = \sum_{i=0}^{k} a_i b_i \in K(C)$ $(a_i \in K)$ with k minimal such that $\sum_{i=0}^{k} a_i b_i(\tau) = 0$.

Definition 6.7.3 We call this $F_\tau[C, \mathcal{B}]$ the *generalized class function* for the triple $C, \mathcal{B}, \tau$. If $\mathcal{B}$ is as in Example 6.7.1 then we again refer to the associated polynomial $H_\tau[C, \mathcal{B}] \in K[X, Y_1, \ldots, Y_d]$ (of total degree ≤ 1 in $Y_1, \ldots, Y_d$ and such that $H_\tau[C, \mathcal{B}](x, y_1, \ldots, y_d) = F_\tau[C, \mathcal{B}]$) as the *generalized class polynomial.* △

Example 6.7.4 It turns out that, already for the case of elliptic curves, allowing the freedom of the choice of basis of may in reality lead to potentially better practical reduction factors. Revisiting our main example $C := X_+^0(119)$, denote by $w := \mathfrak{w}_{7,17}$ the function (6.4) and by $z := x + y$ the sum of the Weierstrass coordinates for the model (6.3). Now consider the basis $\mathcal{B} := \{1, x, z, w, xz, wx, wz, w^2, wxz, w^2x, \ldots\}$ of $\mathcal{L}(\infty\mathcal{D})$. The resulting generalized class polynomials corresponding to the discriminants of Table 6.1 are listed in Table 6.5. We get practical reduction factors in Figure 6.3 that are better than those in Figure 6.1.

A likely explanation for this improvement is that now not only the poles, but also the zeroes are as much restricted to the cusps of $X_+^0(119)$ as possible. Indeed, the points $O = (0 : 1 : 0)$ and $P = (0, 0)$ are the cusps, while $2P$ and $3P$ are rational CM points. Now $\mathrm{div}(w) = 4(P) - 4(O)$, $\mathrm{div}(x) = (P) + (3P) - 2(O)$, and $\mathrm{div}(y) = 2(P) + (2P) - 3(O)$. In particular, the function w is a modular unit. As explained in Remark 6.3.8, modular units in the classical setting give better practical reduction factors than non-units, even though the reduction factors are asymptotically the same. ☆



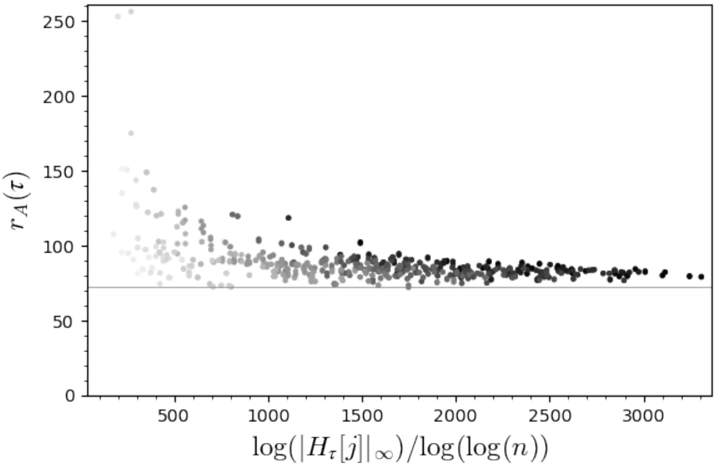


Figure 6.3: Practical reduction factors for $H_\tau[X_+^0(119), \mathcal{B}]$ for fundamental discriminants D with $\gcd(D, N) = 1$ and prime class number $n < 100$.

D	n	$F_\tau[C,\mathcal{B}]$
-52	2	$z-x+1$
-523	5	$xw-xz-x+3w+z$
-5347	13	$xw^3-10xw^2z+42xw^2+48w^3+13xwz+35w^2z+62xw$ $+104w^2+39xz+90wz-11x+41w+39z+1$
-15139	29	$xw^7-33xw^6z+5874xw^6+849w^7-2119xw^5z-3865w^6z$ $+31183xw^5-4249w^6+2200xw^4z-15449w^5z+36423xw^4$ $-29399w^5+6066xw^3z-46282w^4z+46223xw^3-27578w^4+6207xw^2z$ $-30128w^3z+31320xw^2-47581w^3+6757xwz-35595w^2z$ $+8017xw-17181w^2-742xz-10159wz-x-2797w+22z$

Table 6.5: Some conjecturally correct generalized class functions for the curve $C=X_+^0(119)$ using the $\mathcal{L}(\infty\mathcal{D})$-basis $\mathcal{B}:=\{1,x,z,w,xz,wx,wz,w^2,wxz,w^2x,\ldots\}$.

Theorem 6.7.5 *Let $C: y^2+g(x)y=f(x)$ with $f,g\in\mathbf{Q}[X]$ be a hyperelliptic curve such that $4f(x)+g(x)^2$ has odd degree and $\mathrm{Jac}(C)(\mathbf{Q})$ is finite. Set $\mathcal{D}$ to be the unique point at infinity and choose the basis $\mathcal{B}=\{1,x,x^2,y,xy,x^2y,\ldots\}$ of $\mathcal{L}(\infty\mathcal{D})$. Then Theorem 6.3.4 and Proposition 6.3.7 also hold for C and $H_\tau[C,\mathcal{B}]$.*

Proof. The original proof now goes through with only the following change. There are finitely many possibilities for the class c of the divisor $-\sum_\sigma((\sigma(\psi(\tau)))-\mathcal{D})$ by our assumption that $\mathrm{Jac}(C)(\mathbf{Q})$ is finite. For every c, choose a representative $\sum_{i=1}^m((P_i)-\mathcal{D})$ with m minimal and consider a primitive polynomial $T\in\mathbf{Z}[X]$ with roots $x(P_i)$ for $i=1,\ldots,m$. □

Remark 6.7.6 Our proofs of Theorems 6.3.4 and 6.7.5 heavily rely on the fact that Heegner points are torsion. To completely remove the assumption on ranks, one would therefore need to bound the Heegner points, even in the rank-one case. Moreover, the proofs rely on the hyperelliptic equation where we use that $|a|\leq|a+bi|$ for real numbers a and b. Though we expect an analogue of these results to hold for general modular curves, this would require additional ideas. Do note that such an analogue would yield arbitrarily high reduction factors for generalized class polynomials by (6.7). For example, for $C=X_+^0(239)$ of genus 3 we already obtain $r(C)=120$, exceeding the Bröker-Stevenhagen bound. ◊

6.8 Bibliography

[1] Francesc Bars. Bielliptic modular curves. *J. Number Theory*, 76(1):154–165, 1999.

[2] Brian J. Birch. Weber's class invariants. *Mathematika*, 16:283–294, 1969.

[3] Reinier Bröker, Kristin Lauter, and Andrew V. Sutherland. Modular polynomials via isogeny volcanoes. *Math. Comp.*, 81(278):1201–1231, 2012.

[4] Reinier Bröker and Peter Stevenhagen. Constructing elliptic curves of prime order. In *Computational arithmetic geometry*, volume 463 of *Contemp. Math.*, pages 17–28. Amer. Math. Soc., Providence, RI, 2008.

[5] John Cremona. *Algorithms for modular elliptic curves.* Cambridge University Press, Cambridge, 1992.

[6] Chris J. Cummins and Sebastian Pauli. Congruence subgroups of $\mathrm{PSL}(2,\mathbb{Z})$ of genus less than or equal to 24. *Experiment. Math.*, 12(2):243–255, 2003. Interactive database at `https://mathstats.uncg.edu/sites/pauli/congruence/csg1.html`.

[7] Andreas Enge. The complexity of class polynomial computation via floating point approximations. *Math. Comp.*, 78(266):1089–1107, 2009.

[8] Andreas Enge and François Morain. Comparing invariants for class fields of imaginary quadratric fields. In *Algorithmic number theory (Sydney, 2002)*, volume 2369 of *Lecture Notes in Comput. Sci.*, pages 252–266. Springer, Berlin, 2002.

[9] Andreas Enge and François Morain. Generalised Weber functions. *Acta Arith.*, 164(4):309–342, 2014.

[10] Andreas Enge and Reinhard Schertz. Constructing elliptic curves over finite fields using double eta-quotients. *Journal de Théorie des Nombres de Bordeaux*, 16:555–568, 2004.

[11] Andreas Enge and Reinhard Schertz. Modular curves of composite level. *Acta Arith.*, 118(2):129–141, 2005.

[12] Andreas Enge and Reinhard Schertz. Singular values of multiple eta-quotients for ramified primes. *LMS Journal of Computation and Mathematics*, 16:407–418, 2013.

[13] Andreas Enge and Marco Streng. Schertz style class invariants for genus two, 2016. preprint, arXiv:1610.04505.

[14] Andreas Enge and Andrew V. Sutherland. Class invariants by the CRT method. In *Algorithmic number theory*, volume 6197 of *Lecture Notes in Comput. Sci.*, pages 142–156. Springer, Berlin, 2010.

[15] Masahiro Furumoto and Yuji Hasegawa. Hyperelliptic quotients of modular curves $X_0(N)$. *Tokyo J. Math.*, 22(1):105–125, 1999.

[16] Steven Galbraith. *Equations for modular curves.* PhD thesis, St. Cross College, 1996. `https://www.math.auckland.ac.nz/~sgal018/thesis.pdf`.

[17] Alice Gee. Class invariants by Shimura's reciprocity law. volume 11, pages 45–72. 1999. Les XXèmes Journées Arithmétiques (Limoges, 1997).

[18] Alice Gee and Peter Stevenhagen. Generating class fields using Shimura reciprocity. In *Algorithmic number theory (Portland, OR, 1998)*, volume 1423 of *Lecture Notes in Comput. Sci.*, pages 441–453. Springer, Berlin, 1998.

[19] Marc Hindry and Joseph H. Silverman. *Diophantine geometry*, volume 201 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 2000. An introduction.

[20] Daeyeol Jeon. Bielliptic modular curves $X_0^+(N)$. *J. Number Theory*, 185:319–338, 2018.

[21] Henry H. Kim. Functoriality for the exterior square of GL_4 and the symmetric fourth of GL_2. *J. Amer. Math. Soc.*, 16(1):139–183, 2003. With appendix 1 by Dinakar Ramakrishnan and appendix 2 by Kim and Peter Sarnak.

[22] John E. Littlewood. On the class-number of the corpus $p(\sqrt{-k})$. *Proc. Lond. Math. Soc.*, 27:358–372, 1928.

[23] The LMFDB Collaboration. The L-functions and modular forms database. `http://www.lmfdb.org`, 2022. [Online; accessed March and August 2022].

[24] Kurt Mahler. An application of Jensen's formula to polynomials. *Mathematika*, 7:98–100, 1960.

[25] Chloe Martindale. Hilbert modular polynomials. *J. Number Theory*, 213:464–498, 2020.

[26] Enea Milio and Damien Robert. Modular polynomials on Hilbert surfaces. *J. Number Theory*, 216:403–459, 2020.

[27] Morris Newman. Construction and application of a class of modular functions. *Proc. London Math. Soc. (3)*, 7:334–350, 1957.

[28] Andrew P. Ogg. Hyperelliptic modular curves. *Bull. Soc. Math. France*, 102:449–462, 1974.

[29] Karl Rubin and Alice Silverberg. Choosing the correct elliptic curve in the CM method. *Math. Comp.*, 79(269):545–561, 2010.

[30] Reinhard Schertz. Die singulären Werte der Weberschen Funktionen $\mathfrak{f}$, $\mathfrak{f}_1$, $\mathfrak{f}_2$, γ_2, γ_3. *J. Reine Angew. Math.*, 286/287:46–74, 1976.

[31] Reinhard Schertz. Weber's class invariants revisited. *J. Théor. Nombres Bordeaux*, 14(1):325–343, 2002.

[32] Atle Selberg. On the estimation of Fourier coefficients of modular forms. In *Proc. Sympos. Pure Math., Vol. VIII*, pages 1–15. Amer. Math. Soc., Providence, R.I., 1965.

[33] Goro Shimura. *Introduction to the arithmetic theory of automorphic functions.* Kanô Memorial Lectures, No. 1. Iwanami Shoten Publishers, Tokyo; Princeton University Press, Princeton, N.J., 1971. Publications of the Mathematical Society of Japan, No. 11.

[34] Andrew V. Sutherland. Computing Hilbert class polynomials with the Chinese remainder theorem. *Math. Comp.*, 80(273):501–538, 2011.

[35] Andrew V. Sutherland. Accelerating the CM method. *LMS J. Comput. Math.*, 15:172–204, 2012.

[36] The Sage Developers. *SageMath, the Sage Mathematics Software System (Version 9.5)*, 2022. `https://www.sagemath.org`.

[37] Heinrich Weber. *Algebraische Zahlen*, volume 3 of *Lehrbuch der Algebra.* Friedrich Vieweg, 1908.

Chapter 7

Horizontal racewalking using radical isogenies

This chapter consists of a paper written together with Wouter Castryck, Thomas Decru, and Frederik Vercauteren. It has been published as

Wouter Castryck, Thomas Decru, Marc Houben, and Frederik Vercauteren. Horizontal racewalking using radical isogenies. In *Advances in Cryptology – ASIACRYPT 2022*, pages 67–96, Lecture Notes in Computer Science, vol 13792. Springer, Cham. `https://doi.org/10.1007/978-3-031-22966-4_3`.

All authors of this paper contributed equally to the work.

Compared to the published version, we have corrected a few typos and mathematical errors, added a reference in the proof of Theorem 7.6.5 to code in the GitHub repository associated to Conjecture 7.6.4, and extended radical isogeny formulae up to degree $N = 41$ (previously up to $N = 37$). The numbering (of e.g. theorems and definitions) in the published version is different.

Abstract

We address three main open problems concerning the use of radical isogenies, as presented by Castryck, Decru and Vercauteren at Asiacrypt 2020, in the computation of long chains of isogenies of fixed, small degree between elliptic curves over finite fields. Firstly, we present an interpolation method for finding radical isogeny formulae in a given degree N, which by-passes the need for factoring division polynomials over large function fields. Using this method, we are able to push the range for which we have formulae at our disposal from $N \leq 13$ to $N \leq 41$ (where in the range $18 \leq N \leq 41$ we have restricted our attention to prime powers). Secondly, using a combination of known techniques and ad-hoc manipulations, we derive optimized versions of these formulae for $N \leq 19$, with some instances performing more than twice as fast as their counterparts from 2020. Thirdly, we solve the problem of understanding the correct choice of radical when walking along the surface between supersingular elliptic curves over $\mathbb{F}_p$ with $p \equiv 7 \bmod 8$; this is non-trivial for even N and was settled for $N = 2$ and $N = 4$ only, in the latter case by Onuki and Moriya at PKC 2022. We give a conjectural statement for all even N and prove it for $N \leq 14$. The speed-ups obtained from these techniques are substantial: using 16-isogenies, the computation of long chains of 2-isogenies over 512-bit prime fields can be accelerated by a factor 3, and the previous implementation of CSIDH using radical isogenies can be sped up by about 12%.

7.1 Introduction

One of the core operations in isogeny-based cryptography is the fast computation of the codomain curve of a cyclic chain of horizontal $\mathbf{F}_q$-isogenies of some fixed small-to-moderate degree $N \geq 2$ between elliptic curves over a finite field $\mathbf{F}_q$. Here, let us recall that an $\mathbf{F}_q$-isogeny between two elliptic curves over $\mathbf{F}_q$ is called horizontal if their $\mathbf{F}_q$-rational endomorphism rings are isomorphic imaginary quadratic orders. The primary use cases are CRS [10, 22] and CSIDH [7], which are proposals for post-quantum key exchange. However fast horizontal isogenies are also key to various other recent constructions, including digital signatures [2], oblivious transfer constructions [15], verifiable delay functions [12], and schemes for delay encryption [11].

This paper presents a speed-up of such computations. More concretely, we upgrade the radical isogeny approach from [6], where for any given N one produces an iterable formula for computing the elliptic curves in a cyclic chain of N-isogenies, with each step involving the extraction of an Nth root of some radicand $\rho_N \in \mathbf{F}_q$; whence the name "radical". Asymptotically, for fixed N and growing q, the cost of evaluating this formula is dominated by one exponentiation in $\mathbf{F}_q$. This should be compared to one scalar multiplication on an elliptic curve over $\mathbf{F}_q$, which is the dominant cost of the standard approach using Vélu's formulae [26]. In practice however, radical isogenies are useful for small N only, because they come with a large overhead; part of the goal of the current paper is to reduce this overhead.

A first problem is simply *finding* radical isogeny formulae. Indeed, while their existence was argued in [6, §3] by means of the Tate pairing, producing concrete instances is a non-trivial task. The method proposed in [6, §4] relies on finding a zero of the reduced N-division polynomial of a Vélu-type codomain curve over a certain modular function field over $\mathbf{Q}$. As N grows, not only the division polynomial but also this codomain curve and the function field become increasingly complicated, and one quickly reaches the point where this method becomes infeasible. Consequently, the GitHub repository accompanying [6] contains no radical isogeny formulae beyond $N = 13$.

A second problem is that radical isogeny formulae are highly non-unique, with freedom coming from the choice of curve-point model (e.g., the Tate normal form), from the choice of the radicand ρ_N, and from relations in the modular function field. Different radical isogeny formulae for the same value of N can have very different practical performances, and in view of the large overhead it is crucial to try and produce the most efficient version. Here we should mention recent work by Onuki and Moriya [17], who use Montgomery curves to find faster formulae in degrees $N = 3, 4$. Chi-Dominguez and Reijnders [9] have presented projective (= inversion-free) radical isogeny formulae in degrees $2 \leq N \leq 5$ and $N = 7, 9$, but these are constructed directly from the corresponding formulae from [6].

A third problem is that it is not always clear *which* Nth root of ρ_N needs to be chosen in order to walk horizontally. In the CSIDH setting of supersingular elliptic curves over a finite prime field $\mathbf{F}_p$, horizontality comes for free if N is odd; in this case ρ_N has exactly one Nth root in $\mathbf{F}_p$. But even-degree $\mathbf{F}_p$-isogenies, of which non-trivial cyclic chains exist when $p \equiv 7 \bmod 8$ only, are a concern. In this case ρ_N will admit two Nth roots in $\mathbf{F}_p$, and selecting the wrong option will lead to a change of endomorphism ring and, as a result, in a breakdown of the iteration. This can be circumvented by an additional quadratic residuosity check at each step, but this is an annoying extra cost. In [4, Lem. 4] it was shown that this cost can be avoided when $N = 2$, because for the concrete radical isogeny formula presented there, the correct choice always turns out to be the principal square root, i.e. the unique square root which is again a square. This observation was extended to $N = 4$, now in terms of a principal fourth root, first as a conjecture [6, Conj. 2] and recently proved by Onuki and Moriya [17]. As mentioned in [6, §7], the correct generalization to arbitrary even N is not immediately apparent.

Contributions

We contribute significantly to each of the above open problems, which are listed explicitly in [6, §7]. Concretely, we address:

1. *Formula generation.* We develop an entirely different method for finding radical isogeny formulae in any given degree N, which avoids the need for factoring division polynomials over large function fields. The method uses interpolation over the modular curve $X_1(N)$ and is inspired by an alternative, Galois-theoretic proof of the existence of radical isogeny formulae along the lines of [5]. Using this method, we managed to generate radical isogeny formulae in degree as large as $N = 41$.

2. *Formula optimization.* The optimization and/or simplification of rational expressions modulo relations is an old and complicated problem, see for example [16]. In our case however, ad-hoc manipulations seem to yield the best results. We now believe to have found reasonably optimized formulae up to $N = 19$, with e.g. formulae for $N = 11, 13$ that can compete with our (optimized) version of $N = 7$. To highlight one example, for $N = 8$ we present the iteration

$$A \leftarrow \frac{-2A(A-2)\alpha^2 - A(A-2)}{(A-2)^2\alpha^4 - A(A-2)\alpha^2 - A(A-2)\alpha + A} \text{ with } \alpha = \sqrt[8]{\frac{-A^2(A-1)}{(A-2)^4}}$$

   whose counterpart from [6] spanned nearly a quarter of a page.

3. *Ensuring horizontality.* We believe to have found the correct generalization, at least conjecturally, of the observations from [4, Lem. 4], [6, Conj. 2] and [17, §5] for $N = 2, 4$ to arbitrary even N. The surprising new ingredient beyond $N = 4$ is that the principal Nth root needs to be tweaked by the Legendre symbol of a certain coefficient appearing in Tate's normal form; for $N = 4$ this Legendre symbol is always -1 so it goes unnoticed. With the aid of Magma we managed to prove this generalization up to $N = 14$.

One illustrative example where the three contributions resonate is the case $N = 16$. When computing long chains of 2-isogenies, e.g. as in the set-up phase of the delay function from [11], we can use radical 16-isogenies to take 4 horizontal steps "at once", resulting in an asymptotic speed-up by a factor of 4. Experimentally, we observed a speed-up by a factor of about 3 over a 512-bit prime field.

As for CSIDH, we have generated a new prime CRAD-513 capable of handling radical 8- and 9-isogenies, and using our new and optimized formulae we obtained a speed-up of about 12% when compared to the implementation of CSURF-512 from [6]. Furthermore, comparing this to the pre-radical isogenies implementation of CSIDH-512, one sees that the overall speed-up caused by radical isogenies at the 512-bit prime level is about 35%. We expect that there remains room for pushing this quite a bit further, for example by optimizing formulae for $N > 19$.

7.2 Background

Throughout, we let K denote a field, unless otherwise specified. The base point (= neutral element) of an elliptic curve E/K is denoted by $\mathcal{O}_E$, or just $\mathcal{O}$ if E is clear from the context.

7.2.1 Division polynomials

For an elliptic curve $E : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$, $a_i \in K$ in long Weierstrass form we set $b_2 = a_1^2 + 4a_2$, $b_4 = 2a_4 + a_1a_3$, $b_6 = a_3^2 + 4a_6$, $b_8 = a_1^2a_6 + 4a_2a_6 - a_1a_3a_4 + a_2a_3^2 - a_4^2$. For each integer $N \geq 0$ we define the *N-division polynomial* as

$$\Psi_{E,0} = 0, \quad \Psi_{E,1} = 1, \quad \Psi_{E,2} = 2y + a_1x + a_3, \quad \Psi_{E,N} = t \cdot \prod_{Q \in (E[N] \setminus E[2])/\pm} (x - x(Q)),$$

where $t = N$ if N is odd and $t = \frac{N}{2} \cdot \Psi_{E,2}$ if N is even. Note that $\Psi_{E,2}^2 = 4x^3 + b_2x^2 + 2b_4x + b_6$ is a univariate polynomial in x. These division polynomials can be computed efficiently, thanks to the following recurrence relations:

$$\begin{aligned}
\Psi_{E,3} &= 3x^4 + b_2x^3 + 3b_4x^2 + 3b_6x + b_8, \\
\frac{\Psi_{E,4}}{\Psi_{E,2}} &= 2x^6 + b_2x^5 + 5b_4x^4 + 10b_6x^3 + 10b_8x^2 + (b_2b_8 - b_4b_6)x + b_4b_8 - b_6^2, \\
\Psi_{E,2N+1} &= \Psi_{E,N+2}\Psi_{E,N}^3 - \Psi_{E,N-1}\Psi_{E,N+1}^3 \text{ if } N \geq 2, \\
\Psi_{E,2N} &= \frac{\Psi_{E,N}}{\Psi_{E,2}}(\Psi_{E,N+2}\Psi_{E,N-1}^2 - \Psi_{E,N-2}\Psi_{E,N+1}^2) \text{ if } N \geq 3.
\end{aligned}$$

By definition, we have that $\Psi_{E,N}(P) = 0$ for any non-trivial $P \in E[N]$. If one is interested in the points of exact order N, then one can use the *reduced N-division polynomial $\psi_{E,N}$* defined as $\Psi_{E,N}/\mathrm{lcm}_{d|N,d\neq N}\{\Psi_{E,d}\}$. For all primes ℓ, we simply have $\Psi_{E,\ell} = \psi_{E,\ell}$. Observe that for $N > 2$, the reduced N-division polynomial of E is a univariate polynomial in x.

Scalar multiplication by N on E can be expressed explicitly using division polynomials [20, Ex. 3.6]:

$$[N]P = \left(\frac{\phi_{E,N}(P)}{\Psi_{E,N}(P)^2}, \frac{\omega_{E,N}(P)}{\Psi_{E,N}(P)^3} \right), \tag{7.1}$$

with $\phi_{E,N} = x\Psi_{E,N}^2 - \Psi_{E,N+1}\Psi_{E,N-1}$ and $\omega_{E,N} = \frac{1}{2\Psi_{E,N}}(\Psi_{E,2N} - \Psi_{E,N}(a_1\phi_{E,N} + a_3\Psi_{E,N}^2))$.

7.2.2 Tate's normal form

We study elliptic curves E/K that are equipped with a distinguished K-rational point P of finite order N. For $N \geq 4$ such a curve-point pair (E, P) is isomorphic to a unique pair of the form

$$E_{b,c} : y^2 + (1-c)xy - by = x^3 - bx^2, \qquad P = (0,0), \tag{7.2}$$

for some $b, c \in K$. This distinguished model is called the *Tate normal form.* It is worth mentioning that the first few scalar multiples of $(0,0) \in E_{b,c}$ are easy expressions in terms of b and c, e.g.,

$$\begin{aligned} -(0,0) = (0,b),\ 2(0,0) = (b,bc),\ -2(0,0) = (b,0),& \\ 3(0,0) = (c, b-c),\ -3(0,0) = (c,c^2).& \end{aligned}$$

Expressions for higher multiples can be found using (7.1).

Furthermore, for every $N \geq 4$ one can write down a polynomial $F_N \in \mathbf{Z}[b,c]$ whose vanishing, along with the non-vanishing of the discriminant

$$\Delta(E_{b,c}) = b^3(16b^2 - 8bc^2 - 20bc + b + c(c-1)^3),$$

characterizes in any characteristic that the point $(0,0) \in E_{b,c}$ has exact order N. This polynomial can be found as a factor of the constant term of $\psi_{E_{b,c},N}(x) \in \mathbf{Z}[b,c][x]$, or by analyzing $N(0,0)$. It is uniquely determined up to sign. The first few instances are $F_4 = c$, $F_5 = c - b$, $F_6 = c^2 - b + c$, $F_7 = c^3 - b^2 + bc$, $F_8 = bc^2 - 2b^2 + 3bc - c^2$, see again [23, §2]. Thus, when viewing $E_{b,c}$ over the fraction field of $K[b,c]/(F_N)$, one can think of it as a "universal" curve-point pair from which all elliptic curves $E/\overline{K}$ equipped with a point $P \in E$ of order N are obtained through specialization at (unique) concrete values in $\overline{K}$ for b, c.

7.2.3 Radical isogenies

Vélu's formulae from [26] must be fed with the explicit coordinates of the points in $G = \ker\varphi$. In many applications, this kernel is a priori described in a more implicit form. For instance, in CSIDH it typically concerns the "unique subgroup of $E(\mathbf{F}_p)$ of order ℓ" for some odd prime number ℓ. An explicit generator of this subgroup can be found by repeatedly sampling $Q \leftarrow E(\mathbf{F}_p)$ and computing $\frac{p+1}{\ell}Q$ until its order is ℓ, but this scalar multiplication comes at a major cost which can dominate the application of Vélu's formulae itself. Radical isogenies, as introduced in [6], are an attempt at mitigating this.

The key observation behind radical isogenies is that if $\ker\varphi$ is cyclic, say generated by a point $P \in E(K)$ of order $N \geq 2$ coprime to $\operatorname{char} K$, then Vélu's formulae for producing a defining equation of $E' = E/\langle P\rangle$ can be augmented with formulae yielding the coordinates of a point $P' \in E'$ such that

$$E \overset{\varphi}{\to} E' = E/\langle P\rangle \to E'/\langle P'\rangle$$

is cyclic of degree N^2. Consequently, when computing a non-backtracking chain of N-isogenies, from the second step onwards the formulae allow to bypass the scalar multiplication. The formulae depend on N and can be chosen to

- be *radical*, in that they are algebraic expressions in the coefficients of E, the coordinates of P and a radical $\sqrt[N]{\rho_N}$, where the radicand ρ_N is itself an algebraic expression in the coefficients of E and the coordinates of P,

- be *complete*, in that changing the choice of $\sqrt[N]{\rho_N}$, i.e., scaling it with Nth roots of unity, produces generators for the kernel of each N-isogeny that cyclically extends φ,

- have *good reduction*, in the sense that they have coefficients in $\mathbf{Z}[1/N]$ and they can be applied to any elliptic curve E, over any field K with $\operatorname{char} K \nmid N$, equipped with a point $P \in E(K)$ of order N.

In [6] the existence of such formulae is argued using properties of the Tate pairing. The good reduction property is in fact stated as a conjecture [6, Conj. 1].

Remark 7.2.1 When working over $K = \mathbf{F}_q$ for some prime power q satisfying $\gcd(q-1, N) = 1$, one usually wants to choose the unique instance of $\sqrt[N]{\rho_N}$ belonging to $\mathbf{F}_q$; see [6, §5.1]. This instance can be computed as ρ_N^{μ} with $\mu \in \mathbf{Z}$ a multiplicative inverse of N modulo $q-1$. So the cost of evaluating the formulae is asymptotically dominated by one field exponentiation. Unfortunately, the formulae come with a large overhead and, for fixed q, they outperform plain Vélu for small values of N only. The main goal of this paper is to push this crossover point to larger values of N. ♢

Example 7.2.2 (taken from [6, §4]) Consider an elliptic curve E with a point P of order $N = 5$. The Tate normal form of this curve-point pair is $E_{b,b} = y^2 + (1-b)xy - by = x^3 - bx^2$, $P = (0,0)$ for some $b \neq 0, (11 \pm 5\sqrt{5})/2$. Vélu's formulae produce the following equation for $E' = E/\langle P\rangle$:

$$y^2 + (1-b)xy - by = x^3 - bx^2 - 5b(b^2 + 2b - 1)x - b(b^4 + 10b^3 - 5b^2 + 15b - 1).$$

Analyzing the roots of $\psi_{E',5}(x)$ shows that for $\alpha = \sqrt[5]{\rho_5}$ with $\rho_5 = b$ the point

$$\begin{aligned}P' = \big(&5\alpha^4 + (b-3)\alpha^3 + (b+2)\alpha^2 + (2b-1)\alpha - 2b,\\ &5\alpha^4 + (b-3)\alpha^3 + (b^2 - 10b + 1)\alpha^2 + (13b - b^2)\alpha - b^2 - 11b\big)\end{aligned}$$

on E' has order 5 and generates the kernel of a cyclic extension of φ (it is such that $\hat{\varphi}(P') = P$). There are five such cyclic extensions, corresponding to the five possible choices for α. Rewriting the curve-point pair (E', P') into Tate normal form produces the curve $E_{b',b'}$ where b' is given by the iterable formula

$$\rho_5' = b' = \alpha\frac{\alpha^4 + 3\alpha^3 + 4\alpha^2 + 2\alpha + 1}{\alpha^4 - 2\alpha^3 + 4\alpha^2 - 3\alpha + 1}. \tag{7.3}$$

☆

The above example illustrates the strategy from [6] for *finding* radical isogeny formulae. The cases $N = 2, 3$ are easy to handle [6, §4] so we assume that $N \geq 4$. One starts from the "universal" curve-point pair $E = E_{b,c}$, $P = (0, 0)$ over

$$\mathbf{Q}_N(b, c) := \operatorname{Frac}\frac{\mathbf{Q}[b, c]}{(F_N)}$$

and one computes a defining equation for $E' = E/\langle P\rangle$ using Vélu's formulae. One then computes the division polynomial $\psi_{E',N}(x)$ and, for a suitable radicand $\rho_N \in \mathbf{Q}_N(b, c)$, one finds the root $x_0' \in \mathbf{Q}_N(b, c)(\sqrt[N]{\rho_N}\,)$ that is the x-coordinate of a point $P' \in E'$ such that $\hat{\varphi}(P') = P$, using a root-finding algorithm; this step is a severe bottleneck. If successful, then the corresponding y-coordinate $y_0' = y(P')$ can be found by solving a quadratic equation over $\mathbf{Q}_N(b, c)(\sqrt[N]{\rho_N}\,)$. The coordinates x_0', y_0' are the radical isogeny formulae we are after; one hopes, and observes in practice, that the good reduction property comes for free. By writing the curve-point pair (E', P') back in Tate normal form $(E_{b',c'}, (0, 0))$ one obtains formulae for b', c' that can be applied iteratively, as in the case of (7.3).

Concerning the radicand ρ_N, it was argued in [6, §3] that $\rho_N = f_{N,P}(-P)$ works, where $f_{N,P}$ is the function on $E_{b,c}$ with divisor $N(P) - N(\mathcal{O})$ and having leading coefficient 1 when expanded in terms of the uniformizer x/y at $\mathcal{O}$, so that ρ_N is a representative of the Tate pairing $t_N(P, -P)$; see [14, Lem. 1].

7.3 Modular curves and Galois theory

This section recalls some of the theory of Galois coverings of modular curves. We mainly refer to [18] and [19]. Along the way we present an alternative proof of the existence of radical isogeny formulae [6, Thm. 5]. This closely resembles the discussion in [5, §3].

7.3.1 Congruence subgroups

Classically, as Riemann surfaces, modular curves are quotients $X = X_\Gamma = \mathbf{H}^*/\Gamma$ of the extended complex upper half plane $\mathbf{H}^* = \mathbf{H} \cup \mathbf{P}^1(\mathbf{Q})$ by a *congruence subgroup* $\Gamma \subset \mathrm{SL}_2(\mathbf{Z})$, i.e. a subgroup containing $\Gamma(N) \subset \mathrm{SL}_2(\mathbf{Z})$, the kernel of reduction modulo N, for some $N \in \mathbf{Z}_{>0}$. The minimal N for which this last property holds is called the *level* of X. The modular curve X admits a natural Zariski-open subset $Y = \mathbf{H}/\Gamma$, and the (finite collection of) points $X \setminus Y$ are called the *cusps* of X. Modular curves can be seen as irreducible smooth complex projective curves, and they always have a "moduli interpretation", in the sense that they (specifically, the non-cuspidal points) parametrize complex elliptic curves together with some additional structure on the N-torsion subgroup.

To make this latter viewpoint more precise, we will consider a different, slightly more general, method to construct "modular" curves. These modular curves will be more general in the sense that they may be reducible as complex projective curves; but they will be irreducible over $\mathbf{Q}$, and their geometrically irreducible components shall be modular curves in the classical sense. Let $N \geq 1$ be an integer and consider the "universal" elliptic curve

$$E_j : y^2 = 4x^3 - \frac{27j}{j-1728}x - \frac{27j}{j-1728}$$

over $\mathbf{Q}(j)$, whose j-invariant equals the indeterminate j. Let $\mathbf{Q}(j, E_j[N]) \subset \overline{\mathbf{Q}(j)}$ be the field obtained by adjoining the coordinates of all N-torsion points of E_j. Then this is a Galois extension, whose Galois automorphisms are completely determined by their action on $E[N]$. In particular, we have that the Galois group is isomorphic to the automorphism group $\mathrm{GL}_2(\mathbf{Z}/N\mathbf{Z})$ of the N-torsion.

Let $H \subset \mathrm{GL}_2(\mathbf{Z}/N\mathbf{Z})$ be a subgroup containing -1. The fixed field $\mathbf{Q}(j, E_j[N])^H$ is the function field of a smooth projective curve over $\mathbf{Q}$, which we will denote by X_H. This curve has a natural moduli interpretation, in the sense that away from a finite set its geometric points parametrize elliptic curves over $\overline{\mathbf{Q}}$ together with a certain structure on the N-torsion. More explicitly, it parametrizes pairs (E, α) up to H-isomorphism, where $\alpha : E[N] \to (\mathbf{Z}/N\mathbf{Z})^2$ is an isomorphism of abelian groups and two pairs (E_1, α_1) and (E_2, α_2) are called H-isomorphic if there exists an isomorphism $\varphi : E_1 \to E_2$ and an element $h \in H$ such that $\alpha_1 = h \circ \alpha_2 \circ \varphi$; see [19, §3] for more details. E.g. if we take for H the subgroup of $\mathrm{GL}_2(\mathbf{Z}/N\mathbf{Z})$ of upper-diagonal matrices then X_H is the classical modular curve $X_0(N)$, which parametrizes elliptic curves together with a cyclic subgroup of order N.

The connection to modular curves in the classical sense is quite straightforward. If we denote by $\Gamma_H = \pi^{-1}(\mathrm{GL}_2(\mathbf{Z}/N\mathbf{Z})) \subset \mathrm{SL}_2(\mathbf{Z})$ the congruence subgroup that is the inverse image of H under the reduction modulo N map $\pi : \mathrm{SL}_2(\mathbf{Z}) \to \mathrm{GL}_2(\mathbf{Z}/N\mathbf{Z})$, then we have that $X_H \cong X_{\Gamma_H}$ as complex projective curves if and only if $\det(H) = (\mathbf{Z}/N\mathbf{Z})^\times$; in general X_H will be geometrically isomorphic to the disjoint union of $[(\mathbf{Z}/N\mathbf{Z})^\times : \det(H)]$ copies of X_{Γ_H}.

7.3.2 The main suspects

Let $N \geq 3$. The subgroups $H \supset H'$ of $\mathrm{GL}_2(\mathbf{Z}/N^2\mathbf{Z})$ consisting of matrices having respective forms

$$\begin{pmatrix} \pm 1 \bmod N & * \\ 0 \bmod N & * \end{pmatrix}, \qquad \text{and} \qquad \begin{pmatrix} \pm 1 \bmod N & * \\ 0 & * \end{pmatrix}$$

correspond to the modular curves which we denote $X_1(N) = X_H$ and $X_1'(N) = X_{H'}$ respectively. The curve $X_1(N)$ is the classical modular curve parametrizing pairs (E, P) where E is an elliptic curve and $P \in E$ is an N-torsion point. The curve $X_1'(N)$ parametrizes triples (E, P, P') where P' is a *P-distinguished point*, i.e. a point $P' \in E/\langle P\rangle$ that maps to P under the dual isogeny $E/\langle P\rangle \to E$. Alternatively, it parametrizes pairs (E, C), where $C = \{Q, Q+P, \ldots, Q+(N-1)P\}$ is a coset on E modulo the order-N point P, where $NQ = P$.

Let us denote by $K \subset L$ the respective function fields over $\mathbf{Q}$ of these curves:

$$K := \mathbf{Q}(X_1(N)) = \mathbf{Q}(j, E_j[N])^H, \quad L := \mathbf{Q}(X_1'(N)) = \mathbf{Q}(j, E_j[N])^{H'}.$$

Then K, L are the fields $\mathbf{Q}_N(b, c)$ and $\mathbf{Q}_N(b, c, \sqrt[N]{\rho_N})$ from Section 7.2.3. The canonical inclusion $K \hookrightarrow L$ corresponds to the degree-N forgetful map $X_1'(N) \to X_1(N) : (E, P, P') \mapsto (E, P)$. As we will see in the next section, it is possible to deduce from a purely Galois-theoretic argument that the extension L/K is radical.

7.3.3 The Galois structure

Lemma 7.3.1 *Let $N \in \mathbf{Z}_{>0}$ and let $K \subset L$ be a degree N extension of fields whose characteristic does not divide N. Let $\zeta_N \in \overline{L}$ be a primitive Nth root of unity and assume that $L(\zeta_N)$ is Galois over K with Galois group*

$$\mathrm{Gal}(L(\zeta_N)/K) = \mathrm{Gal}(L(\zeta_N)/K(\zeta_N)) \rtimes \mathrm{Gal}(L(\zeta_N)/L),$$

where the first factor is cyclic of order N, say generated by σ, and where the semidirect product is according to the rule

$$\tau_j \circ \sigma^i \circ \tau_j^{-1} = \sigma^{ij} \tag{7.4}$$

for all $i = 0, 1, \ldots, N-1$ and all $\tau_j : \zeta_N \mapsto \zeta_N^j \in \mathrm{Gal}(L(\zeta_N)/L)$. Then there exists an $\alpha \in L$ such that $L = K(\alpha)$ and $\alpha^N \in K$.

Proof. The restricted maps $\sigma^i|_L : L \to L(\zeta_N)$ are pairwise distinct. Indeed, if $i, i' \in \{0, 1, \ldots, N-1\}$ are such that $\sigma^i|_L = \sigma^{i'}|_L$, then

$$\sigma^{i-i'} \in \mathrm{Gal}(L(\zeta_N)/K(\zeta_N)) \cap \mathrm{Gal}(L(\zeta_N)/L) = \{\mathrm{id}\},$$

which can only be true if $i = i'$. From [21, Lem. 0CKL] we get that these restricted maps are linearly independent over $L(\zeta_N)$. Thus there exists $\beta \in L$ such that $\alpha :=$

$\sum_{i=0}^{N-1} \zeta_N^i \sigma^i(\beta)$ is non-zero. From

$$\tau_j(\alpha) = \sum_i \zeta_N^{ij} (\tau_j \circ \sigma^i)(\beta) = \sum_i \zeta_N^{ij} (\sigma^{ij} \circ \tau_j)(\beta) = \sum_i \zeta_N^{ij} \sigma^{ij}(\beta) = \alpha$$

it follows that $\alpha \in L$ as well. Now observe that α was constructed in such a way that $\sigma^i(\alpha) = \zeta_N^{-i}\alpha$ for $i = 0, 1, \dots, N-1$, which has two crucial consequences. On the one hand, it implies that $\mathrm{Gal}(L(\zeta_N)/L)$ is the exact group of automorphisms fixing $K(\alpha)$, or in other words $L = K(\alpha)$. On the other hand, it implies that $\sigma(\alpha^N) = \sigma(\alpha)^N = (\zeta_N \alpha)^N = \alpha^N$, so that α^N is fixed by the entire Galois group, i.e. $\alpha^N \in K$ as wanted. □

Now let K, L as in Section 7.3.2. Below we give an alternative proof of the fact that L/K is a radical extension. Our strategy is to apply Lemma 7.3.1, so we will first prove that $L(\zeta_N)/K$ is Galois, and then find explicitly elements $\sigma, \tau_j \in \mathrm{Gal}(L(\zeta_N)/K)$ satisfying (7.5).

Theorem 7.3.2 *The morphism $X_1'(N) \to X_1(N)$ is a simple radical extension, i.e. the degree N extension of function fields*

$$\mathbf{Q}(j, E_j[N^2])^H \subseteq \mathbf{Q}(j, E_j[N^2])^{H'}$$

can be realized by adjoining $\sqrt[N]{\rho}$ for some function ρ on $X_1(N)$.

Proof. Let $\mathcal{H} \subset H'$ be the subgroup consisting of matrices whose determinant is $\equiv 1 \pmod{N}$. Then the corresponding fixed field $\mathbf{Q}(j, E_j[N^2])^{\mathcal{H}}$ is $L(\zeta_N)$. One can verify that $\mathcal{H}$ is a normal subgroup of H, which implies that $L(\zeta_N)/K$ is Galois of degree $N\varphi(N)$ with Galois group $H/\mathcal{H}$.

In order to understand its structure, we first consider the intermediate extension $L \subseteq L(\zeta_N)$, which is just a cyclotomic extension with Galois group $\{\, \tau_j : \zeta_N \mapsto \zeta_N^j \mid 0 \leq j < N, \gcd(j, N) = 1 \,\} \cong (\mathbf{Z}/N)^*$. When viewed as elements of $H/\mathcal{H}$, these maps can be identified with

$$\tau_j = \begin{pmatrix} 1 & 0 \\ 0 & j \end{pmatrix} \bmod \mathcal{H}.$$

Next, we concentrate on the intermediate extension $K(\zeta_N) \subset L(\zeta_N)$ which is of degree N, and its Galois group can be identified with the cyclic group

$$\left\langle \sigma := \begin{pmatrix} 1 & 0 \\ N & 1 \end{pmatrix} \right\rangle = \left\{ \sigma^i = \begin{pmatrix} 1 & 0 \\ iN & 1 \end{pmatrix} \,\middle|\, i = 0, 1, \dots, N-1 \right\},$$

which, as before, we consider modulo $\mathcal{H}$. It is easy to see that the elements $\tau_j \circ \sigma^i$ are pairwise distinct (e.g. because j is fully determined by the action of $\tau_j \circ \sigma^i$ on ζ_N, and then the uniqueness of i follows at once). Therefore these $N\varphi(N)$ elements must constitute the whole Galois group. The structure of the Galois group is then determined by the rules $\sigma^N = 1$, $\tau_j^{\varphi(N)} = 1$, and

$$\sigma^i \circ \tau_j = \begin{pmatrix} 1 & 0 \\ iN & j \end{pmatrix} = \tau_j \circ \sigma^{ij^{-1}}; \tag{7.5}$$

matching (7.4). The result now indeed follows by applying Lemma 7.3.1. □

Remark 7.3.3 The subgroup $\mathcal{H} \subset \mathrm{GL}_2(\mathbf{Z}/N^2\mathbf{Z})$ introduced in the proof of the Theorem corresponds to a modular curve $\mathcal{X}_1'(N)$ over $\mathbf{Q}$ with function field $L(\zeta_N)$. Since $[(\mathbf{Z}/N^2\mathbf{Z})^\times : \det(\mathcal{H})] = \varphi(N)$ it consists geometrically of $\varphi(N)$ copies of $X_1'(N)$, labeled by the different primitive Nth roots of unity ζ_N.

The level structure induced by $\mathcal{H}$ yields the following moduli interpretation of $\mathcal{X}_1'(N)$: it parametrizes triples (E, C, R), where $(E, C) \in X_1'(N)$ is as in Section 7.3.2 and $R \in E[N]$ is an N-torsion point independent of P (i.e. such that $E[N] = \langle P, R\rangle$), where we identify two such points R_1 and R_2 if their Weil pairing with P yields the same (primitive) Nth root of unity, i.e. if $e_N(P, R_1) = e_N(P, R_2)$. Forgetting R leads to a covering $\mathcal{X}_1'(N) \to X_1'(N)$ of degree $\varphi(N)$.

One can make sense of the Galois action of $L(\zeta_N)/K$ in terms of this moduli interpretation. Given a triple $\mathcal{P} = (E, \{Q, Q+P, \ldots, Q+(N-1)P\}, R)$, the images under σ and τ_j are

$$\begin{aligned} \sigma(\mathcal{P}) &= (E, \{Q+R, Q+R+P, \ldots, Q+R+(N-1)P\}, R), \\ \tau_j(\mathcal{P}) &= (E, \{jQ, jQ+P, \ldots, jQ+(N-1)P, R). \end{aligned}$$

◊

7.4 Radical isogeny formulae through interpolation

We now describe the method we used to compute the radical isogeny formulae. Explicitly, starting from the universal Tate normal curve $E = E_{b,c}$ over $K = \mathbf{Q}_N(b, c)$ together with the point $P = (0, 0) \in E$ of order $N \geq 4$, we would like to find an expression for the coordinates of a P-distinguished point P' on the quotient curve $E' = E/\langle P\rangle$ (whose Weierstrass model, let us assume, is given by Vélu's formulae). According to Section 7.3, these coordinates live over some radical field extension L of K. For simplicity, we will mostly focus on computing the x-coordinate of P', as the computation of the y-coordinate is more or less analogous.

7.4.1 A linear system

Let us denote by $\overline{K}$ an algebraic closure of K, and let $Q \in E(\overline{K})$ be such that $NQ = P$. We would like to find an expression for

$$\beta_0 := \sum_{i=0}^{N-1} x(Q + iP),$$

since by Vélu's formulae this is equivalent to finding the x-coordinate of P'. If we define

$$\gamma_d := \sum_{S \in E[N]} e_N(P,S)^d x(Q+S),$$

then $\gamma_d^N \in K$ for all $d \in \mathbf{Z}$: indeed, let $R \in E(\overline{K})$ be an N-torsion point so that $E[N] = \langle P, R \rangle$ and denote by $e_N : E[N] \times E[N] \to \overline{K}$ the Weil pairing. Then $\zeta_N := e_N(P,R)$ is a primitive Nth root of unity. By Remark 7.3.3, it follows that

$$\gamma_d = \sum_{j=0}^{N-1} e_N(P, jR)^d \sum_{i=0}^{N-1} x(Q + jR + iP) = \sum_{j=0}^{N-1} \zeta_N^{jd} \sigma^j(\beta_0),$$

for some generator $\sigma \in \mathrm{Gal}(L(\zeta_N)/K)$ of $\mathrm{Gal}(L(\zeta_N)/K(\zeta_N))$. Following the last paragraph of the proof of Lemma 7.3.1 now shows that $\gamma_d^N \in K$.

Note that $\gamma_d \in L(\zeta_N)$ depends on the choice of Q. However all of them are related as follows:

Lemma 7.4.1 *Let $Q, Q' \in E[N^2]$ be such that $NQ = NQ' = P$. Then there exists an Nth root of unity $\zeta \in \overline{K}$ such that $\gamma_d(Q) = \zeta^d \gamma_d(Q')$ for all $d \in \mathbf{Z}$. Moreover, for all $d \in \mathbf{Z}$ we have that γ_d/γ_1^d is an element of K that is independent of the choice of Q.*

Proof. We have that Q' differs from Q by an N-torsion point. Note that adding multiples of P to Q clearly does not affect the value of γ_d while adding a multiple kR of R scales it by ζ_N^{-kd}. This shows the first statement with $\zeta = \zeta_N^{-k}$. For the second part, note that the independence on Q already follows from the first part. Now let σ be as above and let τ_j be a generator for the cyclotomic extension $L(\zeta_N)/K(\gamma_1)$. Then $\tau_j(\gamma_d) = \gamma_d$, whereas $\sigma(\gamma_d) = \zeta_N^{-d}\gamma_d$. Since σ, τ_j together generate $\mathrm{Gal}(L(\zeta_N)/K)$ we see that γ_d/γ_1^d is invariant under all Galois automorphisms of $L(\zeta_N)/K$ and we conclude that it is an element of K. □

Defining

$$\beta_j := \sigma^j(\beta_0) = \sum_{i=0}^{N-1} x(Q + jR + iP),$$

we now have the following linear system.

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \zeta_N & \zeta_N^2 & \cdots & \zeta_N^{N-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \zeta_N^{N-1} & \zeta_N^{2(N-1)} & \cdots & \zeta_N^{(N-1)^2} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{N-1} \end{pmatrix} = \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \vdots \\ \gamma_{N-1} \end{pmatrix}.$$

In particular, if we set $\alpha := \gamma_1$ then we see that

$$\beta_0 = \frac{1}{N}\sum_{d=0}^{N-1} \gamma_d = \frac{1}{N}\sum_{d=0}^{N-1} \left(\frac{\gamma_d}{\gamma_1^d}\right) \alpha^d \in K(\alpha) = L. \tag{7.6}$$

We have now reduced the problem of finding radical isogeny formulae (at least the determination of the x-coordinate of P') to finding expressions for the elements $\gamma_d/\gamma_1^d \in K$ for all $d \in \{0, \dots, N-1\}$. In the next subsection we will describe the method we used to do this. Before that we should point out one subtlety. To ensure that (7.6) is well defined we must have $\alpha \neq 0$; in fact, to be able to use the formula in practice, we should know exactly the value of $\alpha^N \in K$. Though, given N, this is not so difficult to establish (or even guess) in practice; a proof of a closed expression for α^N that works for all N can be found in the appendix (from which it also follows that α is never zero), see Theorem 7.7.1.

7.4.2 Finding the formulae

Expressions for $c_d := \gamma_d/\gamma_1^d$ will of course depend heavily on how one represents the field $K = \mathbf{Q}(X_1(N))$. It turns out that the representation $K = \mathbf{Q}_N(b, c)$ as presented in Section 7.2.3 is not always optimal. In order to minimize the complexity of the resulting formulae, as well as the running time complexity of the algorithm used to find them, we will instead employ Sutherland's optimized models of $X_1(N)$ [24]. These models are optimal in the sense that they write K as the fraction field, which we will denote $\mathbf{Q}_N(A, B)$, of $\mathbf{Q}[A, B]/G_N(A, B)$ for some modular polynomial $G_N(A, B)$ whose degree in B matches the gonality of $X_1(N)$ over $\mathbf{Q}$ (at least for $N \leq 40$). In particular, we can theoretically write every element of K, specifically the c_d we are after, as a polynomial in $\mathbf{Q}(A)[B]$, where the degree in B is as small as one could hope for. It is also possible, and relatively easy in fact, to find an explicit expression for $b, c \in K$ in terms of Sutherland's functions A, B, so one can also express the universal Tate normal curve $E_{b,c}$ as a curve $E_{A,B}$ over $\mathbf{Q}_N(A, B)$.

The idea is now to determine the reduction $\overline{c_d} \in \mathbf{F}_p(A)[B]$ of the coefficients c_d modulo several primes p, and then to lift the results to $\mathbf{Q}(A)[B]$ using the Chinese Remainder Theorem. To find the $\overline{c_d}$, we sample many curves $E_{A,B}$ over $\mathbf{F}_p$ for which Q, R, and ζ_N of the previous section are all defined over $\mathbf{F}_p$. For each of these curves, we explicitly compute the coefficients c_d as elements of $\mathbf{F}_p$. Then, as long as the number of samples is sufficiently large, we can determine an expression for $\overline{c_d} \in \mathbf{F}_p(A)[B]$ by means of rational interpolation (this last step can be achieved purely by linear algebra over $\mathbf{F}_p$).

The main problem that arises is how to efficiently generate suitable samples $(A, B) \in X_1(N)(\mathbf{F}_p)$. The requirement that ζ_N be defined over $\mathbf{F}_p$ is rather trivially met by demanding that $p \equiv 1 \pmod{N}$. The condition that $Q, R \in E_{A,B}(\mathbf{F}_p)$, however, is more intricate, and simply generating random curves turns out to be far too inefficient for large N. Instead, we rely on an approach based on the theory of complex multiplication.

The CM Method

The endomorphism ring of an elliptic curve $E/\mathbf{C}$ is isomorphic to either $\mathbf{Z}$ or an order $\mathcal{O}$ in an imaginary quadratic number field. In the latter case we say that E has complex multiplication (CM) by $\mathcal{O}$. The j-invariants of such elliptic curves are algebraic integers. The *Hilbert class polynomial* $H_D(X) \in \mathbf{Z}[X]$ is the minimal polynomial over $\mathbf{Q}$ of the j-invariant of an elliptic curve $E/\mathbf{C}$ with CM by the quadratic order of discriminant D.

Ordinary elliptic curves over a finite field always have CM. An ordinary elliptic curve $E/\mathbf{F}_q$ with CM by the imaginary quadratic order $\mathcal{O}$ of discriminant D exists if and only if there exist $t, u \in \mathbf{Z}$ such that $u^2 D = t^2 - 4q$ and $p \nmid t$ (where $p = \operatorname{char} \mathbf{F}_q$). In this case H_D splits completely over $\mathbf{F}_q$ and its roots are precisely the j-invariants of elliptic curves with CM by $\mathcal{O}$. The trace of Frobenius of such curves is $\pm t$, so they will have $q + 1 \pm t$ points. One can use this to find curves over $\mathbf{F}_q$ with a desired number of points; this is known as the *CM Method*.

Sampling curves with torsion

We now describe how to use the CM method to construct curves $E_{A,B}$ with full N^2-torsion over $\mathbf{F}_p$; this will certainly ensure that the desired points Q, R be defined over $\mathbf{F}_p$. We thus want to find curves with number of points divisible by N^4. One approach is to strengthen the requirement that $p \equiv 1 \pmod{N}$ to $p \equiv 1 \pmod{N^4}$ and construct curves of trace 2 using the CM method, i.e. with CM by an order whose discriminant D satisfies an equation of the form $u^2 D = 2^2 - 4p$ for some $u \in \mathbf{Z}_{>0}$. The structure of the $\mathbf{F}_p$-rational N^∞-torsion also be controlled by D; if we choose D to be a divisor of $(2^2 - 4p)/N^4$ then $E[N^2](\mathbf{F}_p) \cong (\mathbf{Z}/N^2\mathbf{Z})^2$, see e.g. [8, Thm. 7].

Algorithm

We summarize the above discussion in the following pseudo algorithm generating radical isogeny formulae for $N \geq 4$. The SageMath code we used can be found in the GitHub repository accompanying this paper.

(i) Find all prime numbers $p \equiv 1 \pmod{N^4}$ up to a certain bound.

(ii) For each prime number p, determine the roots j_i of the Hilbert class polynomials H_D modulo p for every imaginary quadratic discriminant D of the form $u^2 N^4 D = 4(p-1)$ for some $u \in \mathbf{Z}$.

(iii) For each root j_i, determine the $(A, B) \in X_1(N)(\mathbf{F}_p)$ for which $j(E_{A,B}) = j_i$.

(iv) For each pair (A, B), if $E_{A,B}$ has trace $+2$, determine $c_d \in \mathbf{F}_p$ for all $d \in \{0, \ldots, N-1\}$.

(v) For each d, find a formula for $c_d \in \mathbf{F}_p(A)[B]$ by rational interpolation.

(vi) Lift the formulae to $\mathbf{Q}(A)[B]$ by the Chinese Remainder Theorem.

7.4.3 Iterative formulae

The above describes how to find an expression for the x-coordinate of P' as an element of $L = K(\alpha)$. An analogous method can be used to find an expression for the y-coordinate. By transforming the pair (E', P') to Tate normal form one can then also determine explicit formulae for Sutherland's parameters $A', B' \in L$ corresponding to the point $(E', P') \in X_1(N)(L)$. In this way, we obtain radical isogeny formulae that can be applied iteratively. We list formulae for prime powers $16 < N \leq 41$ in our GitHub repository.[1]

7.5 Optimizing the formulae

When optimizing radical isogeny formulae, one needs to take into account all of the following choices.

- The radicand ρ_N is not unique: it can be scaled with Nth powers in $\mathbf{Q}_N(b, c)$, and it can be raised to exponents that are coprime with N. Switching from one radicand to another results in different radical isogeny formulae with different performances.
- It is not self-evident that the optimized representations of $X_1(N)$ by Sutherland from [24] will result in optimized radical isogeny formulae.
- Elements in $\mathbb{Q}_N(b, c, \alpha)$ can be expressed in several ways since we work modulo the two relations $F_N(b, c) = 0$ and $\alpha^N = \rho_N(b, c)$.
- It is a priori not clear what formulae we are trying to optimize; e.g. for $E' = E/\langle P \rangle$ we can try to find optimal expressions for a P-distinguished point P' on E', or we can try to write E' in Tate normal form immediately.

We will focus on finding efficient enough formulae in this setting, where it seems nigh impossible to prove that they are indeed the most optimal (especially for $N \geq 10$ as we will see further up ahead). Hence we do not claim they are optimal, but they should not be far off and at the very least in certain cases a big improvement compared to the work in [6].

For $N \in \{4, 5, \ldots, 10\} \cup \{12\}$, the Tate normal form can be parametrized by a single parameter, say A. This means that the codomain curve of a radical N-isogeny can be put into a (new) Tate normal form with a single parameter, say A', where we translated the P-distinguished point P' to $(0, 0)$. In practice, this new parameter seems a good candidate to try to optimize, as can be seen from the case of $N = 4, 5$ from [6]. The raw equation for A' can be easily obtained by any algebraic software package for these small N.

To find an efficient representation of A', consider the curve $X_1'(N)$ defined by $\alpha^N - \rho_N, F_N = 0$. Then A' can be seen as a function on this curve and we can compute its divisor. For $N < 10$, an algebraic software package has no issues checking

[1] `https://github.com/KULeuven-COSIC/Horizontal_Radical_Isogenies`

which linear combinations of places in its support constitute principal divisors, and we can use this to peel off (easy) factors from A'. For every $N \in \{4, \dots, 9\}$, there are clear contenders for which factorization is most efficient. We list them all, skipping the case $N = 5$ which can be found in (7.3). Note that for $N \geq 6$, our "factorization" merely amounts to writing A' as the quotient of two easyish expressions in A and α.

$\mathbf{N = 4}$. In this case we have $b = A$, $c = 0$ and for $\alpha^4 = A$ we have that

$$A' = \alpha \frac{4\alpha^2 + 1}{(2\alpha + 1)^4}. \tag{7.7}$$

$\mathbf{N = 6}$. In this case we have $b = A(A-1)$, $c = A - 1$ and for $\alpha^6 = -A^2(A-1)$ we have that

$$A' = \frac{(-3A+2)\alpha^4 + 3A^2\alpha^2 + 2A\alpha - 3A^3 + 4A^2}{\alpha^4 + 2A\alpha^2 + 3A\alpha + A^2}. \tag{7.8}$$

$\mathbf{N = 7}$. In this case we have $b = A^2(A-1)$, $c = A(A-1)$ and for $\alpha^7 = A^4(A-1)$ we have that

$$A' = \frac{\alpha^6 + A\alpha^5 + 2A^3\alpha^2 - A^3\alpha + A^4}{-\alpha^6 + A\alpha^4 + A^3\alpha^2 - 2A^3\alpha + A^4}.$$

$\mathbf{N = 8}$. In this case we have that $b = \frac{A(A-1)}{(A-2)^2}$, $c = \frac{-A(A-1)}{A-2}$ and for $\alpha^8 = \frac{-A^2(A-1)}{(A-2)^4}$ we have that

$$A' = \frac{-2A(A-2)\alpha^2 - A(A-2)}{(A-2)^2\alpha^4 - A(A-2)\alpha^2 - A(A-2)\alpha + A}.$$

$\mathbf{N = 9}$. In this case we have that $b = A^2(A-1)(A^2 - A + 1)$, $c = A^2(A-1)$ and for $\alpha^9 = A^4(A-1)(A^2 - A + 1)^3$ we have that

$$A' = \frac{A(A^2 - A + 1)(\alpha^5 + A(A^2 - A + 1)\alpha^2 + A^2(A^2 - A + 1)^2)}{\alpha^7 - A(A^2 - A + 1)(A-1)\alpha^4 - A^3(A^2 - A + 1)^2\alpha + (A(A^2 - A + 1))^3}.$$

For $N \geq 10$, Magma struggles to efficiently verify whether a given divisor is principal, and those that do get found are less clean than the above factors, so we will optimize these two cases with the more general method for larger N.[2]

If we compute E' as $E/\langle P \rangle$ by means of Vélu's formulae, then E' is in (long) Weierstrass form and we still need to compute an isomorphism to put E' back in Tate normal form E'_t for certain $b', c' \in \mathbb{Q}_N(b, c, \alpha)$. By [20, Prop. 1.3(d)], the isomorphism $\iota : E'_t \to E'$ is determined by a 4-tuple (u, r, s, t), where $P' = (r, t)$ is the P-distinguished point and u is a unit. This u, when seen as a polynomial of degree $N-1$ in $\mathbb{Q}_N(b, c)[\alpha]$, seems to always be efficient to write down and evaluate. Furthermore, the expressions uc' and ub'/c' also enjoy this feature. In particular, a factor that arises in the coefficient of α^i has a high chance of also being there in the coefficient of α^j for $j > i$, which makes this efficient to evaluate in a Horner scheme with rising powers of α. We provide the concrete expressions for $N = 10$ and refer the reader

[2] We remark that for the smaller N it can be extremely fast to let a computer algebra software package verify that a given divisor is *not* principal, but to prove it is principal is harder in the majority of cases.

to our GitHub repository for larger N. Remark that for $N = 10$ we still work with a one-parameter family of curves and the expression uA' is just as efficient as uc' or ub'/c'. The operation counts for all formulae $N \in \{4, 5, \dots, 17\} \cup \{19\}$ can be found in Table 7.1.

$\mathbf{N = 10}$. In this case we have

$$b = \frac{A^3(A-1)(2A-1)}{(A^2-3A+1)^2}, \quad c = \frac{-A(A-1)(2A-1)}{(A^2-3A+1)}, \quad \alpha^{10} = \frac{A^9(A-1)(2A-1)^2}{(A^2-3A+1)^5},$$

and then $A' = v_{A'}/u$ with

$$\begin{aligned} u = {} & 1 + 3\alpha + \frac{4A-1}{A}\alpha^2 + \frac{2c}{b}\alpha^3 - \frac{c(A-4)}{bA}\alpha^4 + \frac{(A-1)(4A-1)}{bA}\alpha^5 + \\ & \frac{(A+1)(A-1)}{bA^2}\alpha^6 + \frac{4c(A-1)}{b^2A}\alpha^7 + \frac{c(A-1)(4A-1)}{b^2A^2}\alpha^8 - \frac{c^2(A-1)}{b^3A}\alpha^9, \\ v_{A'} = {} & A + 2\alpha + \frac{A+1}{A}\alpha^2 + \frac{3c}{b}\alpha^3 + \frac{c(A+1)}{bA}\alpha^4 + \frac{(A-1)(A+1)}{bA}\alpha^5 + \\ & \frac{(A+1)(4A-1)}{bA^2}\alpha^6 + \frac{c(A-1)}{b^2A}\alpha^7 + \frac{c(A+1)(A-1)}{b^2A^2}\alpha^8 + \frac{c^2(A-1)}{b^3A}\alpha^9. \end{aligned}$$

7.6 Ensuring horizontality

If both E and P are defined over a finite field $\mathbf{F}_q$ with $\gcd(q-1, N) = 1$ then, as discussed in [6, §5.1], the isogeny $\varphi : E \to E' = E/\langle P \rangle$ is necessarily horizontal. The radicand $\rho_N \in \mathbf{F}_q$ admits a unique Nth root $\alpha \in \mathbf{F}_q$, and for this choice of α the resulting point $P' \in E'$ is again defined over $\mathbf{F}_q$, so the argument repeats. Thus, if N and $q-1$ are coprime, then walking horizontally using radical isogenies is natural and easy. As explained in Remark 7.2.1, for any fixed N the cost of an iteration is dominated by this Nth root extraction, which amounts to one exponentiation in $\mathbf{F}_q$. But if $\gcd(q-1, N) > 1$ then maintaining horizontality is more subtle.

In the remainder of this section we focus on the CSIDH case of supersingular elliptic curves over a finite prime field $\mathbf{F}_p$, where this issue arises (only) if $p \equiv 7 \bmod 8$ and one navigates with cyclic isogenies of even degree N, see [13, Thm. 2.7]. In this case $\gcd(p-1, N) = 2$ because $N \mid \#E(\mathbf{F}_p) = p+1$. Let us recall that if $p \equiv 7 \bmod 8$ then supersingular elliptic curves over $\mathbf{F}_p$ come in two kinds: curves on the surface of their 2-isogeny volcano, and curves on the floor. The surface is characterized by the existence of three $\mathbf{F}_p$-rational points of order 2; more precisely, the group of $\mathbf{F}_p$-rational points is isomorphic to $\mathbf{Z}_2 \times \mathbf{Z}_{(p+1)/2}$. The points of order 2 can be classified as follows (see Figure 7.1):

- a point $P_{\to}$, whose halves are $\mathbf{F}_p$-rational,
- a point $P_{\leftarrow}$, whose halves are not $\mathbf{F}_p$-rational, but their x-coordinates are,
- a point $P_{\downarrow}$, the x-coordinates of whose halves are not $\mathbf{F}_p$-rational.

	Previous work [6]	This work	Cost per 2-isogeny
2-isogeny	-	$\mathbf{E} + \mathbf{M} + 3\mathbf{m} + 2\mathbf{A}$	1
3-isogeny	$\mathbf{E} + 6\mathbf{M} + 3\mathbf{A}$	$\mathbf{E} + 2\mathbf{M} + 3\mathbf{m} + 3\mathbf{A}$	1.023
4-isogeny	$\mathbf{E} + 4\mathbf{M} + 3\mathbf{A} + \mathbf{I}$	$\mathbf{E} + 3\mathbf{M} + \mathbf{m} + 3\mathbf{A} + \mathbf{I}$	1.008
5-isogeny	$\mathbf{E} + 7\mathbf{M} + 6\mathbf{A} + \mathbf{I}$	$\mathbf{E} + 6\mathbf{M} + \mathbf{m} + 6\mathbf{A} + \mathbf{I}$	1.034
6-isogeny	-	$\mathbf{E} + 9\mathbf{M} + 6\mathbf{m} + 9\mathbf{A} + \mathbf{I}$	1.090
7-isogeny	$\mathbf{E} + 24\mathbf{M} + 20\mathbf{A} + \mathbf{I}$	$\mathbf{E} + 12\mathbf{M} + 2\mathbf{m} + 9\mathbf{A} + \mathbf{I}$	1.043
8-isogeny	-	$\mathbf{E} + 11\mathbf{M} + \mathbf{m} + 9\mathbf{A} + 2\mathbf{I}$	1.151
9-isogeny	$\mathbf{E} + 69\mathbf{M} + 58\mathbf{A} + \mathbf{I}$	$\mathbf{E} + 17\mathbf{M} + 9\mathbf{A} + \mathbf{I}$	1.062
10-isogeny	-	$\mathbf{E} + 57\mathbf{M} + 5\mathbf{m} + 31\mathbf{A} + 3\mathbf{I}$	1.196
11-isogeny	$\mathbf{E} + 599\mathbf{M} + 610\mathbf{A} + \mathbf{I}$	$\mathbf{E} + 50\mathbf{M} + 21\mathbf{m} + 71\mathbf{A} + 2\mathbf{I}$	1.293
12-isogeny	-	$\mathbf{E} + 90\mathbf{M} + 8\mathbf{m} + 35\mathbf{A} + 3\mathbf{I}$	1.296
13-isogeny	$\mathbf{E} + 783\mathbf{M} + 776\mathbf{A} + \mathbf{I}$	$\mathbf{E} + 89\mathbf{M} + 33\mathbf{m} + 120\mathbf{A} + 2\mathbf{I}$	1.448
14-isogeny	-	$\mathbf{E} + 159\mathbf{M} + 16\mathbf{m} + 131\mathbf{A} + 4\mathbf{I}$	1.613
15-isogeny	-	$\mathbf{E} + 149\mathbf{M} + 32\mathbf{m} + 125\mathbf{A} + 2\mathbf{I}$	1.599
16-isogeny	-	$\mathbf{E} + 120\mathbf{M} + 4\mathbf{m} + 40\mathbf{A} + 3\mathbf{I}$	1.388
17-isogeny	-	$\mathbf{E} + 217\mathbf{M} + 55\mathbf{m} + 332\mathbf{A} + 3\mathbf{I}$	1.921
19-isogeny	-	$\mathbf{E} + 329\mathbf{M} + 125\mathbf{m} + 437\mathbf{A} + 3\mathbf{I}$	2.532

Table 7.1: The computational cost of radical N-isogenies for $N \in \{2, 3, \ldots, 17\} \cup \{19\}$ compared to previous work [6, Tbl. 3]. The letters $\mathbf{E}, \mathbf{M}, \mathbf{A}$ and $\mathbf{I}$ denote exponentiation, (full) multiplication (including squaring), addition and inversion respectively. The letter $\mathbf{m}$ denotes multiplication with a small constant. The last column expresses the cost of an N-isogeny relative to a 2-isogeny, based on the evaluation of a chain of 100 000 horizontal N-isogenies over $\mathbf{F}_p$, where p is the CRAD-513 prime from Section 7.7. Remark that the cost of $\mathbf{E}$ is approximately $(1.5 \log p)\mathbf{M}$ with the square-and-multiply algorithm. In particular, the last column would converge to 1 for larger values of p since the cost of a radical isogeny will be dominated by $\mathbf{E}$.

Each of these points spans the kernel of a 2-isogeny. The point $P_\downarrow$ takes us to the floor, while the other two isogenies are horizontal. It can be checked that the dual of an isogeny in the $P_\rightarrow$-direction is in the $P_\leftarrow$-direction, and vice versa. Therefore, non-backtracking chains of horizontal 2-isogenies necessarily happen on the surface and consistently walk in either of these two directions.

7.6.1 Horizontal vs. non-horizontal N-isogenies

Fix $N \geq 2$ even and assume that $p \equiv -1 \bmod \operatorname{lcm}(2N, 8)$, so that every curve E on the surface satisfies

$$E(\mathbf{F}_p)[N] \cong \mathbf{Z}_2 \times \mathbf{Z}_N. \tag{7.9}$$

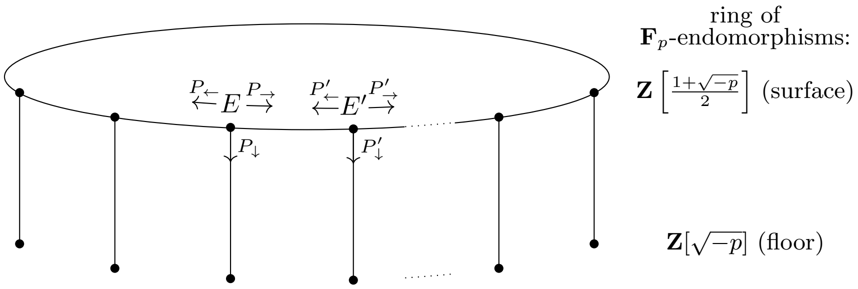

ring of
$\mathbf{F}_p$-endomorphisms:
$\mathbf{Z}\left[\frac{1+\sqrt{-p}}{2}\right]$ (surface)
$\mathbf{Z}[\sqrt{-p}]$ (floor)
$P_{\leftarrow}$
E
$P_{\rightarrow}$
$P'_{\leftarrow}$
E'
$P'_{\rightarrow}$
$P_{\downarrow}$
$P'_{\downarrow}$

Figure 7.1: *Component of the* 2*-isogeny graph over* $\mathbf{F}_p$ *when* $p \equiv 7 \bmod 8$*. The top layer belongs to the surface; the bottom layer belongs to the floor; and* $\sqrt{-p}$ *is identified with the Frobenius endomorphism.*

Then $E(\mathbf{F}_p)$ has 2 or 3 cyclic subgroups of order N, depending on whether $t = \operatorname{ord}_2(N) > 1$ or $t = 1$ (see Lemma 7.6.1 below). Every corresponding isogeny $\varphi : E \to E/\langle P\rangle$ can be decomposed as $\varphi = \theta \circ \psi$, where ψ is the $N/2$-isogeny with kernel $\langle 2P\rangle$ and θ is the 2-isogeny with kernel $\langle \psi(P)\rangle$. The isogeny ψ is necessarily horizontal: indeed, if it would involve a vertical step, then composing with θ would necessarily involve backtracking, rendering φ non-cyclic. However, θ may take us to the floor.

Lemma 7.6.1 *Write* $r = \operatorname{ord}_2(p+1) \geq \operatorname{ord}_2(2N) = t + 1$*.*

(i) If $t = 1$ *then there are* 3 *options for* $\langle P\rangle$*, corresponding to* θ *being in the* $P_{\rightarrow}$*-direction, the* $P_{\leftarrow}$*-direction or the* $P_{\downarrow}$*-direction.*

(ii) If $t \geq 2$ *then there are* 2 *options for* $\langle P\rangle$*, corresponding to* θ *being in the* $P_{\rightarrow}$*-direction or the* $P_{\downarrow}$*-direction.*

(iii) If $r \geq t+2$ *(automatic if* $t = 1$*) then the group corresponding to* θ *being in the* $P_{\rightarrow}$*-direction can be characterized as follows: it is the unique group all of whose elements admit halves in* $E(\mathbf{F}_p)$*.*

Proof. (i) Under the isomorphism (7.9), the cyclic subgroups of order N are generated by $(0,1)$, $(1,1)$ or $(1,2)$. Note that the group $\langle 2P\rangle$ does not depend on this choice, hence neither does ψ. Necessarily, the three groups must then correspond to the three stated options for θ.

(ii) If $t \geq 2$ then only the groups generated by $(0,1)$ or $(1,1)$ remain. Also note that we can further decompose $\psi = \theta' \circ \psi'$, where θ' is a 2-isogeny with kernel $\langle \psi'(2P)\rangle$. Since $\psi'(2P)$ is halvable over $\mathbf{F}_p$, this isogeny is necessarily in the $P_{\rightarrow}$-direction. But then θ cannot be in the $P_{\leftarrow}$-direction, otherwise φ would be non-cyclic.

(iii) If $r \geq t+2$ then $E(\mathbf{F}_p)[2N] \cong \mathbf{Z}_2 \times \mathbf{Z}_{2N}$ from which we see that the group generated by $(0,1)$ under the isomorphism (7.9) is uniquely characterized by its elements being halvable over $\mathbf{F}_p$. But then $\psi(P)$ is also halvable over $\mathbf{F}_p$, from which the claim follows. □

The central question of Section 7.6 is: how do we avoid that $\ker\theta = \langle P_\downarrow \rangle$, within the framework of radical isogenies?

7.6.2 Square vs. non-square radicands

As explained in [6, §5.3], there is a simple algebraic criterion for determining whether quotienting out an order-N point $P \in E$ keeps us on the surface or takes us to the floor. Namely, we stay on the surface if and only if $\rho_N = f_{N,P}(-P)$ is a non-zero square in $\mathbf{F}_p$. In this case ρ_N admits two different Nth roots $\alpha \in \mathbf{F}_p$, which are each other's negatives. The challenge is to select the sign in such a way that the next radicand ρ'_N is again a square. Indeed, for this choice of Nth root the argument repeats and one keeps walking horizontally. Of course, one fallback is to make an arbitrary choice for α, at the cost of an exponentiation in $\mathbf{F}_q$ as before. One then computes the resulting ρ'_N and checks if it is a square. If it is not, then one switches to $-\alpha$.

It was observed in [4, Lem. 4] that for $N = 2$ the extra quadratic residuosity check can be avoided, because the correct choice of α admits an explicit description in terms of the "principal" square root of ρ_2, by which we mean the unique square root which is itself a square.

Remark 7.6.2 More generally, for any non-zero square $\rho \in \mathbf{F}_p$ we will refer to the unique Nth root of ρ that is a square as the principal Nth root. Note that when computing the Nth root through exponentiation, i.e., as $\rho^{(p+1)/2N}$, then it is automatically principal. ◊

Then, in more detail, the observation from [4, Lem. 4] was as follows: the radical isogeny iteration

$$E : y^2 = x^3 + Ax^2 + Bx \quad \to \quad E' : y^2 = x^3 + (A + 6\alpha)x^2 + 4\alpha(A + 2\alpha)x,$$

with $\alpha = \sqrt{B}$, repeatedly quotients out $(0, 0)$. If $(0, 0) \in E$ is the point $P_\rightarrow$, then $(0, 0) \in E'$ is the point $P'_\rightarrow$ if and only if α is the principal square root. This changes if $(0, 0) \in E$ is the point $P_\leftarrow$, in which case $(0, 0) \in E'$ is the point $P'_\leftarrow$ if and only if α is the non-principal square root.

This convenient fact was adapted to $N = 4$, first as a conjecture [6, Conj. 2] but recently this got proved by Onuki and Moriya [17, §5]. We will recall the precise statement of this adaptation in Section 7.6.4, where it will arise as an easy consequence to our generalization to arbitrary even N. But let us first highlight two takeaways that are already apparent from the case $N = 2$:

(i) When considering radical isogeny formulae for even N, then substituting $-\alpha$ for α produces formulae that are equally legitimate, e.g., because -1 is an Nth root of unity. Consequently, one cannot hope for a general rule saying that the $P_\rightarrow$-direction always corresponds to the principal Nth root.

(ii) Even worse, imagine that the rule does apply to some concrete choice of formulae, and now scale the radicand ρ_N with g^N for some arbitrary modular unit $g \in \mathbf{Q}_N(b,c)$, i.e. a function whose zeroes and poles are supported on the cuspidal part of $X_1(N)$; see [23]. The radical isogeny formulae transform into a version in which each occurrence of $\sqrt[N]{\rho_N}$ gets replaced by $\sqrt[N]{\rho_N}\,/g$. For these new formulae, the correct Nth root will depend on the Legendre symbol of the evaluation of g at the point $(E,P) \in X_1(N)$ under consideration.

7.6.3 Conjectural shape of ρ'_N modulo squares (proved for $N \leq 14$)

We ran into the following property of ρ'_N, which unfortunately we could not prove beyond $N = 14$, but which implies a generalization of the aforementioned observations for $N = 2,4$ to arbitrary even N. Concretely, for every even $N \geq 4$ we can consider

$$\phi_{E,2}(x) = x^4 + b(1-c)x^2 - 2b^2x + b^3, \tag{7.10}$$

whose roots are the x-coordinates of the four halves of $P = (0,0)$ on $E = E_{b,c}$. Over $\mathbf{Q}_N(b,c)(\alpha^{N/2})$ this polynomial splits in two quadratic factors, with one quadratic factor corresponding to a pair of points

$$\frac{N}{2}Q, \; \frac{N}{2}Q + \frac{N}{2}P,$$

mapping to $\frac{N}{2}P'$ under φ. The discriminant of said quadratic factor is a modular unit of $X'_1(N)$ that we denote by Δ.

Example 7.6.3 Over $\mathbf{Q}_4(b,c)(\alpha^2)$ the polynomial (7.10) splits as $(x^2-\alpha^2x-\alpha^6)(x^2+\alpha^2x+\alpha^6)$. The roots of the first factor are the x-coordinates of two preimages of $2P'$. The discriminant of that factor is $\Delta = \alpha^4(1+4\alpha^2)$. ☆

Our conjecture is as follows:

Conjecture 7.6.4 *If the radicand $\rho_N = f_{N,P}(-P)$ was chosen, then one has*

$$\rho'_N \equiv \sigma\alpha b\Delta \tag{7.11}$$

modulo multiplication with a non-zero square in $\mathbf{Q}_N(b,c)(\alpha)$, for some $\sigma \in \{\pm 1\}$.

Here, we note:

- The sign σ should be viewed against our first takeaway message *(i)* above: substituting $-\alpha$ for α produces equally valid radical isogeny formulae but flips the sign.
- The congruence sign absorbs squares, so the conjecture is insensitive to replacing ρ'_N with any other representative of $t_N(P', -P')$, or even $t_N(P', \lambda P')$ for

whatever odd λ. However, as discussed in our second takeaway *(ii)* above, in the case of ρ_N the precise representative does matter. Interestingly, scaling with b^N would make the statement somewhat cleaner, as it would remove the mysterious factor b. This suggests that the radicand from Theorem 7.7.1 in the appendix is in fact a more natural choice than $f_{N,P}(-P)$.

It is exactly the presence of this factor b that made it difficult to guess how to go beyond the case $N = 4$; in the case $N = 4$ we have $b = -\alpha^4$ so that modulo squares this factor just appeared as a sign.

Theorem 7.6.5 *Conjecture 7.6.4 is true for* $N \leq 14$.

Proof sketch. From (7.7) and Example 7.6.3 we see that $\rho_4' \equiv \alpha\Delta$ modulo squares, which matches with Conjecture 7.6.4 with $\sigma = -1$ because $-b = \alpha^4$ is a square. So the case $N = 4$ is immediate.

The case $N = 6$ is more illustrative. Take A', α, b as in (7.8) and let

$$\Delta = 4(1 - A)\alpha^3 + 3A^3 - 7A^2 + 4A$$

be the discriminant of the relevant quadratic factor of (7.10). One verifies, aided by the Magma command `IsPrincipal`, that for $\rho_6' = f_{6,P'}(-P') = -A'^2(A' - 1)$ the function $-b\rho_6'/\alpha\Delta$ is a square in the function field of $X_1'(6) : \alpha^6 + A^2(A - 1) = 0$. So this again matches with Conjecture 7.6.4 (now with a minus sign).

In a similar way we have managed to deal with all even N up to 14, with further help coming from the observation that $\rho_N' = f_{N,-P'}(-P') \equiv f_{2,\frac{N}{2}P'}(P')$ modulo squares, see [3, Thm. IX.9(2)]. The right-hand side is a simpler function and therefore easier to handle by Magma. As an example, the Magma code for $N = 14$ can be found in the GitHub repository.[3] □

As mentioned, beyond $N = 14$ we were no longer able to verify Conjecture 7.6.4, although for $N = 16$ we gathered evidence by experimentally verifying Proposition 7.6.6 below for various concrete horizontal supersingular isogeny walks over finite prime fields.

7.6.4 Horizontal isogenies and principal Nth roots

Proposition 7.6.6 *Let $N \geq 4$ be even and consider radical isogeny formulae for computing chains of N-isogenies in terms of the radicand $\rho_N = f_{N,P}(-P)$. Assume that Conjecture 7.6.4 applies to these formulae and let $\sigma = \pm 1$ be the sign involved in its statement.*

Let $p \equiv -1 \bmod \operatorname{lcm}(2N, 8)$ and consider a supersingular elliptic curve $E/\mathbf{F}_p$ on the surface, along with a point $P \in E(\mathbf{F}_p)[N]$ such that the resulting isogeny $\varphi : E \to E' = E/\langle P \rangle$ is horizontal; let θ be the corresponding degree-2 component as in Section 7.6.1 and let $b, c \in \mathbf{F}_p$ be the corresponding Tate normal form coefficients. Let

[3] `https://github.com/KULeuven-COSIC/Horizontal_Radical_Isogenies`

$P' \in E'$ be the point produced by our radical isogeny formulae, where $\alpha = \sqrt[N]{\rho_N(b,c)}$ was computed as

$$\sigma \cdot s \cdot b^{(p-1)/2} \rho_N(b,c)^{(p+1)/2N}.$$

Here the sign s is determined as follows:

(i) if θ walks in the $P_{\rightarrow}$-direction and $r > t+1$ then $s = 1$,

(ii) if θ walks in the $P_{\rightarrow}$-direction and $r = t+1$ then $s = -1$,

(iii) if θ walks in the $P_{\leftarrow}$-direction (only possible if $t = 1$) then $s = -1$.

Then the isogeny $E' \to E'/\langle P' \rangle$ is horizontal.

Proof. Recall that the goal is to choose the instance of α that renders ρ'_N a square. Assuming Conjecture 7.6.4, this happens if and only if $\sigma\alpha b\Delta$ is a square.

In case *(i)* the point P is fully halvable over $\mathbf{F}_p$ thanks to Lemma 7.6.1(iii), so that Δ always evaluates to a square, regardless of the choice of α. So in order for ρ'_N to be a square, it is necessary and sufficient to choose α such that $\sigma\alpha b$ is a square: the claim follows.

If we are in cases *(ii)* or *(iii)* then none of the halves of P belong to $E(\mathbf{F}_p)$. Even stronger: none of these halves can have an $\mathbf{F}_p$-rational x-coordinate, because otherwise such a half H would satisfy $\pi_p(H) = -H$ and therefore $P = \pi_p(P) = -P$; a contradiction. This means that Δ is a non-square, regardless of the choice of α, and we can conclude as before. □

Example 7.6.7 For $N = 4$ we recover [6, Conj. 2], proved in [17]. Indeed, recall that $\sigma = -1$ and that b is always non-square in view of $\rho_4 = -b = \alpha^4$. Thus we have to compute $\alpha = s\rho_4^{(p+1)/8}$ with $s = -1$ if $p \equiv 7 \bmod 16$ and $s = 1$ if $p \equiv 15 \bmod 16$. ☆

We conclude by noting that $b^{(p-1)/2}\rho^{(p+1)/2N} = b^{-1}(b^N\rho_N)^{(p+1)/2N}$, effectively showing that the cost of root computation remains a single exponentiation.

7.7 Implementation

In this section we focus on N-isogenies between supersingular elliptic curves over prime fields $\mathbb{F}_p$ such that computing the required radical can be done deterministically by a single exponentiation. All tests were done in Magma v2.32-2 on an Intel(R) Xeon(R) CPU E5-2630 v2 @ 2.60GHz with 128 GB memory.

7.7.1 Isogeny chains

The main application of these radical isogeny formulae is that they can be used to efficiently compute a cyclic N^k-isogeny for small N and large k. This is similar to the work in [6], but we can now use larger N, have more efficient formulae for smaller N and are not restricted to odd N.

Remark that the radical 5-isogeny formulae from [6] were already optimized. Table 7.1 however shows a modest to strong speed up for radical N-isogenies for $N = 7, 9, 11, 13$. Over the field $\mathbb{F}_p$ with p the 513-bit CRAD-513 prime from Section 7.7.2, they provide a speed-up of respectively 4%, 13%, 55% and 57% compared to the work of [6].

The best known method to compute a chain of 17- or 19-isogenies so far was by sampling 17- or 19-torsion points and then applying Vélu-style formulae to compute the codomain. The cost of this is dominated by the computation of an appropriate torsion point. With the new radical formulae from Section 7.5, we only need to initialize the chain by computing such a torsion point once, and then can iteratively apply the radical isogeny formulae. Working over a prime field of roughly 512 bits, this results in an asymptotic speed-up of chaining 17-isogenies by a factor of 14, and a factor of 10 for chaining 19-isogenies. There is somewhat of a jump in complexity when going to optimized equations from $X_1(19)$ to $X_1(23)$ due to a jump in gonality. In particular, we do not expect radical 23-isogenies to be much of a speed-up over prime fields of characteristic roughly 512 bits,[4] so we did not try to optimize these. Nonetheless, for asymptotically large p the computational cost of a radical isogeny is expected to be dominated by a full exponentiation over $\mathbb{F}_p$.

For composite N, one can make a similar argument with regards to speed-up but the comparison is more subtle. For instance, the cost of computing a 15-isogeny is dominated by one exponentiation and 149 full multiplications according to Table 7.1. Alternatively, a 15-isogeny can also be computed by means of the concatenation of a 3- and 5-isogeny, the cost of which is dominated by two exponentiations and 8 full multiplications. Assuming we work over a prime field of cryptographic size - say at least 128 bits - the 15-isogeny will be the fastest method. However, assuming we have rational 9-torsion available, we have access to highly efficient radical 9-isogeny formulae, so asymptotically a 3-isogeny can be seen as half the cost of a 9-isogeny.

	512 bits	1024 bits	1536 bits
$2^{60,000}$-isogeny	23.38s	97.42s	264.59s
$4^{30,000}$-isogeny	11.93s	49.51s	133.12s
$8^{20,000}$-isogeny	8.77s	34.58s	91.33s
$16^{15,000}$-isogeny	7.92s	29.23s	75.01s
$3^{60,000}$-isogeny	23.39s	98.08s	266.31s
$9^{30,000}$-isogeny	12.77s	49.88s	134.61s

Table 7.2: Comparison in speed with regards to computing a chain of radical ℓ-isogenies over a prime field $\mathbb{F}_p$ for $\ell \in \{2, 3\}$ by means of different prime powers. The bit levels correspond to the size of p.

In general, composite N seem to yield more efficient formulae compared to prime

[4] Especially in the CSIDH setting from Section 7.7.2 where the initializing overhead is less negligible.

N as can be seen in Table 7.1. This stems from the fact that optimized equations for $X_1(N)$ typically have lower degree when N is composite, but also from the radical isogeny formulae themselves which appear to have parameterless integer coefficients (including zero) noticeably more often for composite N. These zero coefficients are even more frequently present in the radical isogeny formulae for prime-power N. In Table 7.2 one can see a comparison for computing low-degree prime-power chains of isogenies for three levels of prime bitsizes.

As can be seen, computing a chain of prime-power degree isogenies can be done more efficiently than a chain of prime degree isogenies for at least these values. The effect is more prominent for larger prime fields, since the exponentiation in those cases is more dominating in the overall cost of the radical isogeny formulae. We did not optimize the formulae for $N = 25$, since an optimal parametrization of $X_1(25)$ is already more complex than $X_1(19)$, and from Table 7.1 it is clear that computing chains of 5-isogenies would most likely be just as fast or faster (at least on the 512-bit level). Assuming the arithmetic for a radical ℓ^{k+1}-isogeny is always more complex than the arithmetic for a radical ℓ^k-isogeny, the asymptotic speed-up that can be gained from going to the next prime power is always bounded by $(k+1)/k$. For this reason, we expect that optimized radical 27- and 32-isogenies would be less efficient than radical 9- and 16-isogenies for all bitsizes in Table 7.2, though from a certain threshold onwards they would be the most efficient option again.

7.7.2 Impact on CSIDH

An application where chains of isogenies can be used is CSIDH [7]. We proceed just as in [6, §6], with the following differences:

- We make use of radical 17- and 19-isogenies.
- The optimzed formulae allow us to sample higher exponents of N-isogenies for $N = 7, 9, 11, 13$.
- We no longer use radical 4-isogenies, instead switching to radical 8-isogenies.

This last point may seem counterintuitive considering that chains of 16-isogenies are faster on the 512-bit prime level, as illustrated in Table 7.2. In CSIDH however, p is chosen such that $p+1$ is divisible by as many small primes as possible. If we want to make use of radical 16-isogenies, we would need to have that $32 \mid p+1$ (instead of $16 \mid p+1$ for radical 8-isogenies). This means that p would need to be roughly one bit larger, making all the other arithmetic more expensive. The trade-off in practice seems to be not worth it, considering the relative small gain from switching from chains of radical 8-isogenies to chains of radical 16-isogenies. The gap in efficiency between radical 4-isogenies and radical 8-isogenies does make a noticeable difference so we will use those. Nonetheless, we still need an extra factor of 2 that divides $p+1$ compared to the suggested prime in [6], so we choose CRAD-513 as the prime

$$p = 2^4 \cdot 3 \cdot \underbrace{(3 \cdot 5 \cdot \ldots 367)}_{\text{72 consecutive primes}} \cdot 379 \cdot 409 - 1.$$

The following sampling interval for the private key was determined heuristically, but can be considered (near) optimal:

$$\begin{aligned}&[-303;303] \times [-198;198] \times [-103;103] \times [-101;101] \times [-91;91]\\&\times [-68;68] \times [-51;51] \times [-41;41] \times [-6;6]^{13} \times [-5;5]^{13}\\&\times [-4;4]^{11} \times [-3;3]^{10} \times [-2;2]^{10} \times [-1;1]^{10}.\end{aligned}$$

Using these parameters, the class group action of the maximal private key can be computed 12% more efficiently than in the case of [6]. For an average private key, this speed-up will be roughly halved but from a constant-time implementation angle, the maximal private key is a more apt benchmark. This implementation in Magma is meant as a comparison to the work of [6], and can not be translated directly to other (constant-time) implementations such as CTIDH [1].

Appendix: an explicit radicand

The goal of the appendix is to prove the following result.

Theorem 7.7.1 *Let $N \in \mathbf{Z}_{>2}$. Let $K = \mathbf{Q}_N(b,c)$ as in Section 7.2.3. Let E/K be the elliptic curve given by $y^2 + (1-c)xy - by = x^3 - bx^2$. Let $P = (0,0) \in E$. Denote by Ψ_j the j-th division polynomial on E. Set $k = \lceil N/2 \rceil$. Then*

$$\left(\sum_{S \in E[N]} e_N(P,S)x(Q+S) \right)^N = N^{2N} \cdot \begin{cases} \frac{\Psi_k^2}{\Psi_{k-1}^2}(P) & \text{if } N \text{ is odd;} \\ \frac{\Psi_{k+1}}{\Psi_{k-1}}(P) & \text{if } N \text{ is even.} \end{cases}$$

Pairings and division polynomials

Let K be a field and let E/K be an elliptic curve. Suppose $P \in E(K)$ is of order N, such that $\operatorname{char} K \nmid N$. Let $Q \in E(\overline{K})$ satisfying $NQ = P$. Let $f \in K(E)$, $g \in \overline{K}(E)$ with respective divisors

$$\operatorname{div} f = N(P) - N(\mathcal{O}), \qquad \operatorname{div} g = \sum_{S \in E[N]} ((Q+S) - (S)).$$

Assume that g is such that $g^N = f \circ [N]$. Denote by $e_N : E[N] \times E[N] \to \mu_N$ the Weil pairing and by $t_N : E(K)[N] \times E(K)/NE(K) \to K^\times/(K^\times)^N$ the Tate pairing. For $\mathcal{P} \in E$, denote by $\tau_{\mathcal{P}} : E \to E$ the translation-by-$\mathcal{P}$ map. Let $\omega \in \Omega_E$ be an invariant differential and denote by $\operatorname{res}_{\mathcal{P}}(-) : \Omega_E \to \overline{K}$ the residue at $\mathcal{P}$ as defined in [25].

Lemma 7.7.2 *For every $Q \in E(K)$ we have*

$$t_N(P,Q) = \frac{\text{“Leading coefficient of } f \text{ at } Q\text{”}}{\text{“Leading coefficient of } f \text{ at } \mathcal{O}\text{”}} \in K^\times/(K^\times)^N.$$

Remark 7.7.3 Note that the leading coefficient of f (meaning the leading coefficient of the expansion of f with respect to a uniformizer) is everywhere well defined up to Nth powers, since the order of vanishing of f is at every point divisible by N (hence a different choice of uniformizer scales the leading coefficient by an Nth power). Also, the quotient in Lemma 7.7.2 is invariant under scaling f by an element of K, hence well-defines an element of $K^\times/(K^\times)^N$ given only the divisor of f. ◊

Proof. If $P = \mathcal{O}$ or $Q = \mathcal{O}$ then both sides are equal to 1, so assume $P \neq \mathcal{O} \neq Q$. We distinguish two cases.

Case $P = Q$. Let $h \in K(E)$ be any function such that $\operatorname{ord}_P(h) = -1$ and $\operatorname{ord}_{\mathcal{O}}(h) = 1$. Then $t_N(P,P) = f(\operatorname{div}(h) + (P) - (\mathcal{O}))$. By Weil reciprocity

$$\prod_R (-1)^{\operatorname{ord}_R(f)\operatorname{ord}_R(h)} \frac{f^{\operatorname{ord}_R(h)}}{g^{\operatorname{ord}_R(f)}}(R) = (-1)^{-2N} \frac{f^{-1}}{h^N} \frac{f^1}{h^{-N}}(P) \prod_{R \neq P,\mathcal{O}} f^{\operatorname{ord}_R(h)}(R).$$

equals 1. Hence

$$t_N(P,P) = \prod_{R\neq P,\mathcal{O}} f^{\mathrm{ord}_R(h)}(R) = \frac{h^N f(P)}{h^N f(\mathcal{O})} \in K^\times/(K^\times)^N.$$

Case $P \neq Q$. Let $h \in K(E)$ be any function such that $\mathrm{ord}_P(h) = 0$, $\mathrm{ord}_Q(h) = -1$, $\mathrm{ord}_{\mathcal{O}}(h) = 1$. Then $t_N(P,Q) = f(\mathrm{div}(h) + (Q) - (\mathcal{O}))$. By Weil reciprocity

$$1 = \prod_R (-1)^{\mathrm{ord}_R(f)\,\mathrm{ord}_R(h)} \frac{f^{\mathrm{ord}_R(h)}}{g^{\mathrm{ord}_R(f)}}(R) = (-1)^{-N}\frac{f}{h^{-N}}(\mathcal{O})\frac{\prod_{R\neq\mathcal{O}} f^{\mathrm{ord}_R(h)}(R)}{h^N(P)}$$

Hence $t_N(P,Q)$ can be rewritten as

$$f(Q)\prod_{R\neq\mathcal{O}} f^{\mathrm{ord}_R(h)}(R) = (-1)^N\frac{h^N(P)}{(h^N f)(\mathcal{O})}f(Q) = \frac{f(Q)}{(h^N f)(\mathcal{O})} \in K^\times/(K^\times)^N.$$

□

Lemma 7.7.4 *Let $R \in E[N]$ such that P, R generate $E[N]$. We have*

$$t_N(P,P) = \left(\sum_{i,j=0}^{N-1} e_N(P,R)^i x(Q + iR + jP)\right)^N \quad \text{in } K^\times/(K^\times)^N.$$

Proof. We rely on the residue theorem [25, Thm. 3], whose use was suggested to us by Alexander Lemmens. This theorem implies that $\sum_{\mathcal{P}\in E}\mathrm{res}_{\mathcal{P}}(xg^{-1}\omega) = 0$, therefore

$$\begin{aligned} -\mathrm{res}_{\mathcal{O}}(xg^{-1}\omega) &= \sum_{S\in E[N]} \mathrm{res}_{Q+S}(xg^{-1}\omega) \\ &= \sum_{S\in E[N]} x(Q+S)\frac{g}{g\circ\tau_S}(Q)\,\mathrm{res}_Q(g^{-1}\omega) \\ &= \mathrm{res}_Q(g^{-1}\omega)\sum_{S\in E[N]} e_N(P,S)x(Q+S). \end{aligned}$$

It follows that (the last equivalence is due to Lemma 7.7.2)

$$\begin{aligned} \left(\sum_{S\in E[N]} e_N(P,S)x(Q+S)\right)^N &= (-1)^N\frac{x^N(g^N\circ\tau_Q)}{g^N}(\mathcal{O}) \\ &= (-1)^N\frac{x^N}{x^N\circ[N]}\frac{(x^N\circ[N])(f\circ[N]\circ\tau_Q)}{f\circ[N]}(\mathcal{O}) \\ &= (-1)^N N^{2N}\frac{x^N(f\circ\tau_P)}{f}(\mathcal{O}) \end{aligned}$$

which equals $t_N(P,P)$ in $K^\times/(K^\times)^N$. □

Now let $K = \mathbf{Q}(b,c)$, where b and c are both transcendental over $\mathbf{Q}$, though possibly algebraically dependent. Let E/K be the elliptic curve given by $y^2 + (1-c)xy - by = x^3 - bx^2$ and set $P := (0,0) \in E$.

For $Q \in E(K)$, we denote by $h_{P,Q} \in K(E)^\times$ any function with divisor $(P)+(Q)-(P+Q)-(\mathcal{O})$. For $j \in \mathbf{Z}$, we define

$$L_j := \left(\left(\frac{x}{y}\right)^{\operatorname{ord}_{\mathcal{O}}(h_{P,jP})-\operatorname{ord}_P(h_{P,jP})} \cdot \frac{h_{P,jP}\circ\tau_P}{h_{P,jP}}\right)(\mathcal{O}).$$

In other words, L_j is the leading coefficient at $\mathcal{O}$ of the Laurent expansion of the function $(h_{P,jP}\circ\tau_P)/h_{P,jP}$ with respect to the uniformizer x/y. Note that, whereas $h_{P,Q}$ is only well-defined up to scalar multiplication, we have that L_j is a well-defined element of $K^\times$.

Lemma 7.7.5 *We have*

$$L_j = \begin{cases} b & \textit{if } jP = -2P \textit{ or } jP = -P; \\ 1 & \textit{if } jP = \mathcal{O}; \\ -b & \textit{if } jP = P; \\ b\cdot\dfrac{y_{jP}}{x_{jP}\cdot x_{(j+1)P}} & \textit{else.} \end{cases}$$

Proof. Using (note that $h_{P,Q}$ as given by the formula below indeed has the desired divisor)

$$h_{P,Q} = \begin{cases} x & \text{if } Q = -P; \\ 1 & \text{if } Q = \mathcal{O}; \\ \dfrac{y}{x - x_{2P}} & \text{if } Q = P; \\ \dfrac{y-(y_Q/x_Q)x}{x-x_{P+Q}} & \text{else,} \end{cases}$$

this is a straightforward check for $Q \in \{-2P, -P, \mathcal{O}, P\}$. If $Q \notin \{-2P,-P,\mathcal{O},P\}$ then in particular $x_{P+Q} \neq 0$. Let $u = x/y$. Then $x\circ\tau_P = bu + O(u^2)$ and $y\circ\tau_P = O(u^2)$, while $x = u^{-2} + O(u^{-1})$ and $y = u^{-3} + O(u^{-2})$. Thus the leading term at $\mathcal{O}$ of $(h_{P,Q}\circ\tau_P)/h_{P,Q}$ becomes

$$\frac{-y_Q/x_Q\cdot b}{-x_{P+Q}} = b\cdot\frac{y_Q}{x_Q\cdot x_{Q+P}}$$

as claimed. □

In what follows, $N > 2$ will always denote an integer and $k = \lceil N/2\rceil$. We will assume that b, c are such that P has order at least $k+1$. Let $f \in K(E)$ be any function with divisor $N(P) - N(\mathcal{O}) + ((k-N)P) - (kP)$.

Lemma 7.7.6 *We have*

$$\left(x^N \cdot \frac{f \circ \tau_P}{f}\right)(\mathcal{O}) = \prod_{j=-\lfloor N/2 \rfloor}^{\lfloor (N-1)/2 \rfloor} L_j.$$

Proof. This follows by noting that

$$\left(\left(\frac{x}{y}\right)^{2N} \cdot x^N\right)(\mathcal{O}) = 1.$$

and that $f = \prod_{j=-\lfloor N/2 \rfloor}^{\lfloor (N-1)/2 \rfloor} h_{P,jP}$ has the desired divisor. □

Define

$$\rho_N := \begin{cases} \dfrac{\Psi_k^2}{\Psi_{k-1}^2}(P) & \text{if } N \text{ is odd;} \\ \\ \dfrac{\Psi_{k+1}}{\Psi_{k-1}}(P) & \text{if } N \text{ is even,} \end{cases} \qquad \text{and} \qquad \pi(N) := \prod_{j=-\lfloor N/2 \rfloor}^{\lfloor (N-1)/2 \rfloor} L_j.$$

Lemma 7.7.7 *For all $N \in \mathbf{Z}_{>2}$, we have $\pi(N) = (-1)^N \rho_N$.*

Proof. We use induction on N. One easily verifies the claim for $N = 3, 4, 5$. Suppose $N = 2k \geq 6$ is even. Then

$$\pi(N)/\pi(N-1) = b \cdot \frac{y_{-kP}}{x_{-kP} \cdot x_{(-k+1)P}}, \quad \text{and} \quad \pi(N+1)/\pi(N) = b \cdot \frac{y_{kP}}{x_{kP} \cdot x_{(k+1)P}},$$

whereas $-\rho_N/\rho_{N-1} = -(\Psi_{k+1}\Psi_{k-1}/\Psi_k^2)(P) = -\rho_{N+1}/\rho_N$. But the middle term $-(\Psi_{k+1}\Psi_{k-1}/\Psi_k^2)(P)$ can be rewritten as $x_{kP} = x_{-kP}$ (from the multiplication-by-k formula using division polynomials; e.g. [20, Ex. 3.7]), so we can conclude using Lemma 7.7.8. □

Lemma 7.7.8 *For all $k \in \mathbf{Z} \setminus \{-1, -2\}$, we have $x_{kP}^2 x_{(k+1)P} = b \cdot y_{kP}$.*

Proof. Using the coordinate-wise addition formula for Weierstrass elliptic curves (e.g. [20, III.2.3]), we find $x_{kP}^2 x_{(k+1)P} = y_{kP}^2 + (1-c)x_{kP}y_{kP} + bx_{kP}^2 - x_{kP}^3 = by_{kP}$. □

Proof of Theorem 7.7.1. In the proof of Lemma 7.7.4, we already saw that the left hand side equals $(-1)^N N^{2N} \left(x^N \cdot \frac{f \circ \tau_P}{f}\right)(\mathcal{O})$. The desired result now follows by combining Lemmas 7.7.6 and 7.7.7. □

7.8 Bibliography

[1] Gustavo Banegas, Daniel J. Bernstein, Fabio Campos, Tung Chou, Tanja Lange, Michael Meyer, Benjamin Smith, and Jana Sotáková. CTIDH: faster constant-time CSIDH. *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, 2021(4):351–387, 2021.

[2] Ward Beullens, Thorsten Kleinjung, and Frederik Vercauteren. CSI-FiSh: Efficient isogeny based signatures through class group computations. In *Asiacrypt (1)*, volume 11921 of *Lecture Notes in Computer Science*, pages 227–247. Springer, 2019. `https://ia.cr/2018/485`.

[3] Ian F. Blake, Gadiel Seroussi, and Nigel P. Smart. *Elliptic Curves in Cryptography.* Cambridge University Press, United Kingdom, 1999.

[4] Wouter Castryck and Thomas Decru. CSIDH on the surface. In Jintai Ding and Jean-Pierre Tillich, editors, *PQCrypto 2020*, volume 12100 of *Lecture Notes in Computer Science*, pages 111–129. Springer, 2020.

[5] Wouter Castryck and Thomas Decru. Multiradical isogenies. In *AGC^2T-18*, Contemp. Math. (to appear). American Mathematical Society, 2022. `https://eprint.iacr.org/2021/1133`.

[6] Wouter Castryck, Thomas Decru, and Frederik Vercauteren. Radical isogenies. In *Proceedings of Asiacrypt 2020 Part II*, volume 12492 of *Lecture Notes in Computer Science*, pages 493–519. Springer, 2020.

[7] Wouter Castryck, Tanja Lange, Chloe Martindale, Lorenz Panny, and Joost Renes. CSIDH: An efficient post-quantum commutative group action. In *Asiacrypt 2018 Pt. 3*, volume 11274 of *Lecture Notes in Computer Science*, pages 395–427. Springer, 2018.

[8] Wouter Castryck, Jana Sotáková, and Frederik Vercauteren. Breaking the decisional diffie-hellman problem for class group actions using genus theory. In *Advances in Cryptology – CRYPTO 2020: 40th Annual International Cryptology Conference, CRYPTO 2020, Santa Barbara, CA, USA, August 17–21, 2020, Proceedings, Part II*, page 92–120, Berlin, Heidelberg, 2020. Springer-Verlag.

[9] Jesus-Javier Chi-Dominguez and Krijn Reijnders. Fully projective radical isogenies in constant-time. In *Topics in Cryptology – CT-RSA 2022: Cryptographers' Track at the RSA Conference 2022, Virtual Event, March 1–2, 2022, Proceedings*, page 73–95, Berlin, Heidelberg, 2022. Springer-Verlag.

[10] Jean-Marc Couveignes. Hard homogeneous spaces, 2006. Unpublished article, available at `https://eprint.iacr.org/2006/291`.

[11] Luca De Feo and Jeffrey Burdges. Delay encryption. In *Proceedings of Eurocrypt 2021 Part I*, volume 12696 of *Lecture Notes in Computer Science*, pages 302–326. Springer, 2021.

[12] Luca De Feo, Simon Masson, Christophe Petit, and Antonio Sanso. Verifiable delay functions from supersingular isogenies and pairings. In Steven D. Galbraith and Shiho Moriai, editors, *Advances in Cryptology – ASIACRYPT 2019*, pages 248–277, Cham, 2019. Springer International Publishing.

[13] Christina Delfs and Steven D. Galbraith. Computing isogenies between supersingular elliptic curves over $\mathbb{F}_p$. *Designs, Codes and Cryptography*, 78(2):425–440, 2016. `https://arxiv.org/abs/1310.7789`.

[14] Robert Granger, Florian Hess, Roger Oyono, Nicolas Thériault, and Frederik Vercauteren. Ate pairing on hyperelliptic curves. In Moni Naor, editor, *Advances in Cryptology - EUROCRYPT 2007*, pages 430–447, Berlin, Heidelberg, 2007. Springer.

[15] Yi-Fu Lai, Steven D. Galbraith, and Cyprien Delpech de Saint Guilhem. Compact, efficient and UC-secure isogeny-based oblivious transfer. In *Proceedings of Eurocrypt 2021 Part I*, volume 12696 of *Lecture Notes in Computer Science*, pages 213–241. Springer, 2021.

[16] Michael Monagan and Roman Pearce. Rational simplification modulo a polynomial ideal. In *ISSAC '06*, pages 239–245. ACM, 2006.

[17] Hiroshi Onuki and Tomoki Moriya. Radical isogenies on montgomery curves. In *Proceedings of PKC 2022 Part I*, volume 13177 of *Lecture Notes in Computer Science*, pages 473–497. Springer, 2022.

[18] David E. Rohrlich. Modular curves, Hecke correspondence, and L-functions. In *Modular forms and Fermat's last theorem*, pages 41–100. Springer, 1997.

[19] Samir Siksek. Explicit arithmetic of modular curves. Summer school notes, available at `https://homepages.warwick.ac.uk/staff/S.Siksek/teaching/modcurves/lecturenotes.pdf`, 2019.

[20] Joseph H. Silverman. *The arithmetic of elliptic curves*, volume 106 of *Graduate Texts in Mathematics*. Springer, second edition, 2009.

[21] The Stacks project authors. The Stacks project. Available at `https://stacks.math.columbia.edu`, 2021.

[22] Anton Stolbunov. Public-key encryption based on cycles of isogenous elliptic curves. Master's thesis, Saint-Petersburg State Polytechnical University, 2004. In Russian.

[23] Marco Streng. Generators of the group of modular units for $\Gamma^1(N)$ over the rationals. *Ann. H. Lebesgue*, 6:95–116, 2023.

[24] Andrew V. Sutherland. Constructing elliptic curves over finite fields with prescribed torsion. *Mathematics of Computation*, 81:1131–1147, 2012.

[25] John Tate. Residues of differentials on curves. *Ann. Sci. École Norm. Sup. (4)*, 1:149–159, 1968.

[26] Jacques Vélu. Isogénies entre courbes elliptiques. *Comptes Rendus de l'Académie des Sciences, Série I*, 273:238–241, 1971.

Chapter 8

Conclusion

The main part of this work, Chapters 4, 5, 6, and 7, consists of four jointly written research papers.

In Chapter 4, we described a novel way in which pairings on elliptic curves can be used to attack the Decisional Diffie–Hellman problem for class group actions on oriented elliptic curves. We showed how the assigned character values associated to connecting ideal classes can be evaluated using the Weil pairing. This was previously only established for the Tate pairing. Our approach works more generally, is conceptually simpler, and speeds up the previous approach in certain cases. The attack only applies in case the class number is even. As a consequence, we recommend to restrict CSIDH and CRS to class groups of odd order.

In Chapter 5, we classified when non-trivial self-pairings on cyclic subgroups compatible with isogenies oriented by an imaginary quadratic order exist. Combining such self-pairings together with isogeny interpolation leads to a new attack strategy against CRS in the case where the degree of the secret isogeny is known. As a result of our classification, this implies the existence of weak instances of CRS; ones in which the discriminant has a large square smooth divisor coprime to the field characteristic.[1] One way to surely mitigate these attacks, is to use a discriminant of the form $-p$ where p is prime. CSIDH, in which the discriminant is of the form $-4p$, also remains unaffected by the strategy. An interesting future question to explore is whether small divisors of the discriminant could be exploited to obtain partial information about the secret isogeny. Furthermore, for some non-trivial self-pairings, we do not yet have an efficient algorithm to compute them; an interesting further topic of research would be to study the existence of efficient Miller-type algorithms for generalized Weil and Tate pairings. It would also be compelling to study whether the results of Chapters 4 and 5 can be unified and extended into a classification of self-pairings on general, not necessarily cyclic, subgroups compatible with oriented isogenies.

In Chapter 6, we devised generalized class polynomials; a multivariate extension of class polynomials. Class polynomials have previously been studied as a generalization of Hilbert class polynomials. The sizes of their coefficients are sometimes smaller by an asymptotic factor, improving their computational applicability in, for example, the CM method. The best known class polynomials obtain an asymptotic size reduction factor of 72. We showed that generalized class polynomials obtain provable asymptotic

[1] At the time of writing, upcoming work has been announced claiming that the condition that the divisor be *square* may be removed.

size reductions that were previously unattainable for a positive proportion of imaginary quadratic discriminants. However, our best such examples still have a reduction factor of at most 72. An interesting further research goal would be to find the first example of a family of generalized class polynomials, say of prime class number, that attain provable asymptotic size reductions beyond 72, or perhaps even exceeding the theoretical univariate bound of 100.83. Another goal is to extend the state-of-the-art method for computing class polynomials, a CRT-based approach by Sutherland, to the case of generalized class polynomials.

In Chapter 7, we studied radical isogenies; a method to compute chains of isogenies of fixed degree based on formulae containing a radical expression. We developed a new way to compute radical isogeny formulae that combines the CM method and Galois theory of function fields of modular curves with CRT-based rational interpolation. This extended the range of degrees in which formulae are available from $N \leq 13$ to all prime $N \leq 41$. Moreover, we simplified formulae and improved their computational performance. We also formulated a conjecture that states, in case of CSIDH, which radical must be taken for the corresponding radical isogeny to be horizontal, and proved this conjecture for all $N \leq 14$. A further goal would be to prove this conjecture for all (even) $N \geq 4$. It would also be interesting to find a method for producing general radical isogeny formulae that is more direct than by means of rational interpolation, for example by obtaining a closed form expression, or a linear recurrence relation satisfied by the formulae.

Samenvatting

Cryptografie gaat over het beveiligen van informatie op een manier waarop alleen de beoogde personen toegang hebben tot die informatie. Stel, bijvoorbeeld, dat een hypothetische persoon, genaamd Alice, een privébericht zou willen sturen naar een andere hypothetische persoon, genaamd Bob. Als Alice en Bob een manier hadden om te communiceren zodat niemand hen zou kunnen afluisteren, dan zou dat makkelijk zijn; ze zouden hun berichten onversleuteld kunnen overbrengen. In de echte wereld is een perfect veilige communicatieverbinding echter vrijwel onmogelijk te garanderen, zeker wanneer communicatie over het internet gebeurt. Alice en Bob zouden dus op een bepaalde manier een soort codetaal moeten afspreken. Maar hoe kunnen ze dat doen, als we aannemen dat kwaadwillende personen al hun communicatie kunnen afluisteren? Eén manier, is met een zogeheten *Diffie–Hellman sleuteluitwisseling*; een methode voor twee partijen om een gezamenlijk geheim af te spreken over een publiek kanaal. Een gangbare manier, toegepast door veel applicaties die gebruik maken van *begin-tot-eind*-versleuteling, is gebaseerd op wiskundige objecten genaamd *elliptische krommen.* Een voorbeeld van een elliptische kromme is de verzameling van punten (x, y) in het vlak die voldoen aan de vergelijking $y^2 = x^3 + 3x^2 - x - 3$; zie Figuur 9.1.

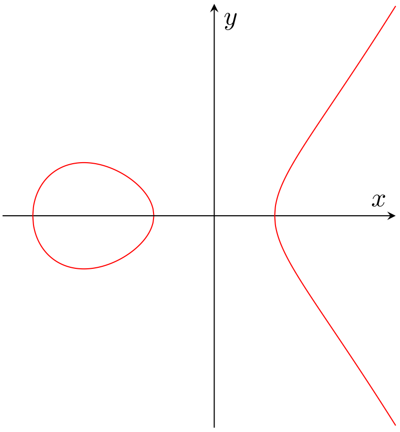


Figure 9.1: Een elliptische kromme gegeven door $y^2 = x^3 + 3x^2 - x - 3$.

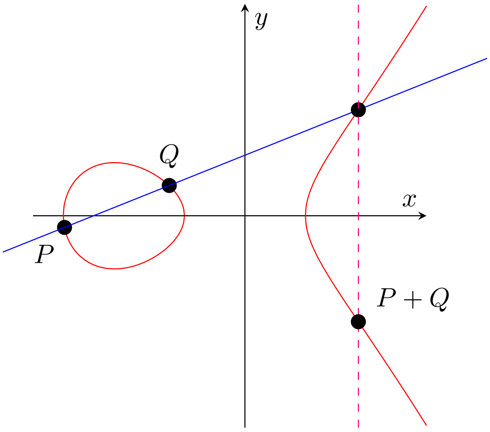


Figure 9.2: Het optellen van twee punten op een elliptische kromme.

Wat elliptische krommen zo speciaal maakt, is dat er een meetkundig recept bestaat om punten op de kromme op te tellen: Stel bijvoorbeeld, dat we twee punten op de kromme, zeg P en Q, zouden willen optellen. We tekenen dan eerst de lijn door P en Q.

Het blijkt dat deze de kromme in precies één ander punt zal snijden.[2] De verticale lijn door dit laatste punt snijdt de kromme weer precies in één ander punt, wat we $P + Q$ noemen; de som van P en Q. In het geval dat $P = Q$, dan definiëren we "de lijn door P en Q" als de raaklijn aan de kromme in P. Op deze manier hebben we een methode om $n \cdot P = \underbrace{P + P + \ldots + P}_{n \text{ keer } P}$ voor elk geheel getal $n > 0$ uit te rekenen. Grote zulke veelvouden kunnen overigens met minder dan $n - 1$ optellingen uitgerekend worden, door een methode genaamd *double-and-add* (Engels voor "verdubbelen en optellen"). Zo kunnen we bijvoorbeeld $20 \cdot P$ uitrekenen als $20 \cdot P = 2 \cdot (2 \cdot ((2 \cdot (2 \cdot P)) + P))$, wat slechts vijf optellingen kost (waarvan vier een verdubbeling zijn; i.e. een punt optellen bij zichzelf). Nu, als Alice en Bob een elliptische kromme willen gebruiken om een gezamenlijk geheim vast te stellen, dan kan dit als volgt.

(i) Alice en Bob spreken, in het openbaar, een punt P op een elliptische kromme af.

(ii) Alice en Bob genereren (grote) geheime gehele getallen a en b.

(iii) Alice rekent het punt $P_A = a \cdot P$ uit, en stuurt het resultaat naar Bob.

(iv) Bob rekent het punt $P_B = b \cdot P$ uit, en stuurt het resultaat naar Alice.

(v) Met behulp van haar geheim en het punt van Bob, berekent Alice $a \cdot P_B = (a \cdot b) \cdot P$.

(vi) Met behulp van zijn geheim en het punt van Alice, berekent Bob $b \cdot P_A = (a \cdot b) \cdot P$.

Aangezien Alice en Bob op hetzelfde punt op de elliptische kromme uitkomen, hebben ze successvol een gezamenlijk geheim vastgesteld; dat wil zeggen, de *sleuteluitwisseling* is voltooid. Deze gezamenlijke sleutel kan vervolgens gebruikt worden om beveiligde berichten naar elkaar te sturen.

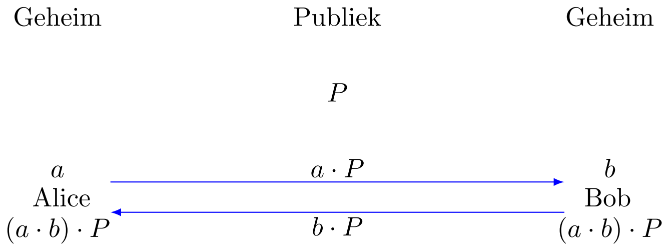


Figure 9.3: Een "Elliptic Curve Diffie–Hellman" sleuteluitwisseling.

De veiligheid van dit protocol steunt op de aanname dat het onmogelijk is om Alice' geheim a vast te stellen uit alleen de publieke informatie van P en $a \cdot P$. Dit heet het *discretelogaritmeprobleem.* Theoretisch gesproken zou het weliswaar mogelijk zijn om

[2] Tenzij de lijn door P en Q verticaal is; dan zeggen we dat $P + Q = O$, waar O het *punt op oneindig* heet. Als de lijn de kromme in één van de punten raakt, dan tellen we dat snijpunt dubbel. Op die manier is "de lijn door P en P" gelijk aan de raaklijn aan de kromme in P.

a uiteindelijk te vinden door $2 \cdot P = P + P$, $3 \cdot P = P + P + P$, $4 \cdot P = P + P + P + P$, enzovoorts, uit te rekenen totdat we $a \cdot P$ tegenkomen, maar dit is praktisch niet haalbaar zodra a heel groot is. Al deze punten nagaan duurt namelijk veel langer dan $a \cdot P$ te berekenen, gegeven a en P, met behulp van de *double-and-add*-methode. Tot op heden zijn er geen snelle algoritmes bekend om het discretelogaritmeprobleem in het algemeen op te lossen. Alhoewel, tenzij we *kwantumcomputers* in beschouwing nemen. Een kwantumcomputer is een speciaal soort computer die zijn rekenkracht baseert op de merkwaardige eigenschappen van subatomaire deeltjes. Op zulke computers bestaan er wél snelle algoritmes die het discretelogaritmeprobleem kunnen kraken. Zo ver we weten, is nog niemand in staat geweest een kwantumcomputer te bouwen die krachtig genoeg is om hedendaagse cryptografie te breken, maar het is onduidelijk of dit in de toekomst wel mogelijk zal zijn. Dit heeft een nieuw vakgebied in het leven geroepen genaamd *post-kwantum cryptografie*. Deze onderzoeksrichting gaat over het zoeken en analyseren van manieren om informatie te versleutelen die veilig zijn tegen aanvallen van kwantumcomputers. *Isogenie-gebaseerde cryptografie* is een deelgebied van deze onderzoeksdiscipline. Isogenieën zijn afbeeldingen tussen elliptische krommen; een soort vervorming die je van de ene elliptische kromme naar de andere brengt. Wanneer je deze vervormingen op een slimme manier kiest, kan je een sleuteluitwisselingsprocedure maken dat erg op het voorgaande protocol lijkt. Voor zo'n procedure beginnen Alice en Bob met het afspreken van een publiek bekende elliptische kromme, maar in plaats van een punt op die kromme te nemen, passen ze nu (geheime) vervormingen toe op de elliptische kromme zelf. Als ze dit doen op een zodanige manier dat de volgorde van hun vervormingen niet uitmaakt, eindigen ze samen met dezelfde elliptische kromme, wat vervolgens hun gedeelde geheim uitmaakt. Een abstracte weergave van dit protocol vindt je in Figuur 9.4

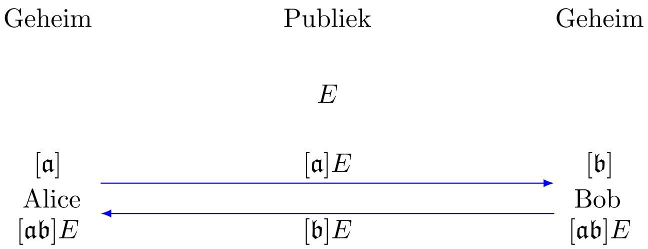


Figure 9.4: Een isogenie-gebaseerde sleuteluitwisseling.

De onderliggende aanname van isogenie-gebaseerde cryptografie is dat het lastig is om, gegeven twee elliptische krommen E_1 en E_2, een vervorming te vinden van die je van E_1 naar E_2 brengt. Dit heet ook wel het *isogenie-pad-probleem*. Men gaat ervan uit dat dit probleem zelfs voor kwantumcomputers lastig is.

In dit proefschrift beschouwen we verschillende computationele problemen die aan isogenieën tussen elliptische krommen kunnen worden toegekend.

Hoofstukken 1, 2, en 3 zijn inleidend en eindigen met een vereenvoudigd overzicht van de hoofdresultaten uit de latere hoofdstukken.

In Hoofdstuk 4 en 5 tonen we aan dat, in zekere instanties, afbeeldingen op elliptische krommen genaamd *pairings* gebruikt kunnen worden om bepaalde computationele moeilijkheidsaannames te weerleggen. Zo vinden we in speciale gevallen efficiënte oplossingen voor het isogenie-pad-probleem, evenals voor een zwakker probleem genaamd het *Diffie–Hellman Beslissingsprobleem*.

In Hoofdstuk 6 ontwikkelen we een meervariabele veralgemening van *Hilbert klassenpolynomen*; veeltermen die elliptische krommen met een bepaalde structuur (gegeven door hun *endomorfismering*) beschrijven. We gaan in het bijzonder in op de computationele voordelen van deze nieuwe veeltermen ten opzichte van eerder bekende klassenpolynomen.

In Hoofdstuk 7 bestuderen we een methode om ketens van isogenieën efficiënt uit te rekenen met behulp van vergelijkingen genaamd *radicale-isogenie-formules*. We ontwikkelen een nieuwe methode om zulke formules uit te rekenen, en verbeteren de efficiëntie van hun evaluatie. Dit zorgt voor een versnelling in het uitvoeren van bepaalde isogenie-gebaseerde protocollen.

Summary

Cryptography is about securing information in such a way that only the intended parties have access to that information. For example, let us say that a hypothetical person, named Alice, would like to send a message to another hypothetical person, named Bob. If Alice and Bob had a way to communicate in such a way that no one else could overhear their conversation, then this would be easy; they could just converse in plain text. However, in the real world, a perfectly secure channel of communication is virtually impossible to guarantee, especially when communication happens over the internet. Somehow, Alice and Bob have to agree on some sort of code language. But how can they do that, if we assume malicious entities can listen in on all of their conversations? One way, is through something called a *Diffie–Hellman key exchange*; a method for two parties to establish a shared secret over a public communication channel. A common way, used by many end-to-end encrypted messaging applications, is based on mathematical objects called *elliptic curves*. An example of an elliptic curve is the collection of points (x, y) in the plane satisfying the equation $y^2 = x^3+3x^2-x-3$; see Figure 9.5.

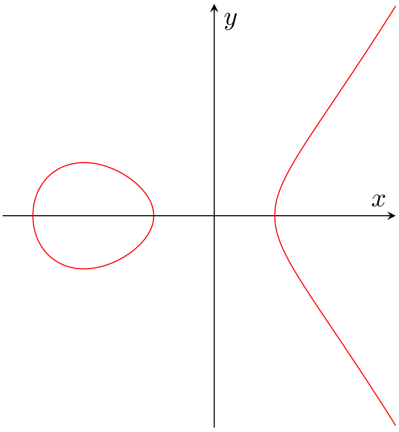


Figure 9.5: An elliptic curve given by the equation $y^2 = x^3 + 3x^2 - x - 3$.

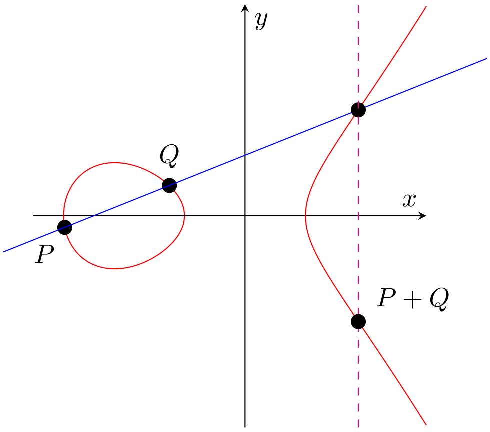


Figure 9.6: Adding two points on an elliptic curve.

What is special about elliptic curves, is that there is a geometric recipe to add points on the curve to each other, which is described as follows. When we would like to add two points P and Q on the curve, we draw the line connecting P and Q, which intersects the curve in exactly one other point.[3] The vertical line through this latter

[3] Unless the line is vertical; then we say $P + Q = O$, where O is called the *point at infinity*. If the

point intersects the curve in exactly one other point, which is $P+Q$; the sum of P and Q. In the case that $P = Q$, then we say that the line connecting P and Q is the tangent to the curve at P. This way, we have a recipe to compute $n \cdot P = \underbrace{P + P + \ldots + P}_{n \text{ times } P}$ for any integer $n > 0$. Such multiples can be computed in a faster way than just adding the point to itself $n-1$ times, through a procedure called *double-and-add*. For example, we can compute $20 \cdot P = 2 \cdot (2 \cdot ((2 \cdot (2 \cdot P)) + P))$ by just five additions (of which four are a doubling; i.e. adding a point to itself). Now, if Alice and Bob would like to establish a common secret, they could execute the following procedure.

(i) Alice and Bob agree publicly on a point P on an elliptic curve.

(ii) Alice and Bob generate (large) secret integers a and b.

(iii) Alice computes the point $P_A = a \cdot P$ and sends the result to Bob.

(iv) Bob computes the point $P_B = b \cdot P$ and sends the result to Alice.

(v) Using her secret and the point from Bob, Alice computes $a \cdot P_B = (a \cdot b) \cdot P$.

(vi) Using his secret and the point from Alice, Bob computes $b \cdot P_A = (a \cdot b) \cdot P$.

Since Alice and Bob both end up at the same point on the elliptic curve, they have successfully established a shared secret; that is, the *key echange* is complete. This common key can then be used to encrypt messages they would like to send to each other securely.

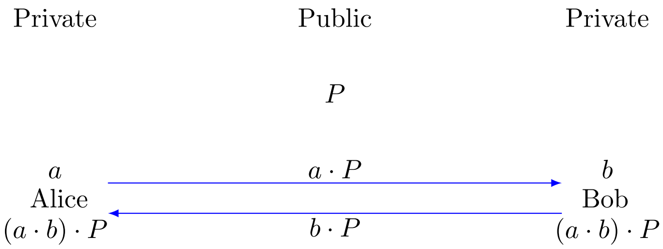


Figure 9.7: An Elliptic Curve Diffie–Hellman key exchange.

The security of this protocol relies on the assumption that it is impossible to recover Alice's secret a only using the publicly available information of P and $a \cdot P$. This is called the *discrete logarithm problem*. Theoretically, one would eventually be able to find a by computing $2 \cdot P = P + P$, $3 \cdot P = P + P + P$, $4 \cdot P = P + P + P + P$, and so on, until one eventually runs into $a \cdot P$. However, this is infeasible when a is really large; much slower than computing $a \cdot P$ given a and P by using the double-and-add method. Currently, no fast algorithms to solve the discrete logarithm problem

line is tangent to the curve in one of the points, then we count that intersection twice. In this way "the line through P and P" is the tangent to the curve at P.

in general are known. That is, unless we take into account *quantum computers.* A quantum computer is a special type of device that bases its computational power on the remarkable physical properties of subatomic particles. On such computers, there are known to exist fast algorithms to solve the discrete logarithm problem. To date, as far as we know, no one was able to build a quantum computer powerful enough to break any practically used cryptographic protocol. However, it is unclear whether such a device will be constructed in the near future. This has sparked a new area of research called *post-quantum cryptography*, which searches for ways to encrypt information that are secure against attacks by quantum computers. One such proposal is called *isogeny-based cryptography.* Isogenies are maps between elliptic curves; a type of transformation that takes you from one elliptic curve to the other. When chosen in a smart way, such maps can be used to establish a key exchange as before. This time, Alice and Bob publicly agree, not on a point on an elliptic curve, but on an elliptic curve itself. They apply successive transformations to the curve in such a way that they end up at the same elliptic curve, which then forms their shared secret. Abstractly, this is pictured in Figure 9.8.

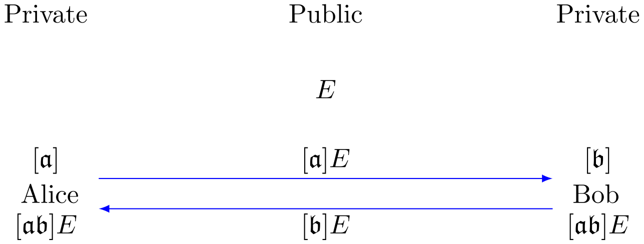


Figure 9.8: An isogeny-based key exchange protocol.

The assumption underlying the security of isogeny-based cryptography, is that it is difficult, given two elliptic curves E_1 and E_2, to find a transformation from E_1 to E_2. This is called the *isogeny path problem.* It is assumed that this problem is difficult even for quantum computers.

In this thesis, we consider several computational problems associated to isogenies between elliptic curves.

Chapters 1, 2, and 3 are introductory and end with a high-level overview of the main results presented in later chapters.

In Chapter 4 and 5, we show how, in certain instances, maps on elliptic curves called *pairings* can be used to disprove computational hardness assumptions related to isogeny-based cryptography. In special cases, we find efficient solutions to the isogeny path problem, as well as to a weaker problem known as the *Decisional Diffie–Hellman Problem.*

In Chapter 6, we develop a multivariate generalization of *Hilbert class polynomials*; polynomials that encode elliptic curves with a certain structure (given by their *endomorphism ring*). We in particular discuss the computational benefits of these novel polynomials compared to previously known class polynomials.

In Chapter 7, we study a method to compute chains of isogenies efficiently through equations called *radical isogeny formulae*. We develop a new method to obtain such formulae, and improve on the efficiency of their evaluation. This leads to a speed-up in the execution of certain isogeny-based cryptographic protocols.

Curriculum Vitae

Marc Houben was born on 16 March 1995 in Utrecht, The Netherlands. He grew up in Houten, where he went to high school College de Heemlanden from 2007, obtaining his diploma in 2013.

After that, he studied at Utrecht University, where in 2016 he obtained bachelor degrees *cum laude* in both Mathematics and Physics. His bachelor's thesis on "Congruences for coefficients of power series expansions of rational functions" was written under supervision of prof. dr. Frits Beukers. He continued studying Mathematics at Utrecht University, obtaining a master's degree *cum laude* in 2018. His master's thesis on "Dynamics on algebraic groups" was written under supervision of prof. dr. Gunther Cornelissen.

In October 2018, Marc started a joint PhD in Mathematics between KU Leuven and Leiden University under supervision of dr. Wouter Castryck and dr. Marco Streng. In November 2019, he began a PhD Fellowship fundamental research from Research Foundation – Flanders (FWO). Since October 2021, Marc is a visitor at the Computer Security and Industrial Cryptography (COSIC) group at KU Leuven.

List of Publications

This thesis is based on the following published papers.

[1] **Weak instances of class group action based cryptography via self-pairings.**
Wouter Castryck, Marc Houben, Simon-Philipp Merz, Marzio Mula, Sam van Buuren, and Frederik Vercauteren.
CRYPTO 2023.

[2] **Horizontal racewalking using radical isogenies.**
Wouter Castryck, Thomas Decru, Marc Houben, and Frederik Vercauteren.
ASIACRYPT 2022.

[3] **On the decisional Diffie–Hellman problem for class group actions on oriented elliptic curves**.
Wouter Castryck, Marc Houben, Frederik Vercauteren, and Benjamin Wesolowski.
ANTS XV. In Res. Number Theory 8.4, Paper No. 99, 18. 2022.

[4] **Generalized class polynomials.**
Marc Houben and Marco Streng.
ANTS XV. In Res. Number Theory 8.4, Paper No. 103, 26. 2022.

The author of the thesis has additionally published the following papers that are not included in the thesis.

[5] **Dynamically affine maps in positive characteristic.**
Jakub Byszewski, Gunther Cornelissen, and Marc Houben, with Appendix B by the authors and Lois van der Meijden.
Contemporary Mathematics 744. pp. 125–156, 2020.

[6] **Gauss congruences for rational functions in several variables.**
Frits Beukers, Armin Straub, and Marc Houben.
Acta Arithmetica 184.4, pp. 341-362, 2018.