

(C)SIDH – Isogeny school week 3

Tanja Lange, Eindhoven University of Technology

<https://www.hyperelliptic.org/tanja/teaching/isogeny-school21/>

18 July 2021

1 Introduction

Isogeny-based cryptography uses maps between elliptic curves to build public-key cryptography. The first such system dates back to 1997, but the first publicly accessible proposals by Couveignes [Cou06] and Rostovtsev–Stolbunov [RS06] are from 2006, after they were used as an attack tool [GHS02, JMV05] in 2002 and 2005 and an isogeny-based hash function [CGL09] was published in 2006. These systems use isogenies between ordinary elliptic curves over finite fields to create a key-exchange system; the key-exchange system is denoted CRS in the following. Stolbunov [Sto11] also shows how to use this construction for building identification schemes. Shor’s attack [Sho97], which breaks elliptic-curve cryptography based on the discrete-logarithm problem, does not affect these constructions, but in 2010 Childs, Jao, and Soukharev [CJS14] showed that CRS can be broken with a sub-exponential quantum attack due to Kuperberg [Kup05] (see week 8). This means that parameters of the CRS scheme need to be scaled up asymptotically, making this already slow system even slower, but it does not mean that the system is fundamentally broken.

In 2011, Jao and De Feo [JF11] designed a different isogeny-based system that uses isogenies between supersingular curves over extension fields and does not have the same weakness as the above-described CRS key-exchange and which according to current knowledge offers exponential security, even against quantum attacks. A minor downside compared to CRS is the more complicated data flow and that the security assumption has changed from the pure isogeny-finding problem to one where additional information is available.

In the years since, research on isogeny-based systems focused mostly on this supersingular-isogeny Diffie-Hellman (SIDH) protocol. Many speedups were found and the security is better understood now; one important attack [GPST16] (see torsion-point attacks in week 6) showed that reusing keys requires extra care such as using the Fujisaki-Okamoto transform [FO99], but for ephemeral use no issues are known. The shared benefit of CRS and SIDH systems is that they require very little bandwidth compared to other systems that are expected to resist attacks using quantum computers.

This text introduces CSIDH and SIDH.

2 Mathematics background: elliptic curves and isogenies

For basics on elliptic curves and isogenies see the lectures from week 1 [Pan21]. This section only settles notation so that the text can be read as standalone.

2.1 Weierstrass curves and twists

Let p be a prime larger than 3 and $n > 0$ an integer. Let \mathbb{F}_{p^n} denote the finite field with p^n elements. An elliptic curve E over \mathbb{F}_{p^n} can be written in short Weierstrass form

$$E : y^2 = x^3 + ax + b, \quad a, b \in \mathbb{F}_{p^n} \text{ and } 4a^3 + 27b^2 \neq 0.$$

Points on the elliptic curves over \mathbb{F}_{p^n} are all pairs $(x, y) \in \mathbb{F}_{p^n}^2$ which satisfy the curve equation along with an extra point ∞ . These points form a group with ∞ as neutral element. This same group operation works for points over any extension field of \mathbb{F}_{p^n} . We use E/\mathbb{F}_{p^n} to denote that E is defined over \mathbb{F}_{p^n} ; we use $E(\mathbb{F}_{p^n})$ to denote the group of points on E over \mathbb{F}_{p^n} , and we use E when speaking about properties that hold independently of the extension field.

Two elliptic curves E_1/\mathbb{F}_{p^n} and E_2/\mathbb{F}_{p^n} are *isomorphic over \mathbb{F}_{p^n}* if there exists a polynomial map over \mathbb{F}_{p^n} that maps points (x, y) on E_1 to points on E_2 in a one-to-one way which is compatible with the group operation. If E_1 and E_2 are given in short Weierstrass form with $ab \neq 0$ any isomorphism has the form $(x, y) \mapsto (u^2x, u^3y)$ for some $u \neq 0$.

The systems we present later will consider elliptic curves up to isomorphism, i.e., work with *isomorphism classes*. They thus require a unique representative for each class. The typical choice of invariant for isomorphism classes is the j -invariant which for curves in Weierstrass form is $j = 1728 \cdot 4a^3 / (4a^3 + 27b^2)$. For the example $(x, y) \mapsto (u^2x, u^3y)$ map given above the curve coefficients satisfy $a_2 = u^4a_1$ and $b_2 = u^6b_1$ which gives the same j for E_1 and E_2 . Note, however, that the j -invariant uniquely describes isomorphism classes over an algebraic closure of \mathbb{F}_{p^n} , two curves having the same j -invariant need not be isomorphic over \mathbb{F}_{p^n} .

A twist of a curve E/\mathbb{F}_{p^n} is a curve E'/\mathbb{F}_{p^n} that is isomorphic to the first over some extension field. The degree of the twist is the smallest degree of extension of \mathbb{F}_{p^n} over which the isomorphism is defined. A quadratic twist of a curve E/\mathbb{F}_{p^n} is a curve that is isomorphic to the first over $\mathbb{F}_{p^{2n}}$ and not over \mathbb{F}_{p^n} . Let ν be a nonsquare over \mathbb{F}_{p^n} , then $y^2 = x^3 + a\nu^2 + b\nu^3$ is a quadratic twist of $y^2 = x^3 + ax + b$.

The number of points on an elliptic curve over \mathbb{F}_{p^n} is roughly p^n : for $\#E(\mathbb{F}_{p^n}) = p^n + 1 - t$ the integer t lies in the interval $[-2p^{n/2}, 2p^{n/2}]$. An elliptic curve over \mathbb{F}_{p^n} is *supersingular* if $t \equiv 0 \pmod{p}$, else the curve is *ordinary*.

The *order* of a point $P \in E(\mathbb{F}_{p^n})$ is the smallest integer $k > 0$ such that $[k]P = \infty$, where $[k]P$ means the addition of k copies of P . Since there are only finitely many points in $E(\mathbb{F}_{p^n})$ the order of every point is finite; furthermore, the order of any point $P \in E(\mathbb{F}_{p^n})$ divides the group order $\#E(\mathbb{F}_{p^n})$. For every prime ℓ not dividing p there are either 1, ℓ , or

ℓ^2 points P with $[\ell]P = \infty$. The first case corresponds to only $P = \infty$ satisfying $[k]P = \infty$. The second case corresponds to $\ell - 1$ points of order ℓ and ∞ satisfying the equations; these points form a cyclic group that behaves like \mathbb{Z}/ℓ . The third case corresponds to $\ell^2 - 1$ points of order ℓ and ∞ satisfying the equations; these points form a product of two cyclic groups that behaves like $\mathbb{Z}/\ell \times \mathbb{Z}/\ell$. In particular, in this case all points of order ℓ are given as a linear combination $[a]P + [b]Q$, $a, b \in [0, \ell - 1]$, where P and Q are points of order ℓ and Q is not a multiple of P .

2.2 Montgomery curves

There are other representations of elliptic curves, most notably Montgomery form [Mon87]

$$BY^2 = X^3 + AX^2 + X, \quad A, B \in \mathbb{F}_{p^n} \text{ and } B(A^2 - 4) \neq 0$$

and Edwards form [BL07] which have advantages for implementations and, sometimes, for exposition. We will use Montgomery curves for most of the exposition. They are very similar to Weierstrass curves and have neutral element ∞ , typically thought high up in the direction of the y -axis, and $-(X, Y) = (X, -Y)$. They have the following addition formulas for the main cases of adding $P = (X_1, Y_1)$ and $Q = (X_2, Y_2)$ giving (X_3, Y_3) with $X_3 = B\lambda^2 - X_1 - X_2 - A$, $Y_3 = \lambda(X_1 - X_3) - Y_1$ and

$$\lambda = \begin{cases} (Y_1 - Y_2)/(X_1 - X_2) & \text{for } \begin{cases} P \neq \pm Q \\ P = Q \neq -Q \end{cases} \\ (3X^2 + AX + 1)/(2B) & \end{cases}$$

Montgomery curves are particularly interesting because they offer efficient differential addition when using only X coordinates. Differential addition means that you can add P and Q if you know their difference $Q - P$. In the case of X -only arithmetic this means computing $X(P + Q)$ given $X(P)$, $X(Q)$, and $X(Q - P)$.

The Montgomery ladder uses differential addition to compute scalar multiples. The ladder starts with $P_0 = \infty$ and $P_1 = P$ and scans the scalar from the most significant bit to the least significant bit, each step updates $P_0 = [2]P_0$, $P_1 = P_0 + P_1$ if the bit is 0 and $P_0 = P_0 + P_1$, $P_1 = [2]P_1$ if the bit is 1. Note that the difference $P_1 - P_0 = P$ is invariant under the update steps and that for each addition the difference is thus P or ∞ (for the doubling).

When using X -only arithmetic we use 0 to represent ∞ as well as $(0, 0)$. The formulas for doubling compute $X_3 = (X_1^2 - 1)^2 / (4X_1(X_1^2 + AX_1 + 1))$. For the differential addition let X_- denote the X -coordinate of the difference. The X -coordinate of the sum is given by $X_3 = ((X_1 + 1)(X_2 - 1) + (X_1 - 1)(X_2 + 1))^2 / (X_-((X_1 + 1)(X_2 - 1) - (X_1 - 1)(X_2 + 1))^2)$.

See the EFD for more curve shapes and efficient formulas, incl. formulas for handling these computations in projective coordinates. Week 10 will also cover more on efficient arithmetic.

For ν a nonsquare over \mathbb{F}_{p^n} and $A \neq 0$, a quadratic twist of $E : BY^2 = X^3 + AX^2 + X$ is given by $E' : \nu BY^2 = X^3 + AX^2 + X$. Each $X \in \Phi_p$ satisfies either that $X^3 + AX^2 + X$ is a

square in \mathbb{F}_p , thus there are two points $(X, \pm\sqrt{X^3 + AX^2 + X})$ in $E(\mathbb{F}_p)$; or $X^3 + AX^2 + X$ is not a square in \mathbb{F}_p , thus there are two points $(X, \pm\sqrt{\nu(x^3 + AX^2 + X)})$ in $E'(\mathbb{F}_p)$; or $X^3 + AX^2 + X = 0$, thus $(X, 0)$ is a point in $E(\mathbb{F}_p)$ and in $E'(\mathbb{F}_p)$. Combining these three cases gives two points for each $X \in \mathbb{F}_p$, thus in addition to one point at infinity per curve we have $\#E(\mathbb{F}_p) + \#E'(\mathbb{F}_p) = 2p + 2$. Therefore $\#E(F\Phi_p) = p + 1 - t$ implies $\#E'(\mathbb{F}_p) = p + 1 + t$.

Montgomery curves always have a point $(0, 0)$ of order 2 and at least one of the following: $X^2 + AX + 1 = (X - X_1)(X - X_2)$, giving $(X_1, 0), (X_2, 0)$ of order 2 or there is a point of order 4. Hence, the group order is always divisible by 4.

3 Isogenies

For mathematically rigorous definitions of the following see the lecture notes of week 1 [Pan21].

3.1 Definitions

An *isogeny* between two elliptic curves E_1/\mathbb{F}_{p^n} and E_2/\mathbb{F}_{p^n} is a non-constant rational function that maps points from E_1 to points on E_2 and is compatible with the group law. In particular, it maps the neutral element ∞_1 on E_1 to the neutral element ∞_2 on E_2 . Unlike isomorphisms, isogenies need not be 1-to-1 but can (and typically do) have several points map to ∞_2 . For isogenies of interest in cryptography¹, the *degree* ℓ of an isogeny is the number of points on E_1 , taken over any extension field of \mathbb{F}_{p^n} , mapping to ∞_2 . Note, this does not mean that E_2 has fewer points than E_1 over \mathbb{F}_{p^n} , just that some points on E_2 are not in the image of points on E_1 over \mathbb{F}_{p^n} . In fact, by Tate's theorem, E_1/\mathbb{F}_{p^n} and E_2/\mathbb{F}_{p^n} are isogenous if and only if they have the same number of points over \mathbb{F}_{p^n} , i.e., $\#E_1(\mathbb{F}_{p^n}) = \#E_2(\mathbb{F}_{p^n})$. The set of curves that are isogenous to E is called the *isogeny class of E* . Note that if E is supersingular then all curves in its isogeny class are supersingular; similarly, ordinary curves are isogenous to ordinary curves.

For each isogeny $\varphi : E_1 \rightarrow E_2$ there exists a *dual isogeny* $\hat{\varphi} : E_2 \rightarrow E_1$ which has the same degree ℓ and for which it holds that the composition $\hat{\varphi} \circ \varphi = [\ell]_{E_1}$ is the multiplication-by- ℓ map on E_1 and likewise $\varphi \circ \hat{\varphi} = [\ell]_{E_2}$.

3.2 Vélu's formulas

The requirement that an isogeny be compatible with the group operation means that points of some order m coprime to ℓ are mapped to points of order m . An isogeny φ of degree ℓ has as kernel (i.e., points mapped to ∞_2) a cyclic subgroup of order ℓ , i.e., containing ℓ points, and each kernel uniquely defines an isogeny. These points may be defined over \mathbb{F}_{p^n} or over some extension field. The isogeny is defined over \mathbb{F}_{p^n} if the kernel is fixed under the p^n -Frobenius map (not pointwise but as a set). The typical way of computing an isogeny is to start with such a subgroup and then to use Vélu's formulas [V71], which give an

¹technically, for separable isogenies

explicit equation of the image curve and φ in terms of the coordinates of the points in the kernel. The computational complexity of these formulas grows linearly in ℓ and requires computations in the extension field over which the points in the kernel are defined.

We will need Vélu's formulas for Montgomery curves. Let P have prime order ℓ on E_A . For $1 \leq i < \ell$ let X_i be the X -coordinate of $[i]P$.

Let

$$\tau = \prod_{i=1}^{\ell-1} X_i, \quad \sigma = \sum_{i=1}^{\ell-1} \left(X_i - \frac{1}{X_i} \right), \quad f(x) = x \prod_{i=1}^{\ell-1} \frac{xX_i - 1}{x - X_i}.$$

Then the ℓ -isogeny with kernel $\langle P \rangle$ is given by

$$\varphi_\ell : E_A \rightarrow E_{A'}, (X, Y) \mapsto (f(X), c_0 Y f'(X))$$

where $A' = \tau(A - 3\sigma)$, and $c_0^2 = \tau$.

The main operation is to compute the X_i , just some elliptic-curve additions. Note that $[\ell - i]P = -[i]P$ and both have the same x -coordinate, so only $(\ell - 2)/2$ multiples of P need to be computed.

3.3 Isogeny graphs

The ℓ -isogeny graph over \mathbb{F}_{p^n} is an undirected graph that has as nodes the isomorphism classes of elliptic curves over \mathbb{F}_{p^n} and two such classes are connected if there exists an ℓ isogeny from one curve in the class to one in the other class. (By combining the isogeny with an isomorphism each curve in the class can be reached; note that different schemes use different extensions of \mathbb{F}_{p^n} for defining the isomorphisms). The graph is undirected because for each isogeny $\varphi : E_1 \rightarrow E_2$, the dual isogeny $\hat{\varphi}$ provides a map back.² In an ℓ -isogeny graph each node has zero, one, two, or $\ell + 1$ edges. This number depends on the structure of the set of points of order ℓ on the elliptic curve, which is a subgroup of $\mathbb{Z}/\ell \times \mathbb{Z}/\ell$. Each order- ℓ subgroup defines an isogeny over some extension field and zero, one, two, or $\ell + 1$ curves that are ℓ -isogenous to this curve may be defined over the same field.

The set of all curves over a field \mathbb{F}_{p^n} splits into multiple disjoint components consisting of curves ℓ -isogenous to one another. In the following, we only consider curves having the same number of points, i.e. curves that are isogenous under an isogeny of some degree. Note that this does not imply that all of these curves are connected under an ℓ isogeny for some fixed ℓ .

3.4 Isogeny graph for ordinary curves

²For isogenies landing back on the same curve one might need to introduce directions in order not to over count.

Components of the ℓ -isogeny graph consisting of ordinary curves form (parts of) *volcanoes* (see Figure 1 for an illustration). A full volcano consists of a circular rim (the top of the volcano) and each of the nodes of the rim additionally has $\ell - 1$ edges pointing downwards, each node on lower levels has ℓ edges pointing downwards. CRS uses only the top of the volcano and thus pick isogeny classes where that part is large. Each node in the top has exactly two neighbors and repeated application of ℓ isogenies makes multiple steps in the same direction. For different isogeny degrees ℓ_i the rim might split into multiple rims or multiple rims get combined into a larger one. The maximal size of the rim is the class number of the endomorphism ring; for details on what this means see [Sut12] and the lectures on class groups in week 1 [Bia21]. An important property for the CRS system is that the action of isogenies on ordinary elliptic curves is commutative, i.e., the order of applying ℓ_1 and ℓ_2 isogenies does not change the isomorphism class of the image curve.

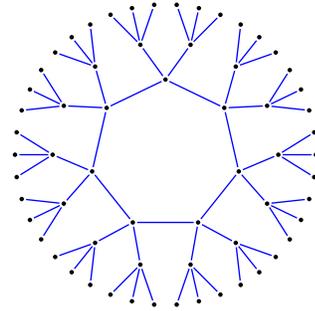


Figure 1: Isogeny graph for 3-isogenies forming a volcano with rim and two more levels. Image credit: Lorenz Panny.

3.5 Isogeny graph for supersingular curves

All isomorphism classes of supersingular elliptic curves over extensions of \mathbb{F}_p have their j -invariant defined over \mathbb{F}_{p^2} , so considering isomorphism classes of supersingular elliptic curves over \mathbb{F}_{p^2} covers all classes of supersingular curves, where the isomorphisms are taken over extension fields of \mathbb{F}_{p^2} .

There are of roughly $p/12$ isomorphism classes of elliptic curves over \mathbb{F}_{p^2} . The graph of ℓ isogenies is (almost) $\ell + 1$ regular and Ramanujan. This means that the graph is very well connected and any node in the graph can be reached in few steps from any other node (rapid mixing). It also means that there is no sense of direction – in computing a sequence of ℓ isogenies one can avoid going back but each step offers a choice of ℓ other edges forward.

Figure 2 shows an example for 2 and 3 isogenies for isomorphism classes of curves over \mathbb{F}_{419^2} , where each node corresponds to a j -invariant. The isomorphism classes are arranged around the circle; orange edges are 2 isogenies and blue ones 3 isogenies. This figure is typical for the full isogeny graph of supersingular curves. Note that endomorphism rings of supersingular curves are orders in a quaternion alge-

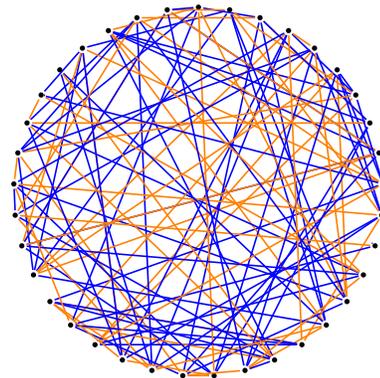


Figure 2: Nodes: Supersingular elliptic curves \mathbb{F}_{419^2} , edges are 2 and 3 isogenies. Image credit: Lorenz Panny.

are arranged around the circle so

that 3-isogenous curves are adjacent, showing the rim of the volcano of the 3-isogeny graph. The other isogenies are drawn to match this arrangement of isomorphism classes; if stratified, each of them would resemble a rim as well.

Note that each E_A on the left has E_{-A} on the right. This is not a coincidence: $\nu = -1$ is not a square for $p \equiv 3 \pmod{4}$ and thus the quadratic twist of E_A is given by $-Y^2 = X^3 + AX^2 + X$ which is isomorphic to $E_{-A} : Y^2 = X^3 - AX^2 + X$. Also note that Figure 2 would have identified these two classes as they are isomorphic over a quadratic extension of \mathbb{F}_p .

The graph also shows how to compute a 9-isogeny as a sequence of 3-isogenies iteratively. Rather than finding a point of order 9 over the starting curve, find one of order 3, compute the domain of that 3-isogeny, find a point of order 3 there, and compute the domain of that 3-isogeny. This means that there is no need to deal with extension fields to compute this sequence. The only requirement is that the second point is chosen so that the second isogeny is not the dual of the first. This gives an easy sense of direction, much unlike in the full graph in Figure 2.

This example is stereotypical for the CSIDH construction handled in Section 5: 3, 5, and 7 divide $p+1$ with multiplicity one and this holds for both the curve and its quadratic twist. So for each prime ℓ dividing $p+1$ a basis of the ℓ -torsion group can be given consisting of a point $P = (X_P, Y_P)$ with $X_P, Y_P \in \mathbb{F}_p$ and a point $Q = (X_Q, iY_Q)$ with $X_Q, Y_Q \in \mathbb{F}_p$ and $i^2 = -1$ in \mathbb{F}_p . Isogenies from points defined over \mathbb{F}_p continue in the forward direction (depicted counterclockwise in the graph) while those from points of the shape of Q define isogenies in the backwards direction.

4 The CRS system

The CRS system [Cou06, RS06] resembles closely the regular Diffie-Hellman key exchange. It uses the overlay of multiple isogeny graphs for the same set of ordinary curves isogenous to one another (when allowing the isomorphism to be over a larger field). The system parameters fix a finite field \mathbb{F}_p , a starting curve E of known order N , and a set of primes $\ell_i > 2$, $1 \leq i \leq r$, so that for each of these primes an ℓ_i isogeny can be computed with not too much effort and in a unique manner. For a properly-chosen curve the number of curves isogenous to it are on the order of \sqrt{p} (more precisely, it is the class number of the endomorphism ring of that curve, see week 1 [Bia21]).

For each of the primes ℓ_i Alice picks an exponent a_i and computes the curve E_A which is $\prod \ell_i^{a_i}$ isogenous to E . This is typically computed as a sequence of a_1 isogenies of degree ℓ_1 , a_2 isogenies of degree ℓ_2 , etc., see the description in Section 3.6 for how to compute this piece wise to avoid large extension fields. Computing the next ℓ_1 isogeny on $\varphi_{\ell_1}(E)$ works similar to the computation on E and continues on the rim of the ℓ_1 isogeny volcano. Alice's public key is E_A and her secret key is the exponent vector (a_1, a_2, \dots, a_r) .

Similarly, Bob picks (b_1, b_2, \dots, b_r) and computes and publishes E_B .

To check that Bob's key is valid Alice verifies that the number of points on E_B is N .

Starting from Bob’s curve E_B , Alice computes the curve E_{BA} which is $\prod \ell_i^{a_i}$ isogenous to E_B ; similarly Bob computes the curve E_{AB} which is $\prod \ell_i^{b_i}$ isogenous to E_A . The resulting curves are isomorphic because it does not matter whether Alice’s or Bob’s isogenies are applied first, because the endomorphism ring of ordinary curves is commutative. Thus $j(E_{BA}) = j(E_{AB})$ and Alice and Bob use a key derived from $j(E_{BA})$ as their shared secret.

The cost of computing the isogenies depends on the degrees ℓ_i and, when using Vélu’s formulas, the extension field of \mathbb{F}_p over which points of order ℓ_i are defined. If all a_i are chosen from an interval of size m , at most m^r different curves can be reached. In order to be able to reach all $\approx \sqrt{p}$ isogenous curves efficiently it is important to overlay multiple isogeny graphs, i.e. choose r sufficiently large, else a lot more steps are needed. This means that the system cannot stick to primes ℓ_i for which points of order ℓ_i are defined over \mathbb{F}_p but needs to move to more primes and extension fields. The timing in [Sto10] makes the system too slow to be practical, and the security of the chosen parameters was later revised downwards.

In 2018 De Feo, Kieffer, and Smith published [DKS18] a new way to construct curves for CRS but even with the many ideas that went into the construction and a large computation the resulting system was slow, yet much faster than earlier ones.

5 The CSIDH system

In 2018, Castryck, Lange, Martindale, Panny, and Renes tackled the problem of how to build an efficient CRS-like system, i.e., one in which all isogenies could be defined over the base field. The main observation in their CSIDH [CLM⁺18] scheme is that the \mathbb{F}_p endomorphism ring of supersingular curves over \mathbb{F}_p is commutative, as in the case of ordinary elliptic curves.

CSIDH works with isogenies between \mathbb{F}_p -isomorphism classes of supersingular curves defined over \mathbb{F}_p . The number of points on a supersingular curve over \mathbb{F}_p is $p + 1$ making it very easy to control which isogenies are defined over \mathbb{F}_p by choosing p appropriately. CSIDH puts $p = 4\ell_1 \cdot \ell_2 \cdots \ell_r - 1$ for some r and distinct odd primes ℓ_i . This means that the resulting curve has points of order ℓ_i over \mathbb{F}_p and Vélu’s formulas can be used efficiently. Note also that $p \equiv 3 \pmod{4}$, so all considerations from Section 3.6 apply and the isogeny graph has the beautiful regular structure depicted in Figure 3. CSIDH uses isomorphism classes in the largest connected component of the graph, denote this set by Z .

Of course, CSIDH uses much larger p and many more ℓ_i than in that example. The smallest version, CSIDH-512, proposed in [CLM⁺18] has $\log_2 p = 511$ and the ℓ_i in the definition of p being the first 73 odd primes and $\ell_{74} = 587$ (the smallest prime after ℓ_{73} to make $4 \prod_{i=1}^{74} \ell_i - 1$ prime).

CSIDH chooses another system parameter m . The original CSIDH proposal chooses r and m so that $(2m + 1)^r \approx \sqrt{p}$. By now many other choices have been proposed taking into account the relative costs of each isogeny as well as balancing pre- and post-quantum security. For CSIDH-512 the choice was $m = 5$.

Alice’s secret key is the exponent vector (a_1, a_2, \dots, a_r) with $a_i \in [-m, m]$ stating how

many times the ℓ_i -isogeny is computed. A positive exponent a_i indicates taking points P over \mathbb{F}_p of order ℓ_i while a negative one means taking points of the form $Q = (X_Q, iY_Q)$, see Section 3.6.

To compute her public key, Alice computes the elliptic curve that is reached after a_i isogenies of degree ℓ_i for $1 \leq i \leq r$. As the A coefficient in Montgomery form uniquely determines each \mathbb{F}_p -isomorphism class, CSIDH uses this coefficient to represent curves and thus Alice’s public key is the A coefficient of the resulting curve.

Bob does the same with his exponent vector (b_1, b_2, \dots, b_r) and publishes his resulting curve via the Montgomery coefficient A' .

The shared key of Alice and Bob is the curve reached with $(a_1 + b_1, a_2 + b_2, \dots, a_r + b_r)$ which each of them can compute by applying their sequence of isogenies to the other party’s public key curve.

5.1 Description using ideal class groups

A cleaner description of the mathematics behind CSIDH can be given using ideal class groups, see [Bia21] for details.

The elliptic curves in Z , meaning the main component shown in Figure 3, all have \mathbb{F}_p -endomorphism ring $\mathcal{O} = \mathbb{Z}[\sqrt{-p}]$. The ideal class group $\text{cl}(\mathcal{O})$ of the endomorphism ring acts on Z . As \mathcal{O} is commutative we get a commutative group action.

Let π be the Frobenius endomorphism. The ideal in \mathcal{O} above ℓ_i is $\mathfrak{l}_i = (\ell_i, \pi - 1)$.

We described moving in the positive direction with an ℓ_i -isogeny as picking a point of order ℓ_i defined over \mathbb{F}_p as the generator of the kernel. Being defined over \mathbb{F}_p means invariant under π , i.e., being in $\ker(\pi - 1)$. This means that the subgroup corresponding to \mathfrak{l}_i is $E[\mathfrak{l}_i] = E(\Phi_p)[\ell_i]$.

The conjugate of \mathfrak{l}_i is $\bar{\mathfrak{l}}_i = (\ell_i, \pi + 1) = \mathfrak{l}_i^{-1}$. Thus the subgroup corresponding to $\bar{\mathfrak{l}}_i$ is $E[\bar{\mathfrak{l}}_i] = \{Q \in E[\ell_i] \mid \pi(Q) = -Q\}$. Note that for a point $Q = (X_Q, iY_Q)$ with $X_Q, Y_Q \in \mathbb{F}_p$ the Frobenius is $(X_Q^p, i^p Y_Q^p) = (X_Q, -iY_Q) = -Q$, i.e., Frobenius operates by negation, explaining our choice of kernel points for the negative direction.

In this view, Alice’s secret key is the ideal class $\prod_{i=1}^r \mathfrak{l}_i^{a_i}$ and the key space is $\{\mathfrak{l}_1^{e_1} \cdots \mathfrak{l}_n^{e_n} \mid (e_1, \dots, e_n) \text{ is ‘short’} \subseteq \text{cl}(\mathcal{O})\}$. If the structure of \mathcal{O} is known it is possible to sample from the entire ideal-class group and use lattice algorithms to find an efficient short representation. With a quantum computer this structure can be computed in polynomial time, however, without one only subexponential algorithms are known and the computation for the CSIDH-512 parameters in [BKV19] was a major undertaking and no larger examples have been computed to date.

5.2 How to compute the CSIDH group action

Before doing any computation on a public key a user needs to validate it. For CSIDH with keys given by the Montgomery coefficient A this amounts to testing that $E : Y^2 = X^3 + AX^2 + X$ is supersingular, i.e., that $\#E(\mathbb{F}_p) = p + 1$. The factorization of $p + 1$ is known and by sampling random points on E one can check quickly that all ℓ_i appear

in the group order. In fact, one can stop sooner and use the Hasse interval, for details see [CLM⁺18].

To compute one ℓ_i -isogeny in the positive direction we want to find a point $P \in E[\ell_i](\mathbb{F}_p)$. The typical way is to sample a random X coordinate and check whether $X^3 + AX^2 + X$ is a square in \mathbb{F}_p , if so complete to a point $P' = (X, \pm\sqrt{X^3 + AX^2 + X})$ in $E(\mathbb{F}_p)$ and compute $P = [(p+1)/\ell_i]P'$. If $P \neq \infty$ it has order ℓ_i , else start with a fresh choice.

Note that the computations in Vélu's formulas (Section ??) use only the X coordinate of points, so it is not actually necessary to compute the Y coordinate of P' as Montgomery curves have very efficient computations of scalar multiples using only the X coordinate. For most choices of fields it is most efficient to do these computations in projective coordinates avoiding most inversions.

To compute one ℓ_i -isogeny in the negative direction we proceed similarly except for finding an X with $X^3 + AX^2 + X$ not being a square. A benefit of using X -only arithmetic is that no extension fields appear.

The CSIDH authors noticed that a lot of time is spent on finding points and doing scalar multiplications $[(p+1)/\ell_i]P'$. It is unavoidable to sample a new point for each of the a_i isogenies of degree ℓ_i but one can use a point a point T of order 15 on the first curve to compute a 3-isogeny followed by a 5-isogeny by first taking $[5]T$ as kernel to compute $\varphi_3 : E \rightarrow E'$, then computing $T' = \varphi_5(T)$ on E' and using T' as kernel compute $\varphi_5 : E' \rightarrow E''$. Note that the application of φ_3 means that T' has order 5. CSIDH thus merges computations for all ℓ_i that have the same sign in their exponent.

The following Algorithm 1 is taken from [CLM⁺18].

Algorithm 1: Evaluating the class-group action.

Input: $A \in \mathbb{F}_p$ and a list of integers (e_1, e_2, \dots, e_n) .

Output: A' such that $[e_1^{e_1} \cdots e_n^{e_n}]E_A = E_{A'}$ (where $E_{A'} : Y^2 = X^3 + A'X^2 + X$).

While some $e_i \neq 0$ **do**

Sample a random $x \in \mathbb{F}_p$.

Set $s \leftarrow +1$ if $x^3 + Ax^2 + x$ is a square in \mathbb{F}_p , else $s \leftarrow -1$.

Let $S = \{i \mid e_i \neq 0, \text{sign}(e_i) = s\}$. **If** $S = \emptyset$ **then** start over with a new x .

Let $k \leftarrow \prod_{i \in S} \ell_i$ and compute $T \leftarrow [(p+1)/k]P$.

For each $i \in S$ **do**

Compute $R \leftarrow [k/\ell_i]T$. **If** $R = \infty$ **then** skip this i .

Compute an isogeny $\varphi : E_A \rightarrow E'_A : y^2 = x^3 + A'x^2 + x$ with $\ker \varphi = R$.

Set $A \leftarrow A'$, $T \leftarrow \varphi(T)$, $k \leftarrow k/\ell_i$, and finally $e_i \leftarrow e_i - s$.

Return A .

As pointed out in [CLM⁺18] this computation is highly irregular and timing can give information on how often each ℓ_i -isogeny is used, thus leaking the exponent vector (e_1, e_2, \dots, e_r) . Note that since 2018 many improvements to this basic approach have appeared, including versions that can be used for constant-time implementations, see weeks 10 and 11 for details.

6 Security analysis of CRS and CSIDH

For CRS and CSIDH, an attacker is confronted with the problem of finding E_{AB} given E , E_A , and E_B . Analogous to the situation of discrete logarithms and Diffie-Hellman, there is no known attack faster than computing Alice’s or Bob’s secret isogeny to get E_{AB} .

Obviously, the number of keys has to be large enough to protect against brute force searches or their more intelligent meet-in-the-middle variants. These search for the key in time square-root of the search space, so in roughly $\sqrt[p]{p}$ if the key space is large enough so that all isogenous curves can be reached.

A quantum attacker can use the attack by Childs, Jao, and Soukharev [CJS14] which requires a subexponential number of calls to an oracle which compute isogenies in quantum superposition. A recent analysis [BLMP19] for the case of CSIDH shows that the cost of each such oracle call contributes significantly to the cost of the attack so that for low security levels the main concern is to defend against the above non-quantum attacks. Recent papers [BS20, Pei20] have focused on the number of oracle calls and improving Kuperberg’s algorithm, see week 8.

7 The SIDH system

The SIDH system [JF11, DFJP14] uses isomorphism classes of supersingular curves over \mathbb{F}_{p^2} . The number of points $\#E(\mathbb{F}_{p^2}) = p^2 - t + 1$ has $t \in [-2p, 2p]$ and $t \equiv 0 \pmod{p}$, thus $t \in \{-2p, -p, 0, p, 2p\}$. Of these, the isomorphism classes considered in crypto have $t = -2p$, thus $\#E(\mathbb{F}_{p^2}) = (p+1)^2$. Given that the isomorphism classes are considered over extension fields, a curve and its quadratic twist land in the same isomorphism class. Thus the j -invariant refers to two curves, one with $(p+1)^2$ points, which is the one we consider in the sequel, and its twist with $(p-1)^2$ points.

The graph of 2 and 3-isogenies in that case looks like in Figure 2. Choosing $p = 2^{n_A}3^{n_B} - 1$ and $E : Y^2 = X^3 + X$ as starting curve means that the curves have $(p+1)^2 = 2^{2n_A}3^{2n_B}$ points and that all $2^{2n_A} - 1$ points of order 2^{n_A} and all $3^{2n_B} - 1$ of order 3^{n_B} are defined over \mathbb{F}_{p^2} . Computing degree 2^{n_A} or 3^{n_B} -isogenies is very efficient using Vélú’s formulas as the kernels are subgroups defined over \mathbb{F}_{p^2} . The same holds for all curves in the graph. The parameters are chosen so that $2^{n_A} \approx 3^{n_B} \approx \sqrt{p}$. Here and in the following A and B simply stand for Alice and Bob and are not related to the coefficients of curves in Montgomery form.

In SIDH, Alice works with 2^{n_A} isogenies and Bob works with 3^{n_B} isogenies. Note that unlike in CRS and CSIDH the degrees are known publicly, but there are about \sqrt{p} choices left because of all the directions; to see this we need to look at the number of choices for a kernel of these maps. The points of order 2^{n_A} form a space of dimension two, we can find points P_A and Q_A of order 2^{n_A} with $Q_A \neq \langle P_A \rangle$. Then the subgroups of order 2^{n_A} are given by $\langle P_A \rangle$, $\langle P_A + Q_A \rangle$, $\langle P_A + 2Q_A \rangle$, \dots , $\langle P_A + (2^{n_A} - 1)Q_A \rangle$ and $\langle Q_A \rangle$; each of these subgroups determines a unique 2^{n_A} isogeny.

The system parameters for SIDH are p and E as above, a basis P_A, Q_A of the points of

order 2^{n_A} on E , and similarly a basis P_B, Q_B of the points of order 3^{n_B} on E .

Alice picks a secret $0 \leq a < 2^{n_A}$, computes $T_A = P_A + aQ_A$ and the isogeny φ_A with kernel $\langle T_A \rangle$, landing at E_A , which is (part of) her public key. Similarly, Bob picks a secret $0 \leq b < 3^{n_B}$, computes $T_B = P_B + bQ_B$ and the isogeny φ_B with kernel $\langle T_B \rangle$ to E_B . In SIDH isomorphism classes are represented by the j -invariants, a suitable choice as isomorphisms over extensions of \mathbb{F}_{p^2} are considered

One difficulty in defining this system is that Alice cannot compute an isogeny φ'_A on E_B that matches φ_A translated by φ_B without having more information on φ_B , but φ_B is Bob's secret, so cannot be given to Alice. The way out found by Jao and De Feo [JF11] is to include additional points in the public keys of Alice and Bob, namely Bob also computes and publishes $\varphi_B(P_A)$ and $\varphi_B(Q_A)$. With that information, Alice can compute $T'_A = \varphi_B(P_A) + a\varphi_B(Q_A) = \varphi(T_A)$ and the isogeny φ'_A with kernel $\langle T'_A \rangle$, landing at E_{BA} . The use of the image points means that E_{BA} and Bob's E_{AB} are isomorphic. SIDH uses the j -invariant of the resulting curve to compute a shared key.

$$\begin{array}{ccc} E & \xrightarrow{\varphi_A} & E_A \\ \downarrow \varphi_B & & \downarrow \varphi'_B \\ E_B & \xrightarrow{\varphi'_A} & E_{AB} \end{array}$$

Figure 4: SIDH key exchange for Alice and Bob.

In summary, Alice's secret key is a and her public key is $(E_A, \varphi_A(P_B), \varphi_A(Q_B))$. Bob's secret key is b and his public key is $(E_B, \varphi_B(P_A), \varphi_B(Q_A))$.

Computing φ_A in one go would be very inefficient because the cost grows linearly with the degree and φ_A has degree 2^{n_A} . Hence, SIDH decomposes this 2^{n_A} isogeny into n_A computations of 2 isogenies. These start with a 2 isogeny φ_2 with kernel $\langle 2^{n_A-1}T_A \rangle$ and compute the image $\varphi_2(T_A)$, which is a point of order 2^{n_A-1} on a 2-isogenous curve, so that computing φ_A is the same as computing φ_2 followed by a 2^{n_A-1} isogeny with kernel $\langle \varphi_2(T_A) \rangle$. This is the same approach as described in [Pan21, Section 2.2.3]. Likewise, the images of P_B and Q_B are pushed through the 2 isogenies. A fast sequence of steps is proposed in [DFJP14] to reduce the cost of this computation.

One problem, recognized by [GPST16], is that Alice cannot validate the key she receives from Bob. An evil Bob can perform a reaction attack on Alice to learn a by sending her malformed public keys, e.g. sending $(E_B, \varphi_B(P_A), \varphi_B(Q_A) + 2^{n_A-1}\varphi_B(P_A))$ has Alice compute the same j -invariant as Bob if and only if a is even (because then $T'_A = \varphi_B(P_A) + a\varphi_B(Q_A) = \varphi_B(P_A) + a(\varphi_B(Q_A) + 2^{n_A-1}\varphi_B(P_A))$ as $2^{n_A}\varphi_B(P_A) = \infty_B$), learning one bit of Alice's secret. See Section 9 and talks in week 5 for how to deal with this issue and week 6 for details on the attack.

8 Security analysis of SIDH

For SIDH an attacker is confronted with the problem of finding E_{AB} given $E, E_A, \varphi_A(P_B), \varphi_A(Q_B), E_B, \varphi_B(P_A)$, and $\varphi_B(Q_A)$. The additional points have raised some concern but no attack for balanced $2^{n_A} \approx 3^{n_B}$, such as the parameters proposed in SIKE (see below), is known; see also the talk on torsion-point attacks in week 6.

Alice’s and Bob’s key spaces have size $\approx \sqrt{p}$, so meet-in-the-middle attacks run in time roughly $\sqrt[4]{p}$. Analysis [ACC⁺18] has shown that the cost of these attacks is typically underestimated, meaning that smaller parameters would offer sufficient security, see week 6 for details.

On the quantum side, similarly [JS19] showed that the attack costs of the so-called claw finding attack with $\sqrt[4]{p}$ underestimate security when taking into account the full cost of quantum computation (RAM model) so that choosing parameters to protect against non-quantum attacks suffices to remain secure against quantum attackers, see week 8.

9 Complete instantiations of isogeny-based encryption

The only isogeny-based submission to the NIST competition is SIKE [JAC⁺] by Jao, Azarderakhsh, Campagna, Costello, De Feo, Hess, Jalali, Koziel, LaMacchia, Longa, Naehrig, Renes, Soukharev, and Urbanik. SIKE is based on SIDH and uses a transformation to achieve CCA security and prove that ciphertexts are properly generated. Of the candidates that advanced to the second round, SIKE has the smallest public keys and the smallest combined size of message and public key.

SIDH for use in TLS 1.3 has been tested by Cloudflare <https://blog.cloudflare.com/sidh-go/> in a hybrid construct with regular elliptic-curve cryptography.

References

- [ACC⁺18] Gora Adj, Daniel Cervantes-Vázquez, Jesús-Javier Chi-Domínguez, Alfred Menezes, and Francisco Rodríguez-Henríquez. On the cost of computing isogenies between supersingular elliptic curves. In *SAC*, volume 11349 of *Lecture Notes in Computer Science*, pages 322–343. Springer, 2018.
- [Bia21] Jean-Francois Biasse. Ideal class groups, 2021. <http://www.usf-crypto.org/class-groups/>.
- [BKV19] 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.
- [BL07] Daniel J. Bernstein and Tanja Lange. Faster addition and doubling on elliptic curves. In *ASIACRYPT*, volume 4833 of *Lecture Notes in Computer Science*, pages 29–50. Springer, 2007.
- [BLMP19] Daniel J. Bernstein, Tanja Lange, Chloe Martindale, and Lorenz Panny. Quantum circuits for the CSIDH: optimizing quantum evaluation of isogenies. In

- EUROCRYPT*, Lecture Notes in Computer Science, page to appear. Springer, 2019. <https://ia.cr/2018/1059>.
- [BS20] Xavier Bonnetain and André Schrottenloher. Quantum security analysis of CSIDH. In *EUROCRYPT (2)*, volume 12106 of *Lecture Notes in Computer Science*, pages 493–522. Springer, 2020.
- [CGL09] Denis Xavier Charles, Eyal Z. Goren, and Kristin E. Lauter. Cryptographic hash functions from expander graphs. *J. Cryptology*, 22(1):93–113, 2009. <https://ia.cr/2006/021>.
- [CJS14] Andrew M. Childs, David Jao, and Vladimir Soukharev. Constructing elliptic curve isogenies in quantum subexponential time. *J. Mathematical Cryptology*, 8(1):1–29, 2014. <https://arxiv.org/abs/1012.4019>.
- [CLM⁺18] Wouter Castryck, Tanja Lange, Chloe Martindale, Lorenz Panny, and Joost Renes. CSIDH: An efficient post-quantum commutative group action. In *ASIACRYPT*, volume 11274 of *Lecture Notes in Computer Science*, pages 395–427. Springer, 2018. <https://ia.cr/2018/383>.
- [Cou06] Jean-Marc Couveignes. Hard Homogeneous Spaces, 2006. IACR Cryptology ePrint Archive 2006/291. <https://ia.cr/2006/291>.
- [DFJP14] Luca De Feo, David Jao, and Jérôme Plût. Towards quantum-resistant cryptosystems from supersingular elliptic curve isogenies. *Journal of Mathematical Cryptology*, 8(3):209–247, 2014. IACR Cryptology ePrint Archive 2011/506. <https://ia.cr/2011/506>.
- [DKS18] Luca De Feo, Jean Kieffer, and Benjamin Smith. Towards practical key exchange from ordinary isogeny graphs. In *ASIACRYPT*, volume 11274 of *Lecture Notes in Computer Science*, pages 365–394. Springer, 2018. <https://ia.cr/2018/485>.
- [FO99] Eiichiro Fujisaki and Tatsuaki Okamoto. Secure integration of asymmetric and symmetric encryption schemes. In *CRYPTO*, volume 1666 of *Lecture Notes in Computer Science*, pages 537–554. Springer, 1999.
- [GHS02] Steven D. Galbraith, Florian Hess, and Nigel P. Smart. Extending the GHS Weil descent attack. In *EUROCRYPT*, volume 2332 of *Lecture Notes in Computer Science*, pages 29–44. Springer, 2002.
- [GPST16] 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. IACR Cryptology ePrint Archive 2016/859. <https://ia.cr/2016/859>.

- [JAC⁺] David Jao, Reza Azarderakhsh, Matthew Campagna, Craig Costello, Luca De Feo, Basil Hess, Amir Jalali, Brian Koziel, Brian LaMacchia, Patrick Longa, Michael Naehrig, Joost Renes, Vladimir Soukharev, and David Urbanik. SIKE. Submission to [Nat16]. <http://sike.org>.
- [JF11] 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://eprint.iacr.org/2011/506/20110918:024142>.
- [JMV05] David Jao, Stephen D. Miller, and Ramarathnam Venkatesan. Do all elliptic curves of the same order have the same difficulty of discrete log? In *ASIACRYPT*, volume 3788 of *Lecture Notes in Computer Science*, pages 21–40. Springer, 2005. <https://ia.cr/2004/312>.
- [JS19] Samuel Jaques and John M. Schanck. Quantum cryptanalysis in the RAM model: Claw-finding attacks on SIKE. Cryptology ePrint Archive, Report 2019/103, 2019. <https://ia.cr/2019/103>.
- [Kup05] Greg Kuperberg. A subexponential-time quantum algorithm for the dihedral hidden subgroup problem. *SIAM J. Comput.*, 35(1):170–188, 2005. <https://arxiv.org/abs/quant-ph/0302112>.
- [Lan20] Tanja Lange. SD8 (Post-Quantum Cryptography) – Part 6: Isogeny-Based Cryptography. Technical Report N 2274, ISO/IEC JTC 1/SC 27/WG 2, 2020. <https://www.din.de/resource/blob/721042/4f1941ac1de9685115cf53bc1a14ac61/sc27wg2-sd8-data.zip>.
- [Mon87] Peter L. Montgomery. Speeding the Pollard and elliptic curve methods of factorization. *Mathematics of Computation*, 48(177):243–264, 1987.
- [Nat16] National Institute of Standards and Technology. Post-quantum cryptography standardization, December 2016. <https://csrc.nist.gov/Projects/Post-Quantum-Cryptography/Post-Quantum-Cryptography-Standardization>.
- [Pan21] Lorenz Panny. Elliptic curves and isogenies: The good bits, 2021. https://yx7.cc/docs/misc/isog_bristol_notes.pdf.
- [Pei20] Chris Peikert. He gives c-sieves on the CSIDH. In *EUROCRYPT (2)*, volume 12106 of *Lecture Notes in Computer Science*, pages 463–492. Springer, 2020.
- [RS06] Alexander Rostovtsev and Anton Stolbunov. Public-key cryptosystem based on isogenies, 2006. IACR Cryptology ePrint Archive 2006/145. <https://ia.cr/2006/145>.

- [Sho97] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput.*, 26(5):1484–1509, 1997. <https://arxiv.org/abs/quant-ph/9508027>.
- [Sto10] Anton Stolbunov. Constructing public-key cryptographic schemes based on class group action on a set of isogenous elliptic curves. *Adv. in Math. of Comm.*, 4(2):215–235, 2010.
- [Sto11] Anton Stolbunov. *Cryptographic Schemes Based on Isogenies*. PhD thesis, Norwegian University of Science and Technology, 2011.
- [Sut12] Andrew V. Sutherland. Isogeny volcanoes. In *ANTS X*, volume 1 of *The Open Book Series*, pages 507–530. Mathematical Sciences Publishers, 2012. <https://arxiv.org/abs/1208.5370>.
- [V71] Jacques Vélu. Isognies entre courbes elliptiques. *Comptes Rendus de l'Académie des Sciences de Paris*, 273:238–241, 1971.

Acknowledgment

This text grew out of a chapter I contributed to Standing Document 8 (SD8) [Lan20] of ISO SC 27 / JTC 1 working group 2. Please let me know of any typos you find. I would like to acknowledge feedback by Chris Swart and Matt Campagna on versions of the SD8 file.