ON THE MODULAR OSIDH PROTOCOL

LEONARDO COLÒ AND DAVID KOHEL

ABSTRACT. We introduce the use of modular curves with level structure to improve the complexity of the modular approach to the previously introduced OSIDH protocol. In addition, we augment the protocol with the use of 'descending' isogenies at non-split primes for the chosen orientation. This has the advantage of making effective use of all small primes for added efficiency, enlarges the size of the class group of the oriented order to secure against class group attacks, and attains mixing in a larger set of supersingular points. We illustrate the approach with a specific choice of orientation and conclude with a security analysis of the revised protocol.

1. INTRODUCTION

The oriented supersingular isogeny Diffie–Hellman (OSIDH) protocol [5] is based on an effective class group action on supersingular elliptic curves augmented by an orientation — an explicit and effective quadratic subring $\mathcal{O}$ of the endomorphism ring of a supersingular elliptic curve. In particular, the orientation allows two parties to arbitrarily extend two commuting isogeny directions, cut out by smooth ideals in $\mathcal{O}$, in order to compose two random walks in the class groups, acting on a set of curves orientated by $\mathcal{O}$.

In this protocol, an elliptic curve E is constructed by means of an ℓ-isogeny chain, orienting E by an order $\mathcal{O}_n = \mathbb{Z} + \ell^n \mathcal{O}_K$ of large smooth index ℓ^n in a maximal order $\mathcal{O}_K$ of class number 1. In order to conceal the orientation during key exchange, the split isogeny directions are precomputed as q-isogeny chains on the curve obtained by a random walk, and passed between the two parties. Each party composes isogeny chains to reconstruct their random smooth isogeny on the exchanged curve, resulting in a common secret. As observed by Dartois and De Feo [6], this initial scheme leaks too much information about the class group $\mathcal{C}\ell(\mathcal{O}_n)$, which can be used to construct an endomorphism, and reconstruct the ascending chain from $\mathcal{O}_n$ back to $\mathcal{O}_K$. In order to defeat this attack, we introduce descents at non-split (inert or ramified) primes and at larger split primes, which serves to make efficient use of all small primes, and provides exponential growth of the discriminant to defeat the class group attack.

Any isogeny-based protocol essentially concerns only properties of the isogeny graph, for which the vertices are isomorphism classes of curves, captured by their j-invariant (over the algebraic closure). The supersingular isogeny graph used in OSIDH (and SIDH and its variants) are graphs for which the vertices are j-invariants of supersingular elliptic curves, points on the modular curve $X(1)$, and the edges are conveyed by correspondences in $X(1) \times X(1)$, defined by a modular polynomial $\Phi_\ell(X, Y)$ which vanishes on the pairs of ℓ-isogenous j-invariants. These modular polynomials are too large to use in practice for more than tiny primes ℓ. In what follows we replace the modular curve $X(1)$ with a modular curve of higher level, and in particular develop the necessary theory for working with supersingular Weber invariants and their correspondences.

2. ISOGENY GRAPHS AND ORIENTATIONS

Modular isogeny graphs. We recall that an isogeny graph $G = G_S(E)$, of an elliptic curve E/k is a graph whose vertices are elliptic curves $\bar{k}$-isogenous to E, and whose directed edges are isogenies of prime degree $\ell \in S$. If $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ is a congruence subgroup of level N, then $G_S(E, \Gamma)$ is the graph whose vertices are pairs $(E, \Gamma(P, Q))$ where $\Gamma(P, Q)$ is the orbit of an ordered basis (P, Q) of $E[N]$ such that the Weil pairing $e_n(P, Q) = \zeta_N$ is a fixed root of unity in $\bar{k}$, and whose edges are isogenies of prime degree $\ell \in S$. The orbit is defined with respect to the left action of Γ, given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} (P, Q) = (aP + bQ, cP + dQ).$$

We identify $G_S(E)$ with $G_S(E, \mathrm{SL}_2(\mathbb{Z}))$ and for each inclusion $\Gamma_1 \subset \Gamma_2$ we obtain a morphism of graphs

$$G_S(E, \Gamma_1) \to G_S(E, \Gamma_2).$$

We can identify the vertices of $G_S(E, \Gamma)$ with points on the modular curve $X(\Gamma)$ and edges with points on the modular curve $X(\Gamma \cap \Gamma_0(\ell))$ for ℓ in S different from the level of Γ, and otherwise $X(\Gamma \cap \Gamma_0(\ell^i))$, where i is the smallest exponent such that Γ is not contained in $\Gamma_0(\ell^i)$. When $S = \ell$, we will write simply $G_\ell(E)$ and $G_\ell(E, \Gamma)$.

Orientations. We introduced the notion of orientation on elliptic curves and isogeny graphs in [5], which we briefly recall here. Let E be a supersingular elliptic curve over a finite field k of characteristic p, and denote by $\mathrm{End}(E)$ the full endomorphism ring. We assume moreover that k contains $\mathbb{F}_{p^2}$ and E is in an isogeny class such that $\mathrm{End}_k(E) = \mathrm{End}(E)$. We denote by $\mathrm{End}^0(E)$ the $\mathbb{Q}$-algebra $\mathrm{End}(E) \otimes_{\mathbb{Z}} \mathbb{Q}$, the unique quaternion algebra over $\mathbb{Q}$ ramified at p and ∞, and write $\mathrm{SS}(p)$ for the set of supersingular elliptic curves over $\overline{\mathbb{F}}_p$ up to isomorphism. For a congruence subgroup Γ in $\mathrm{SL}_2(\mathbb{Z})$, we write $\mathrm{SS}(p, \Gamma)$ for the set of isomorphism classes equipped with a Γ level structure. This lets us define supersingular graphs $G_S(\mathrm{SS}(p))$ and $G_S(\mathrm{SS}(p, \Gamma))$ on these vertex sets.

Let K be a quadratic imaginary field of discriminant D_K with maximal order $\mathcal{O}_K$. If E is a supersingular elliptic curve over $\overline{\mathbb{F}}_p$, there exists an embedding $\iota : K \to \mathrm{End}^0(E)$ if and only if p is inert or ramified in $\mathcal{O}_K$, and there exists an order $\mathcal{O} \subseteq \mathcal{O}_K$ such that $\iota(\mathcal{O}) = \iota(K) \cap \mathrm{End}(E)$.

Definition. *A K-orientation on an elliptic curve E/k is a homomorphism $\iota : K \hookrightarrow \mathrm{End}^0(E)$. An $\mathcal{O}$-orientation on E is a K-orientation such that the image of the restriction of ι to $\mathcal{O}$ is contained in* $\mathrm{End}(E)$. *We write* $\mathrm{End}((E, \iota))$ *for the order* $\mathrm{End}(E) \cap \iota(K)$ *in* $\iota(K)$. *An $\mathcal{O}$-orientation is primitive if ι induces an isomorphism of $\mathcal{O}$ with* $\mathrm{End}((E, \iota))$.

With this notation, we let $\mathrm{SS}_K(p)$ be the set of K-oriented supersingular elliptic curves over $\overline{\mathbb{F}}_p$, up to K-isomorphism, let $\mathrm{SS}_{\mathcal{O}}(p)$ be the subset of $\mathcal{O}$-oriented curves, and we denote the subset of primitive $\mathcal{O}$-oriented curves by $\mathrm{SS}^{pr}_{\mathcal{O}}(p)$. An element of $\mathrm{SS}^{pr}_{\mathcal{O}}(p)$ consists of the data of

- a supersingular elliptic curve $E/\overline{\mathbb{F}}_p$,
- a primitive orientation $\iota : \mathcal{O} \hookrightarrow \mathrm{End}(E)$.

The additional structure of a p-orientation is a homomorphism $\rho : \mathcal{O} \longrightarrow \overline{\mathbb{F}}_p$. We note that $\mathrm{End}(E)$ is equipped with a p-orientation $\rho : \mathrm{End}(E) \hookrightarrow \overline{\mathbb{F}}_p$ given by its action on the 1-dimensional vector space of invariant differentials, with kernel $\mathfrak{P}$:

$$\alpha^* \omega_E = \rho(\alpha)\omega \text{ for all } \alpha \in \mathrm{End}(E).$$

For this fixed $\mathrm{End}(E)/\mathfrak{P} \hookrightarrow \mathbb{F}_{p^2} \subseteq \overline{\mathbb{F}}_p$, a p-orientation on $\mathcal{O}$ is determined by ι:

$$\mathcal{O} \longrightarrow \mathcal{O}/\mathfrak{p} \xrightarrow{\iota} \mathrm{End}(E)/\mathfrak{P} \subseteq \overline{\mathbb{F}}_p.$$
$$\rho$$

We denote by $\mathrm{SS}_{\mathcal{O}}(\rho)$ the set of oriented supersingular elliptic curves with ρ induced by ι and $\mathrm{End}(E)/\mathfrak{P} \hookrightarrow \overline{\mathbb{F}}_p$, and $\mathrm{SS}_{\mathcal{O}}(\overline{\rho})$ the opposite p-orientation class. In the notation of Belding [**?**, Section 2.3.2], the set $\mathrm{SS}^{pr}_{\mathcal{O}}(\rho)$ consists of *normalized* optimal embeddings.

Remark. The p-orientation ρ restricts to the same subset $\mathrm{SS}^{pr}_{\mathcal{O}}(\rho)$ of $\mathrm{SS}^{pr}_{\mathcal{O}}(p)$ as the image of the canonical lift:

$$\rho : \mathscr{E}\ell\ell(\mathcal{O}) \longrightarrow \mathrm{SS}^{pr}_{\mathcal{O}}(p),$$

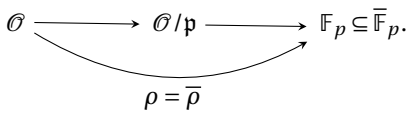
in the notation of Onuki [11].

Remark. The Frobenius automorphism σ induces an isomorphism $\mathrm{SS}_{\mathcal{O}}(\rho) \to \mathrm{SS}_{\mathcal{O}}(\overline{\rho})$ taking (E, ι) to (E^σ, ι^σ), noting that $\overline{\rho} = \sigma \circ \rho$.

Since K embeds in B, the prime p is either ramified or inert. In the former case, $\rho = \bar{\rho}$, hence

$$\mathrm{SS}_{\mathcal{O}}(p) = \mathrm{SS}_{\mathcal{O}}(\rho) = \mathrm{SS}_{\mathcal{O}}(\bar{\rho}).$$

This follows since the image of ρ lies in $\mathbb{F}_p$:



In the latter case, $\mathrm{SS}_{\mathcal{O}}(p)$ decomposes into the disjoint union of the set of normalized oriented curves and its conjugate:

$$\mathrm{SS}_{\mathcal{O}}(p) = \mathrm{SS}_{\mathcal{O}}(\rho) \cup \mathrm{SS}_{\mathcal{O}}(\bar{\rho}) = \mathrm{SS}_{\mathcal{O}}(\rho) \cup \mathrm{SS}_{\mathcal{O}}(\rho)^{\sigma}.$$

With this notation for $\mathrm{SS}_{\mathcal{O}}(\rho)$, distinguished from $\mathrm{SS}_{\mathcal{O}}(p)$, we restate the theorem from Colò and Kohel [5].

The set $\mathrm{SS}_{\mathcal{O}}(\rho)$ admits a transitive group action:

$$\begin{aligned} \mathcal{C}\ell(\mathcal{O}) \times \mathrm{SS}_{\mathcal{O}}(\rho) &\longrightarrow \mathrm{SS}_{\mathcal{O}}(\rho) \\ ([\mathfrak{a}], E) &\longmapsto E/E[\mathfrak{a}] \end{aligned}$$

where $\mathfrak{a}$ is any representative ideal coprime to the index $[\mathcal{O}_K : \mathcal{O}]$ so that the isogeny $\phi_{\mathfrak{a}} : E \to E/E[\mathfrak{a}]$ is horizontal. When restricted to primitive $\mathcal{O}$-oriented curves, we obtain the following classical result, extending the standard result for CM elliptic curves

Theorem. $\mathrm{SS}^{pr}_{\mathcal{O}}(\rho)$ *is a torsor for* $\mathcal{C}\ell(\mathcal{O})$.

Remark. In particular, we note that, if it is not empty, $\mathrm{SS}_K(\rho)$ is an infinite set, the disjoint union of its finite subsets $\mathrm{SS}^{pr}_{\mathcal{O}}(\rho)$ for all $\mathcal{O} \subset \mathcal{O}_K$.

Remark. This result is restated as [11, Th. 3.4], and [11, Th. 3.3] implies the equality

$$\mathrm{SS}^{pr}_{\mathcal{O}}(p) = \mathrm{SS}^{pr}_{\mathcal{O}}(\rho) \cup \mathrm{SS}^{pr}_{\mathcal{O}}(\bar{\rho}).$$

In [**?**, Prop. 4.2], the fact that these two orbits are disjoint when p is inert in $\mathcal{O}$ is asserted.
In this context the class group action is given by

$$(\mathcal{C}\ell(\mathcal{O}) \rtimes \langle\sigma\rangle) \times \mathrm{SS}^{pr}_{\mathcal{O}}(p) \longrightarrow \mathrm{SS}^{pr}_{\mathcal{O}}(p)$$

By lifting σ to complex conjugation, the dihedral group $\mathcal{C}\ell(\mathcal{O}) \rtimes \langle\sigma\rangle$, can be identified with the Galois group $\mathcal{G}a\ell(H_{\mathcal{O}}/\mathbb{Q})$ acting on the set of canonical lifts $\mathcal{E}\ell\ell(\mathcal{O})$.

Oriented isogeny graphs. The notion of orientation lets us define K-oriented isogeny graphs $G_S(E,K)$ whose vertices are elliptic curves in the $\bar{k}$-isomorphism class of E up to K-isomorphism. In view of the theorem, the graph is infinite and decomposes into finite subgraphs $G_S(E,\mathcal{O})$ of curves oriented by an order $\mathcal{O}$ in K. When E is supersingular we obtain the graphs $G_S(\mathrm{SS}_K(\rho))$ and $G_S(\mathrm{SS}_{\mathcal{O}}(\rho))$ on the sets of oriented supersingular curves. For any congruence subgroup Γ, of level coprime to the characteristic, we have covering graphs $G_S(E,K,\Gamma) \to G_S(E,K)$ and $G_S(E,\mathcal{O},\Gamma) \to G_S(E,\mathcal{O})$, and we write $G_S(\mathrm{SS}_K(\rho,\Gamma))$ and $G_S(\mathrm{SS}_{\mathcal{O}}(\rho,\Gamma))$ for E supersingular.

Remark. The action of ideals through isogenies lets us define an action on the graph $G_S(\mathrm{SS}_{\mathcal{O}}(\rho,\Gamma))$ with Γ level structure, by a ray class group $\mathcal{C}\ell(\mathcal{O},\Gamma)$ preserving the level structure.

$$\begin{aligned} \mathcal{C}\ell(\mathcal{O},\Gamma) \times \mathrm{SS}_{\mathcal{O}}(\rho,\Gamma) &\longrightarrow \mathrm{SS}_{\mathcal{O}}(\rho,\Gamma) \\ \big([\mathfrak{a}], (E,\Gamma(P,Q)\big) &\longmapsto (\phi_{\mathfrak{a}}(E), \Gamma(\phi_{\mathfrak{a}}(P), \phi_{\mathfrak{a}}(Q))) \end{aligned}$$

In the following section we work with a congruence subgroup Γ_{24} of level 48 fixing the Weber invariants. This lets us develop efficient isogeny relations in terms of Weber functions to navigate the graph $G_S(\mathrm{SS}_{\mathcal{O}}(\rho,\Gamma_{24}))$, after which we apply the morphisms given by the forgetful functors:

$$G_S(\mathrm{SS}_{\mathcal{O}}(\rho,\Gamma_{24})) \to G_S(\mathrm{SS}(\rho,\Gamma_{24})) \to G_S(\mathrm{SS}(\rho)).$$

3. On clouds and eddies

In what follows we describe the local structure of oriented isogeny graphs and their associated class group actions. Given an order $\mathcal{O}$ in an imaginary quadratic field K, for simplicity of notation we write $\mathcal{O}(M)$ for the order $\mathbb{Z}+M\mathcal{O}$ of index M, and for fixed prime ℓ we write $\mathcal{O}_n$ for $\mathcal{O}(\ell^n)$. Moreover we denote the kernel

$$U(\mathcal{O},M) = \ker(\mathcal{C}\ell(\mathcal{O}(M)) \to \mathcal{C}\ell(\mathcal{O}))$$

which will be identified with the stabilizer subgroup of an isomorphism class of a curve oriented by $\mathcal{O}$.

Isogeny chains and clouds. In order to study the structure of ℓ-isogeny graphs, we introduce some notation for paths and neighborhoods in the graph.

Definition. *An ℓ-isogeny chain of length n from E_0 to E is a sequence of isogenies of degree ℓ:*

$$E_0 \xrightarrow{\phi_0} E_1 \xrightarrow{\phi_1} E_2 \xrightarrow{\phi_2} \dots \xrightarrow{\phi_{n-1}} E_n = E.$$

The ℓ-isogeny chain is without backtracking if $\ker(\phi_{i+1}\circ\phi_i) \neq E_i[\ell]$ *for each* $i = 0,\dots,n-1$.

Remark. An ℓ-isogeny chain corresponds to a path in the underlying ℓ-isogeny graph $G_\ell(E)$. The concept of backtracking, however, is more subtle in the ℓ-isogeny graph with level structure Γ. In particular, an ℓ-isogeny chain with $\phi_{i+1}\circ\phi_i = E_i[\ell]$ to $E_{i+2} = E_i$ may induce a nontrivial automorphism of Γ-orbits in $E_i[N]$. One of the interests of introducing level structure on graphs is to avoid backtracking, loops and cycles of order 2 in an ℓ-isogeny graphs.

We recall that a K-orientation on E_0 induces a well-defined notion of horizontal, descending and ascending ℓ-isogenies, and the depth at ℓ of a curve E in the isogeny class, see [5, §2]. If E_0 admits a K-orientation, we say that the isogeny chain is horizontal, descending, or ascending if each ϕ_i is horizontal, descending or ascending, respectively. Since there is at most one ascending direction, an ℓ-isogeny chain without backtracking is descending if and only if ϕ_0 is descending.

In order to discuss the local neighborhood of a graph, and conduct a breadth-first search, we introduce the notion of an ℓ-isogeny cloud around E.

Definition. An *ℓ-isogeny cloud of radius r* at E is a subgraph of $G_\ell(E)$, whose paths from E extend to length r.

A cloud of radius r is therefore the neighborhood of E in the ℓ-isogeny graph consisting of curves at distance less than or equal to r.

Whirlpools, vortices, and eddies. Suppose that ℓ is a prime, $\mathcal{O}$ an ℓ-maximal order in K, and E an $\mathcal{O}$-oriented elliptic curve. Set $\mathcal{O}_r = \mathcal{O}(\ell^r)$ for all positive integers r. The subgraph of descending ℓ-isogenies in the ℓ-isogeny cloud of radius r at E in $G_\ell(E,K)$ admits an action of $U(\mathcal{O},\ell^r)$. In view of the exact sequence of class groups,

$$1 \to U(\mathcal{O},\ell^r) \to \mathcal{C}\ell(\mathcal{O}_r) \to \mathcal{C}\ell(\mathcal{O}) \to 1,$$

we consider the corresponding decomposition into subgraphs.

Definition. *Let $\mathcal{O}$ be an ℓ-maximal order and E an $\mathcal{O}$-oriented elliptic curve. A* vortex *at ℓ is the ℓ-isogeny subgraph $G_\ell(E,\mathcal{O})$ of $G_\ell(E,K)$ equipped with the action by $\mathcal{C}\ell(\mathcal{O})$. A whirlpool the union of the subgraphs $G_\ell(E,\mathcal{O}_n)$ in $G_\ell(E,K)$ equipped with the compatible actions of $\mathcal{C}\ell(\mathcal{O}_n)$. An eddy at E is the subgraph of ℓ-isogenies descending from E, equipped with the compatible actions of $U(\mathcal{O},\ell^n)$. The restriction of the whirlpool (or eddy) to $G_\ell(E,\mathcal{O}_r)$ is called the whirlpool (or eddy) of depth r.*

Even at a non-split prime ℓ, in which every ℓ-isogeny is descending, the distinction between cloud and eddy, a set and G-set, respectively, is nontrivial. The structure of a cloud can be constructed from ℓ-isogeny relations, on which the G-set structure can be enumerated, but an effective G-set structure on the eddy is entirely determined by its transitive action on one descending path. An effective action of $U(\mathcal{O},\ell^r)$ yields a compression from an ℓ-cloud of order ℓ^r elements to a ℓ-isogeny chain of length r.

4. MODULAR ISOGENY CHAINS

A modular ℓ-isogeny chain is determined by a set of (supersingular) moduli points on a modular curve $\mathscr{X} = X(\Gamma)/\mathbb{F}_p$ for some $\Gamma \subset \Gamma(N)$, and edge relations given by points in the cover,

$$\mathscr{X}(\Gamma_0(\ell)) \longrightarrow \mathscr{X} \times \mathscr{X}$$

over a given pair of moduli points, or by $\mathscr{X}(\Gamma_0(\ell^{t+1})) \to \mathscr{X} \times \mathscr{X}$ when $\Gamma \subset \Gamma_0(\ell^t)$. Here $\mathscr{X}(\Gamma_0(\ell))$ is the modular curve

$$\mathscr{X}(\Gamma_0(\ell)) = X(\Gamma_0(\ell) \cap \Gamma)$$

Remark. When working with a level structure Γ, the oriented points are associated to a class group $\mathscr{C}\ell(\mathscr{O}_K, \Gamma)$ and ℓ-isogeny chains with $\mathscr{C}\ell(\mathscr{O}_n, \Gamma)$. When representing the ideal classes by binary quadratic forms (or lattices in $\mathbb{C}$), the equivalence class is determined by a form or lattice with basis up to equivalence by Γ rather than by the full group $\mathrm{PSL}_2(\mathbb{Z})$. When $\Gamma = \Gamma_0(N)$, the class group $\mathscr{C}\ell(\mathscr{O})$ of the order of index N is efficiently modeled by reduction by $\mathrm{PSL}_2(\mathbb{Z})$ of binary quadratic forms of discriminant $\mathrm{disc}(\mathscr{O}) = N^2 D_K$, but can be equivalently modeled by forms of discriminant D_K and reduction by $\Gamma_0(N)$.

One advantage of working with a modular curve $\mathscr{X}$ of higher level, still of genus 0, is that the logarithmic coefficient size of the modular polynomials $\Phi_q(x, y)$ with respect to a degree one function is reduced by a factor of the degree $\mathscr{X} \to X(1)$. If we consider the smallest correspondence, for $q = 2$, for the j-invariant we have the modular polynomial

$$\begin{aligned} x^3 - x^2y^2 + y^3 + 1488x^2y + 1488xy^2 - 162000x^2 - 162000y^2 + 40773375xy \\ + 8748000000x + 8748000000y - 157464000000000. \end{aligned}$$

In comparison, for the group $\Gamma_0(3)$ of index 4, this becomes:

$$x^3 - x^2y^2 - 24x^2y - 24xy^2 - 729xy + y^3$$

and for the full congruence subgroup $\Gamma(3)$ of index 6 we have:

$$x^3 - x^2y^2 + 9xy + y^3 - 54$$

It is worth noting that the sparsity of monomials of the latter polynomial can be explained by the transformation

$$\Phi_q(\zeta_3 x, \zeta_3^q y) = \zeta_3^{q+1} \Phi_q(x, y),$$

of the family of modular polynomials with respect to the particular normalized modular function on $\mathscr{X} = X(3)$. In particular, only the monomials $x^i y^j$ satisfying $i + qj \bmod q + 1 \bmod 3$ can occur.

Weber modular polynomials. The best known reduction in coefficient size as well as in sparsity of coefficients is obtained for the Weber function $\mathfrak{f}$ of level 48,

$$\mathfrak{f}(\tau) = \zeta_{48}^{-1} \frac{\eta\left(\frac{\tau+1}{2}\right)}{\eta(\tau)},$$

which generates a degree-72 cover of the j-line, given by

$$j = \frac{(\mathfrak{f}^{24} - 16)^3}{\mathfrak{f}^{24}}.$$

The modular polynomials with respect to $\mathfrak{f}$ are the integral polynomials $\Phi_q(x, y)$ such that

$$\Phi_q(\mathfrak{f}(\tau), \mathfrak{f}(q\tau)) = 0.$$

Although the Weber function does not generate the full modular curve $X(48)$, which has genus 2689, it still satisfies a transformation giving the following symmetry of its induced modular polynomials.

Lemma 1. *The modular functions $\Phi_q(x, y)$ of prime level q with respect to the Weber function satisfies the transformation:*

$$\Phi_q(\zeta_{24} x, \zeta_{24}^q y) = \zeta_{24}^{q+1} \Phi_q(x, y),$$

with respect to a primitive 24-th root of unity ζ_{24}.

Asymptotically, modular polynomials have q^2 monomials, but in a practical range the sparseness is dictated by this transformation — on the order of $q^2/24$ monomials

$$\Phi_5(x,y) = x^6 - x^5y^5 + 4xy + y^6$$
$$\Phi_7(x,y) = x^8 - x^7y^7 + 7x^4y^4 - 8xy + y^8$$
$$\Phi_{11}(x,y) = x^{12} - x^{11}y^{11} + 11x^9y^9 - 44x^7y^7 + 88x^5y^5 - 88x^3y^3 + 32xy + y^{12}$$

— makes these polynomials attractive for constructing isogeny invariants. For example, the modular polynomial $\Phi_{71}(x,y)$ has exactly $3 \cdot 71 = 213$ nonzero coefficients, ignoring the symmetry $\Phi_q(x,y) = \Phi_q(y,x)$, which implies an even smaller number of distinct coefficients.

In the interest of constructing ℓ-isogeny chains, especially for $\ell = 2$ or $\ell = 3$, we note that the 48-level structure gives the modular polynomials $\Phi_2(x,y)$ and $\Phi_3(x,y)$ a particular form. We descend the 2-level structure by setting $t = -\mathfrak{f}^8$, so that

$$j = \left(\frac{t^3+16}{t}\right)^3$$

With respect to this function, we obtain the modular polynomial:

$$\Psi_2(x,y) = (x^2 - y)y + 16x$$

and the Weber modular polynomial $\Phi_2(x,y) = -\Psi_2(-x^8,-y^8)$ remains irreducible.[1] A similar descent of the 3-level to the function $r = \mathfrak{f}^3$, gives the modular polynomial

$$\Psi_3(x,y) = x^4 - x^3y^3 + 8xy + y^4,$$

such that $\Psi_3(r(\tau), r(3\tau)) = 0$, for which $\Phi_3(x,y) = \Psi_3(x^3,y^3)$ is irreducible. For a given supersingular Weber invariant, these relations determine orbits under multiplication by ζ_8 or ζ_3, but in view of the global relation of Lemma 1, the lift to the orbit can be chosen to be compatible with isogeny relations of other prime degrees.

Weber modular functions. The classically defined triple of Weber functions,

$$\mathfrak{f}(\tau) = \zeta_{48}^{-1}\frac{\eta\left(\frac{\tau+1}{2}\right)}{\eta(\tau)}, \quad \mathfrak{f}_1(\tau) = \frac{\eta\left(\frac{\tau}{2}\right)}{\eta(\tau)}, \quad \mathfrak{f}_2(\tau) = \sqrt{2}\cdot\frac{\eta(2\tau)}{\eta(\tau)},$$

are modular functions with rational integral q-expansions in $q^{1/24} = e^{2\pi i\tau/24}$ satisfy

$$j = \frac{(\mathfrak{f}^{24}-16)^3}{\mathfrak{f}^{24}} = \frac{(\mathfrak{f}_1^{24}+16)^3}{\mathfrak{f}^{24}} = \frac{(\mathfrak{f}_2^{24}+16)^3}{\mathfrak{f}^{24}},$$

on which the generators S and T of $\mathrm{SL}_2(\mathbb{Z})$ induce (see Gee [10]):

$$(\mathfrak{f},\mathfrak{f}_1,\mathfrak{f}_2)\circ S = (\mathfrak{f},\mathfrak{f}_2,\mathfrak{f}_1) \text{ and } (\mathfrak{f},\mathfrak{f}_1,\mathfrak{f}_2)\circ T = (\zeta_{48}^{-1}\mathfrak{f}_1,\zeta_{48}^{-1}\mathfrak{f},\zeta_{48}^{2}\mathfrak{f}_2).$$

It is well-known that the Weber functions satisfy $\mathfrak{f}^8 = \mathfrak{f}_1^8 + \mathfrak{f}_2^8$, and based on the identity

$$\zeta_{48}^{-1}\eta\left(\frac{\tau+1}{2}\right)\eta\left(\frac{\tau}{2}\right)\eta(2\tau) = \eta(\tau)^3$$

it follows from the definition of the Weber functions that $\mathfrak{f}\mathfrak{f}_1\mathfrak{f}_2 = \sqrt{2}$.

Setting $(\mathfrak{u}_0,\mathfrak{u}_1,\mathfrak{u}_2) = (\mathfrak{f},\zeta_{16}\mathfrak{f}_1(\tau),\zeta_{16}^{-1}\mathfrak{f}_2(\tau))$ the triple $(\mathfrak{u}_0,\mathfrak{u}_1,\mathfrak{u}_2)$ satisfies the common relations

$$j = \frac{(\mathfrak{u}_0^{24}-16)^3}{\mathfrak{u}^{24}} = \frac{(\mathfrak{u}_1^{24}-16)^3}{\mathfrak{u}^{24}} = \frac{(\mathfrak{u}_2^{24}-16)^3}{\mathfrak{u}^{24}},$$

and one verifies that the three orbits $\{\zeta_{24}^j\mathfrak{u}_i : j \in \mathbb{Z}/24\mathbb{Z}\}$ run over the 72 roots of the modular polynomial

$$(X^{24}-16)^3 - j(q)X^{24} \in \mathbb{Q}(\zeta_{24})[[q^{1/24}]],$$

[1] More correctly, the modular polynomial $\Psi_2(x,y)$ satisfies $\Psi_2(\mathfrak{f}_1^8(\tau),\mathfrak{f}_1^8(2\tau)) = 0$, where $\mathfrak{f}_1$ is the conjugate Weber function

$$\mathfrak{f}_1^8(\tau) = -\mathfrak{f}(\tau+3)^8 = \left(\frac{\eta(\frac{\tau}{2})}{\eta(\tau)}\right)^8$$

and hence $\Psi_2(\mathfrak{f}^8(\tau),\mathfrak{f}^8(2\tau-3)) = 0$. Nevertheless, this modular relation describes a 2-isogeny relation of the underlying curves, extending the parametrized 2-isogeny to a 4-isogeny, and will be used for defining our 2-isogeny chains.

and moreover, the functions $\mathfrak{u}_i$ satisfy transformations

$$(\mathfrak{u}_0,\mathfrak{u}_1,\mathfrak{u}_2)\circ S = (\mathfrak{u}_0,\zeta_8^{-1}\mathfrak{u}_2,\zeta_8\mathfrak{u}_1) \text{ and } (\mathfrak{u}_0,\mathfrak{u}_1,\mathfrak{u}_2)\circ T = (\zeta_{24}\mathfrak{u}_1,\zeta_{12}^{-1}\mathfrak{u}_0,\zeta_{24}\mathfrak{u}_2).$$

Weber curves. The map determined by the normalized Weber functions $(\mathfrak{u}_0^m:\mathfrak{u}_1^m:\mathfrak{u}_2^m:1)$ determines a *Weber modular curve* $\mathscr{W}_{3n}$ in $\mathbb{P}^3$

$$\mathscr{W}_{3n}:\begin{cases} X_0^n+X_1^n+X_2^n=0,\\ X_0X_1X_2=\sqrt{2}^m X_3^3\end{cases}$$

for m and n such that $mn=8$, with quotient Weber curve $\mathscr{W}_n$ defined as the image of $(\mathfrak{u}_0^{3m}:\mathfrak{u}_1^{3m}:\mathfrak{u}_2^{3m}:1)$ in $\mathbb{P}^3$:

$$\mathscr{W}_n:\begin{cases} X_0^n+X_1^n+X_2^n=48X_3^n,\\ X_0X_1X_2=\sqrt{2}^{3m}X_3^3.\end{cases}$$

These defining relations follow directly from the relations $\mathfrak{f}^8=\mathfrak{f}_1^8+\mathfrak{f}_2^8$ and $\mathfrak{f}\mathfrak{f}_1\mathfrak{f}_2=\sqrt{2}$, and the curves are equipped with maps $\mathscr{W}_{mn}\to\mathscr{W}_n$ for each product mn dividing 24.

Modular group. To each factorization $mn=24$, the Weber curve $\mathscr{W}_n$ in $\mathbb{P}^3$, defined by the triple of Weber functions $(\mathfrak{u}_0^m,\mathfrak{u}_1^m,\mathfrak{u}_2^m)$, comes equipped with an action of $\mathrm{PSL}_2(\mathbb{Z})$. We denote the kernel of the action by Γ_n, identifying the Weber curves with the modular curve $X(\Gamma_n)$. The action of $\mathrm{PSL}_2(\mathbb{Z})$ on Weber functions induces a representation in $\mathrm{GL}_3(\mathbb{Q}(\zeta_n))$ determined by the images of the generators S and T.

$$\begin{array}{ccc}\mathrm{PSL}_2(\mathbb{Z}) & \longrightarrow & \mathrm{GL}_3(\mathbb{Q}(\zeta_n))\\ S,T & \longmapsto & \begin{pmatrix}1&0&0\\0&0&\overline{\zeta}_8^m\\0&\zeta_8^m&0\end{pmatrix},\begin{pmatrix}0&\zeta_{24}^m&0\\\overline{\zeta}_{24}^{2m}&0&0\\0&0&\zeta_{24}^m\end{pmatrix}\end{array}$$

The image group is finite, whose kernel Γ_n is a normal congruence subgroup of $\mathrm{PSL}_2(\mathbb{Z})$. We reduce the computation of Γ_n by first proving that $\Gamma_1\cong\Gamma(2)$, then noting that $\Gamma_1\subset\Gamma_3\cap\Gamma_8=\Gamma_{24}$, which reduces to determining Γ_3 and the chain $\Gamma_1\subset\Gamma_2\subset\Gamma_4\subset\Gamma_8$.

Let $\boldsymbol{\mu}_m=\langle\zeta_m\rangle$ be the group of m-th roots of unity, and define the antidiagonal group,

$$\nabla(\boldsymbol{\mu}_m^2)=\left\{\left(\zeta_m^i,\zeta_m^j,\zeta_m^{-i-j}\right)\mid 0\le i,j<m\right\},$$

and the diagonal group $\Delta(\boldsymbol{\mu}_m)=\left\{\left(\zeta_m^i,\zeta_m^i,\zeta_m^i\right)\mid 0\le i<m\right\}$, each of which acts by coordinate scaling on $\mathbb{A}^3$, and if $m\equiv 0 \bmod 3$, then $\nabla(\boldsymbol{\mu}_m^2)\cap\Delta(\boldsymbol{\mu}_m)=\Delta(\boldsymbol{\mu}_3)$, otherwise the intersection is trivial. For each divisor mn of 24, the action of the antidiagonal group $\nabla(\boldsymbol{\mu}_m^2)$, extended to $\mathbb{P}^3$, stabilizes $\mathscr{W}_{mn}$.

Proposition 2. *For each product mn dividing 24, the group $\nabla(\boldsymbol{\mu}_m^2)$ acts on $\mathscr{W}_{mn}$ with quotient $\mathscr{W}_n$.*

- *If $m\not\equiv 0 \bmod 3$, then the group $\nabla(\boldsymbol{\mu}_m^2)$ acts faithfully and $\Gamma_{mn}/\Gamma_n\cong\nabla(\boldsymbol{\mu}_m^2)$.*
- *If $m\equiv 0 \bmod 3$, then group $\nabla(\boldsymbol{\mu}_m^2)$ acts with kernel $\Delta(\boldsymbol{\mu}_3)$, and $\Gamma_{mn}/\Gamma_n\cong\nabla(\boldsymbol{\mu}_m^2)/\Delta(\boldsymbol{\mu}_3)$.*

In particular, if m divides 8, the degree of $\mathscr{W}_{mn}\to\mathscr{W}_n$ is m^2 and the degree of $\mathscr{W}_{3n}\to\mathscr{W}_n$ is 3.

Proposition 3. *The Weber kernel group Γ_1 equals $\Gamma(2)$ and $\mathscr{W}_1\cong X(2)$.*

Proposition 4. *The Weber kernel group Γ_3 equals $\Gamma(2)\cap\Gamma_{ns}^+(3)$, and for each n dividing 8*

$$\Gamma_{3n}=\Gamma_n\cap\Gamma_{ns}^+(3).$$

It thus suffices to characterize the groups Γ_n for n dividing 8.

Proposition 5. *The Weber kernel group Γ_2 equals $\Gamma(4)$ and $\mathscr{W}_2=X(4)$.*

Proposition 6. *The Weber kernel group Γ_4 equals $\Gamma_s(8)$ and $\mathscr{W}_4=X_s(8)$.*

It remains to characterize the group Γ_8 under which the triple of functions $(\mathfrak{u}_0^3,\mathfrak{u}_1^3,\mathfrak{u}_2^3)$ is invariant. This group is not the split Cartan subgroup $\Gamma_s(16)$, but we can show that

$$\Gamma(16)\subset\Gamma_8\subset\Gamma_s(8)=\left\{\begin{pmatrix}a&b\\c&d\end{pmatrix}:\ b\equiv c\equiv 0 \bmod 8\right\},$$

and that the group $\Gamma_8/\Gamma(16)$ is cyclic of order 4 generated by

$$T^2U^2T^{-2}U^{-2} \equiv \begin{pmatrix} 13 & 8 \\ 8 & 5 \end{pmatrix} \bmod 16,$$

where $U = STS^{-1}$. The equality is easily verified in $\mathrm{SL}_2(\mathbb{Z}/16\mathbb{Z})$ and the word expression on the left maps to the identity under the above homomorphism to $\mathrm{GL}_3(\mathbb{Q}(\zeta_8))$, showing that the element is in the kernel of the action on $\mathscr{W}_8$. Moreover, the matrix on the right lifts to $\mathrm{SL}_2(\mathbb{Z})$. Given that the degree of $\mathscr{W}_8 \to \mathscr{W}_4 = X_s(8)$ is 4, and $X(16) \to X(8)$ is of degree 16, we obtain the following description of the kernel group Γ_8.

Proposition 7. *The Weber kernel group* Γ_8 *is the group generated by* $\Gamma(16)$ *and* $\begin{pmatrix} 13 & 8 \\ 8 & 5 \end{pmatrix}$.

In particular we note that Γ_8 contains the diagonal matrix in the center of $\mathrm{SL}_2(\mathbb{Z}/16\mathbb{Z})$:

$$\begin{pmatrix} 13 & 8 \\ 8 & 5 \end{pmatrix}^2 \equiv \begin{pmatrix} 9 & 0 \\ 0 & 9 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}/16\mathbb{Z})$$

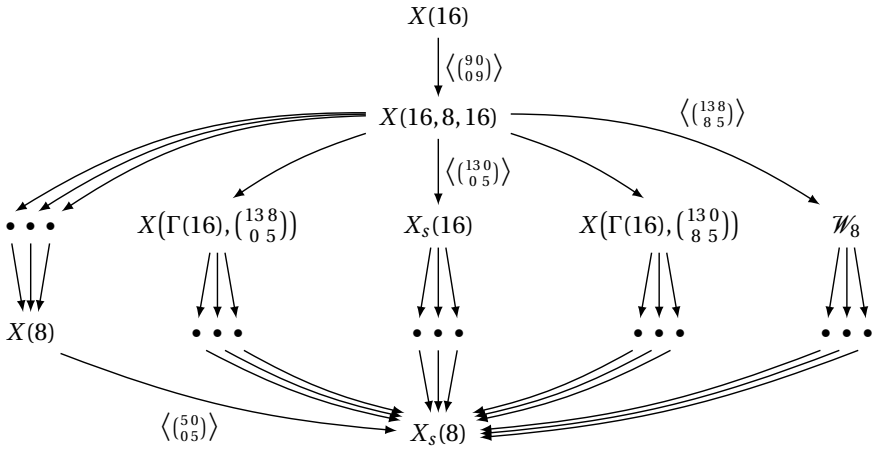
hence contains the subgroup

$$\Gamma(16,8,16) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \,:\, a \equiv d \equiv 1 \bmod 8, b \equiv c \equiv 0 \bmod 16 \right\}.$$

Given that $\Gamma_s(8)/\Gamma(16)$ is an abelian group:

$$\Gamma_s(8)/\Gamma(16) = \left\langle \begin{pmatrix} 13 & 0 \\ 0 & 5 \end{pmatrix}, \begin{pmatrix} 1 & 8 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 8 & 1 \end{pmatrix} \right\rangle \cong C_4 \times V_4,$$

so that $\Gamma_s(8)/\Gamma(16,8,16) \cong C_2 \times V_4 = C_2^3$; we have the following diagram of modular curves between $X(16)$ and $X_s(8)$, where the curves represented by dots are intermediate quotients by the subgroups of $V_4 = \langle T^8, U^8 \rangle$.



In what follows we will show that the supersingular points split completely over $\mathbb{F}_{p^2}$ on the quotients of $X(16,8,16)$ not covering $X(8)$, for every odd prime p.

Supersingular fields of definition.

Theorem 8. *For any positive integer N, the supersingular invariants on the modular curve $X_0(N)$ are defined over $\mathbb{F}_{p^2}$, and if $p \equiv \pm 1 \bmod N$, then the supersingular invariants also split over $\mathbb{F}_{p^2}$ on $X_1(N)$.*

Proof. For any elliptic curve E in the isogeny class of a curve over $\mathbb{F}_p$, the full endomorphism ring $\mathcal{O}$ is defined over $\mathbb{F}_{p^2}$. Since the action of $\mathcal{O}/N\mathcal{O} \cong \mathbb{M}_2(\mathbb{Z}/N\mathbb{Z})$ on the $E[N]$ is defined over $\mathbb{F}_{p^2}$, it follows that the Galois action on $E[N]$, which commutes with $\mathcal{O}/N\mathcal{O}$, acts through the center $(\mathbb{Z}/N\mathbb{Z})^*$ of $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$, and more precisely, Frobenius acts as $-p$ on $E[N]$. Consequently, the lines are Galois stable and every cyclic N-isogeny is defined over $\mathbb{F}_{p^2}$. In view of the action of Frobenius, if $p \equiv \pm 1 \bmod N$, the Galois action on the N-torsion of E or its twist is trivial, so the supersingular moduli are defined in $\mathbb{F}_{p^2}$. □

Remark. Equivalently, for $X_0(N)$ we can state that every supersingular j-invariant j_0 splits completely under the map $X_0(N) \to X(1)$, or that the polynomial $\Phi_N(x, j_0)$ splits completely, where $\Phi_N(x,y)$ is the classical modular polynomial. For $X_1(N)$, the splitting of the supersingular points is recognized by the factorization of the N-division polynomial ψ_N.

As a consequence, the split Cartan modular curve $X_s(N)$, defined by the congruence subgroup

$$\Gamma_s(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : b \equiv c \equiv 0 \bmod N \right\},$$

parametrizing elliptic curves with a disjoint pair of cyclic N-isogenies, also splits the supersingular moduli.

Corollary 9. *For any positive integer N, then the supersingular invariants on the split Cartan modular curve $X_s(N)$ are defined over $\mathbb{F}_{p^2}$. In particular if $p \equiv \pm 1 \bmod N$, then the supersingular invariants on the modular curve $X(N)$ are defined over $\mathbb{F}_{p^2}$.*

Proof. The first statement follows from the splitting of N-isogenies over $\mathbb{F}_{p^2}$. In addition if $p \equiv \pm 1 \bmod N$, the points of each kernel are fixed, hence a basis is defined over $\mathbb{F}_{p^2}$ (up to twist). □

Remark. For the levels N in $\{1,2,3,4,6\}$, the unit group $(\mathbb{Z}/N\mathbb{Z})^*/\{\pm 1\}$ is trivial so the supersingular points split for all p. This corresponds to the geometric equalities $X_1(N) = X_0(N)$ and $X(N) = X_s(N)$.

The Weber moduli are functions on $X(48)$ which map through $\mathcal{W}_{24}$. To show the splitting of supersingular points on $\mathcal{W}_{24}$ it suffices to prove it for $\mathcal{W}_3$ and $\mathcal{W}_8$. However, $X(6)$ covers $\mathcal{W}_3$, so the supersingular moduli on $\mathcal{W}_3$ split over $\mathbb{F}_{p^2}$ by the previous theorem. To prove that they split on $\mathcal{W}_8$ it is necessary to consider the factorization

$$\begin{array}{ccc} X(16,8,16) & \longrightarrow & X(8) \\ \Big\langle \left(\begin{smallmatrix} 13 & 8 \\ 8 & 5 \end{smallmatrix}\right) \Big\rangle \Big\downarrow & & \Big\downarrow \Big\langle \left(\begin{smallmatrix} 5 & 0 \\ 0 & 5 \end{smallmatrix}\right) \Big\rangle \\ \mathcal{W}_8 & \longrightarrow & X_s(8), \end{array}$$

where $X(16,8,16)$ is the quotient of $X(16)$ by the diagonal matrix group $\langle \pm 9 I_2 \rangle \subset \mathrm{SL}_2(\mathbb{Z}/16\mathbb{Z})/\{\pm 1\}$.

The supersingular points split in $X_s(8)$ by the previous theorem. On the other hand, for the classes $p \bmod 8$ in the coset $\{\pm 5\} \subset (\mathbb{Z}/8\mathbb{Z})^*/\{\pm 1\}$ form an obstruction to lifting supersingular points to $X(8)$ over $\mathbb{F}_{p^2}$. Clearly, since $\langle 9 I_2 \rangle \subset \Gamma(16,8,16)/\Gamma(16)$, for the primes p such that $p \bmod 16$ lie in the kernel

$$\langle -1, 9 \rangle = \{\pm 1, \pm 9\} \subset (\mathbb{Z}/16)^*/\{\pm 1\} \longrightarrow (\mathbb{Z}/8)^*/\{\pm 1\},$$

the supersingular invariants in $X(16,8,16)$ split over $\mathbb{F}_{p^2}$. It remains to show that the obstruction vanishes also on the coset $\{\pm 3, \pm 5\}$. However, this follows since the subgroup of $\Gamma_8/\Gamma(16)$ surjects on the diagonal subgroup of $\Gamma_s(8)/\Gamma(8)$:

$$\left\langle \begin{pmatrix} 13 & 8 \\ 8 & 5 \end{pmatrix} \right\rangle \longrightarrow \left\langle \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix} \right\rangle$$

under $\mathrm{SL}_2(\mathbb{Z}/16\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/8\mathbb{Z})$, corresponding to the fact that $\mathcal{W}_8$ does not factor through $X(8)$. This establishes the following theorem.

Theorem 10. *The supersingular Weber invariants on $\mathcal{W}_{24}$ are defined over $\mathbb{F}_{p^2}$.*

Remark. A point (u_0, u_1, u_2) on $\mathcal{W}_{24}$ over the j-invariant j_0 consists of a triple of common roots of the polynomial $(x^{24} - 16)^3 - j_0 x^{24}$, and the set roots is precisely $\{\zeta_{24}^i u_j : 0 \le i < 24, 0 \le j < 3\}$. The property that j_0 splits completely under $\mathcal{W}_{24} \to X(1)$ over $\mathbb{F}_{p^2}$ is equivalent to this polynomial splitting completely over $\mathbb{F}_{p^2}$.

5. Weber initialization

In what follows we denote by u a supersingular value of the Weber function, $r = u^3$, $t = -u^8$ and $s = t^3$. This gives the following relations with the j-invariant:

$$j = \frac{(u^{24} - 16)^3}{u^{24}} = \frac{(r^8 - 16)^3}{r^8} = \left(\frac{t^3 + 16}{t} \right)^3 = \frac{(s+16)^3}{s}.$$

While we will only evoke the elliptic curves associated to Weber invariants, we note that such a curve can be viewed as a fiber in the family:

$$y^2 + xy = x^3 - \frac{1}{u^{24} - 64}x$$

over u on the Weber curve $\mathscr{X}$.

The OSIDH protocol is initialized with oriented chains from an effective CM order. In the table below we give associated values for the CM j-invariants and s-invariants (where $s = -u^{24}$) for the discriminants of the first class number one CM orders, and their index 2 orders. This gives initial values with which to build the public ℓ-isogeny chains.

D	j_0	s_0	t_0	D	j_1	s_1	t_1
-3	0	-2^4	$-(\sqrt[3]{2})^4$	-12	$2^4 15^3$	-2^8	$-(\sqrt[3]{2})^8$
-4	12^3	2^3	2	-16	66^3	2^9	2^3
-7	-15^3	-1	-1	-28	255^3	-2^{12}	-2^4
-8	20^3	2^6	2^2	-32	j_1	t_1^3	$2^3(\sqrt{2}+1)$

In what follows we describe the initialization of a Weber ℓ-isogeny chain for $\ell = 2$. In view of the previous form of modular polynomials at 2 we use $t = -u^8$ and the modular polynomial

$$\Psi_2(x, y) = (x^2 - y)y + 16x$$

to construct the 2-isogeny graph on t-values.

Remark. Associated to a t-value t_i is a u-value u_i satisfying $t_i = -u_i^8$, which one can obtain by extracting three square roots (making arbitrary choices of signs, the result is determined up to an 8-th root of unity). It remains to be determined whether the sequence (u_i) of Weber values is sufficient, or whether a sequence of points

$$U_i = (u_1, u_2, u_3)_i \in \mathscr{W}_{24}(\mathbb{F}_{p^2})$$

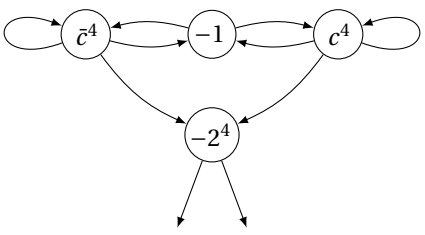
should be computed such that $u_1^8 + u_2^8 + u_3^8 = 0$ and $u_1 u_2 u_3 = \sqrt{2} \in \mathbb{F}_{p^2}$.

Discriminant -7. The fundamental discriminant -7 is an interesting starting point in that the endomorphism ring is small enough to be effective — generated by an endomorphism of degree 2 — while avoiding any pathologies associated with the extra automorphisms for $D = -3$ and $D = -4$. In fact, choosing $p = 1 \bmod 12$ we can assure that no supersingular point has extra automorphisms.

Let $t_0 = -1$ and let c be a root of $x^2 - x + 2$ in $\mathbb{F}_{p^2} = \mathbb{F}_p[c]$, with conjugate $\bar{c} = 1 - c = 2/c$. We note that c^4 and $\bar{c}^4$ are also t-values over $j = -15^3$, and since $\Psi_2(-1, c^4) = \Psi_2(-1, \bar{c}^4) = 0$, the two extensions correspond to the horizontal 2-isogenies (endomorphisms of the underlying elliptic curve). Subsequently, we find

$$\Psi_2(c^4, c^4) = \Psi_2(c^4, -2^4) = 0.$$

The former enters a cycle of degree-2 endomorphisms, while the latter induces a descending isogeny, as depicted in the 2-isogeny graph below.



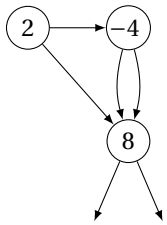
This suggests the following initialization of the 2-isogeny chain of t-values $(t_0, t_1, t_2, \ldots)$ beginning with $(-1, c^4, -2^4, \ldots)$. Successive solutions to $\Psi_2(t_i, t_{i+1}) = 0$ are necessarily descending with respect to the orientation by $\mathbb{Q}(\sqrt{-7})$. A random choice of root t_{i+1} of $\Psi_2(t_i, x)$ completes this initialization of the 2-isogeny chain to any desired depth n.

The above discussion, applicable in any characteristic, leaves open the question of the field of definition of the u-values, in particular whether $u_1 = \sqrt[8]{-c^4}$ is in $\mathbb{F}_{p^2}$. We give an affirmative response to this question in general in Theorem 10.

Discriminant -4. The t-invariants over $j = 12^3$ fall in two orbits of points, $\{2, 2\omega, 2\omega^2\}$ of multiplicity 2, and $\{-4, -4\omega, -4\omega^2\}$ of multiplicity 1. These points at the surface are linked by a 2-isogeny (the degree 2 endomorphism by $1+i$), since $\Psi_2(2,-4) = 0$ and to 2-depth 1, to $t = 8$:

$$\Psi_2(2,8) = \Psi_2(-4,8) = 0.$$

Given that $\Psi_2(\omega x, \omega^2 y) = \omega\Psi_2(x,y)$, the choice of representative in the orbit gives rise to one of three distinct components of the 2-isogeny graph, with the component of 2 depicted below.



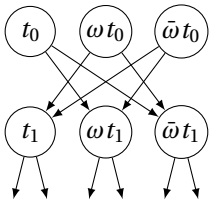
This suggests the initialization $(t_0, t_1, t_2, \ldots) = (2, 8, 8c, \ldots)$ of the 2-isogeny chain, where c is a root of $x^2 - 8x - 2$, extended by random selection of a root t_{i+1} of $\Psi_2(t_i, x)$.

Remark. The full 2-isogeny graph has ascending edges from the depth one points $t_1' = -4 + 3\sqrt{2}$ and its conjugate $t_1'' = -4 - 3\sqrt{2}$ to $t_0 = 2$, as well as a descending isogeny from each of t_1' and t_1'' to depth 2. In general, if an isogeny is descending its only extension to a 2-isogeny chain is descending (since this is a property of the underlying j-invariant chain).

Discriminant -3. In what follows we will give preference to the discriminant -3, which we treat next. Let $t_0 = -(\sqrt[3]{2})^4 = -2\sqrt[3]{2}$. Then $\{t_0, t_0\omega, t_0\omega^2\}$ is the set of t-values over $j = 0$, each of multiplicity 3, where as above $\omega^2 + \omega + 1 = 0$. Setting $t_1 = -t_0^2$, we verify that

$$\Psi_2(t_0, t_1\omega) = \Psi_2(t_0, t_1\omega^2) = 0,$$

Since 2 is inert, every path from t_0 is descending, so we may initialize the 2-isogeny chain with $(t_0, t_1\omega)$. The graph of descending isogenies from the surface vortex appears as follows.



As above there are additional t-invariants at each depth greater than 0 which admit ascending and descending isogenies. At depth 1, we have $\{t_1', t_1'\omega, t_1'\omega^2\}$ and $\{t_1'', t_1''\omega, t_1''\omega^2\}$, which ascending to the surface points $\{t_0, t_0\omega, t_0\omega^2\}$ and descending (in bijection) with the points at depth 2. Any descending isogenies must rejoin this graph of descending isogenies from the surface.

Lifting 2-isogeny chains from t-invariants to Weber points. The modular curve $\mathcal{W}_3$ has affine model

$$x_0 + x_1 + x_2 = 0,\ x_0x_1x_2 = 16,$$

from which we derive $x_0x_1^2 + x_0^2x_1 + 16 = 0$, and projects to the t-value $t = -x_0$. Let (y_0, y_1, y_2) be an affine point 2-isogenous to (x_0, x_1, x_2). The two possible extensions, determined by the modular polynomial $\Psi_2(x,y) = (x^2 - y)y + 16x$ on $(-x_0, -y_0)$ are given by the solutions

$$y_0 = x_0x_1 \text{ or } y_0 = x_0x_2.$$

We may assume that $y_0 = x_0x_1$, determined by the prior choice of point (x_0, x_1, x_2), as opposed to (x_0, x_2, x_1). The neighboring 2-isogenous point (y_0, y_1, y_2) is determined by a choice of solution to the quadratic equation:

$$y_1^2 + y_0y_1 + x_2.$$

noting that the other root is $y_2 = -y_0 - y_1$. The third coordinate is redundant, given the dependence on (y_0, y_1). The lifts to points of $\mathscr{W}_{24}$ are choices of points (u_0, u_1, u_2) and (v_0, v_1, v_2) such that

$$(u_0^8, u_1^8, u_2^8) = (x_0, x_1, x_2),\ u_0u_1u_2 = \sqrt{2},$$

and

$$(v_0^8, v_1^8, v_2^8) = (y_0, y_1, y_2),\ v_0v_1v_2 = \sqrt{2},$$

for a fixed element $\sqrt{2}$ of $\mathbb{F}_{p^2}$. The values of (u_0, u_1) and (v_0, v_1) can be arbitrarily chosen, up to a power of ζ_8, after which the product relation determines u_2 and v_2.

6. Modular OSIDH Protocol

The OSIDH protocol [5] made exclusive use of the class group action at split primes in $\mathcal{O}$. In this work we extend the protocol to include descent in the eddies at non-split primes (inert or ramified) or at large primes which are not cost-effective for use for longer isogeny walks.

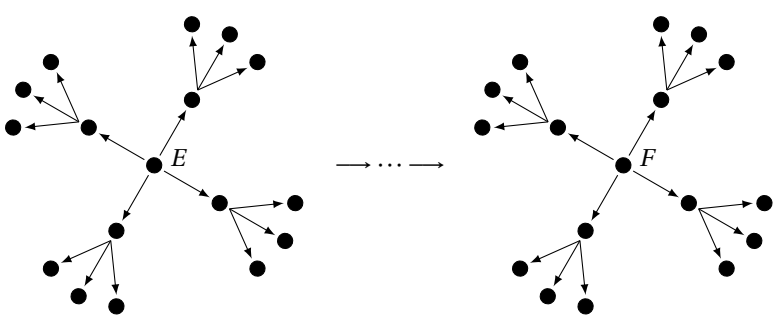
Isogeny computation. The standard practice in supersingular elliptic curve-base cryptosystems, such as SIDH [7], CSIDH [4] and SQISign [9], is to use torsion points of smooth order, whose existence is assured by the factorization of $p^2 - 1$. This practice is justified by the the analysis of De Feo, Kieffer and Smith [8], who describe algorithms for isogeny computation with modular curves and torsion points. In particular, for the modular curve approach they describe explicit algorithms, "Elkies first step" and "Elkies walk" and analyze their complexity. In conclusion, the alternative "Vélu walk" using torsion points is found more efficient.

By means of precomputed modular isogeny walks (chains and clouds), in this protocol we replace the Elkies steps with pushforward algorithms which require only gcd's of modular polynomials to "compose" isogenies, at the cost of memory. These algorithms are described below. We use Weber modular polynomials, whose sparseness and small coefficient size renders the construction of their specializations from bivariate to univariate polynomials more efficient.

The use of a torsion point of order q^r or a basis of the q^r torsion would provide an efficient method of compressing the precomputed q-isogeny chains and q-isogeny clouds, respectively. This would be of particular interest for the larger primes, and the use of the modular approach for small primes would vastly simplify the sieving step for finding suitable primes.

Pushforward of chains. The pushforward of a q-isogeny chain by an ℓ-isogeny, both from a curve E, was described in [5] in terms of construction of isogeny ladders. This process is iterated to push forward the q-isogeny chain along a ℓ-isogeny chain $E = F_0 \to \cdots \to F_n = F$ to create the image q-isogeny chain from F.

Pushforward of clouds. The pushforward of a q-isogeny cloud by an ℓ-isogeny chain involves pushing forward each of the $q+1$ neighboring q-isogenies along the isogeny chain, iterating to depth r. This is represented graphically in the following image (for $q = 3$).



The ladder construction requires taking gcd's with polynomials of degree $\ell+1$ and degree $q+1$. As the matching of the $q+1$ neighbors progresses, the degree of the latter polynomial decreases, but the total cost remains proportional to the size of the cloud.

The protocol. We describe the new modular OSIDH key exchange protocol using horizontal isogenies at split primes and descents at non-split primes. The protocol is broken down into three phases of parameter setup, key generation, and finally key exchange. Only the final stage requires computation in real time, as the first two steps concern the initialization of the protocol and systems of public-private keys.

The protocol requires setting up a set $\mathscr{P}_s$ of split primes in $\mathcal{O}_K$ for exchange of horizontal isogeny chains, shared by parties A and B, and two sets $\mathscr{P}_A$ and $\mathscr{P}_B$ of non-split prime powers for the exchange of isogeny clouds. The depth n at ℓ of the initial descent, and the conductors M_A and M_B, products of the prime powers $\mathscr{P}_A$ and $\mathscr{P}_B$, respectively, determine the conductors of the exchanged oriented orders $\mathcal{O}_n(M_A)$ and $\mathcal{O}_n(M_B)$. In the following section we give particular parameters choices for the sets $\mathscr{P}_A$ and $\mathscr{P}_B$ in order to assure that $M_A \approx M_B$.

Phase 0 - Parameter setup. The initialization phase consists of fixing a discriminant D_K and prime ℓ, and constructing a common public isogeny chain (of moduli) from a given supersingular curve $E_0/\mathbb{F}_{p^2}$ with CM by the order of discriminant D_K. The terminus E_n of this isogeny chain is denoted E.

$$E_0 \longrightarrow E_1 \longrightarrow \ldots \longrightarrow E_n = E$$

The protocol is based on a prior specification of disjoint prime sets $\mathscr{P}_s$ of split primes, and $\mathscr{P}_A$ and $\mathscr{P}_B$ which partition the non-split (inert or ramified) and larger split primes. In the case of split or ramified primes, only descending isogenies are used, which have the effect of increasing the class group size.

We will first look at the set $\mathscr{P}_s$: at this stage we need to precompute the initial directions in the class group (this is the analog of Elkies first step) at split primes - this step was also present in the original OSIDH protocol [5]. This means that, to each prime and exponent pair (q,r) in $\mathscr{P}_s$, q-isogeny chains of length $2r$ are constructed from E_0 for each prime $\mathfrak{q}$ over q, and pushed forward to E. One direction is declared positive and the other negative, so that the concatenated chains are in bijection with $[-2r,2r]$.

Secondly, we let the two parties precompute labeled clouds at non-split primes including ℓ. For each prime and exponent pair (q,r) in $\mathscr{P}_A$ and $\mathscr{P}_B$ the q-isogeny eddy of depth r is constructed around E_0 and pushed forward to E.

This initialization data is made public, after which each party in an exchange can play the role of A or B, initializing one or more public and private key data sets as follows.

Phase I - Key generation. Following the setup procedure, A and B will compute their secret key.

On one side, A begins with $F = E$. As in the original instantiation of OSIDH, for each prime q_i in $\mathscr{P}_S$, A chooses a random $s_i \in [-r_i, r_i]$. For $j = 1$ up to t, she constructs the q_j-isogeny walk of length s_j from the current F, relabeling the remaining curves in the interval $[-r_j + s_j, r_j + s_j] \subset [-2r_j, 2r_j]$ to $[-r_j, r_j]$, and pushing forward the curves in the intervals $[-r_i, r_i]$ for each $i < j$ and the intervals $[-r_i + s_i, r_i + s_i]$ for each $i > j$. She also pushes forward the q-clouds at each prime q in $\mathscr{P}_A$ and $\mathscr{P}_B$.

A should then compute the non-split key: for each prime and exponent pair (q_j, r_j) in $\mathscr{P}_A$, A chooses a random walk of length r_j in the cloud to a new curve F and pushes forward the remaining unused q-clouds for q in $\mathscr{P}_A$ as well as all q in $\mathscr{P}_B$ to F.

The data F and q-isogeny chains at primes q in $\mathscr{P}_s$ and q-clouds at primes q in $\mathscr{P}_B$ constitute A's public key. The indices s_j, for $1 \le j \le t$ and the isogeny walks in $\mathscr{P}_A$ form her private key.

In parallel, B constructs an equivalent public key on a curve G with data for the primes in $\mathscr{P}_s$ and in $\mathscr{P}_A$, saving his private key. The public key data for A and B can be certified by a certification authority.

Phase II - Key establishment.

A obtains the public key data from B (or a certification authority) and reconstructs her isogeny walk from G using B's data for $\mathscr{P}_s$ and $\mathscr{P}_A$.

B obtains the public key data from A (or a certification authority) and reconstructs his isogeny walk from F using A's data for $\mathscr{P}_s$ and $\mathscr{P}_B$.

The curve H resulting from A's and B's random walks serves as secret key.

7. Security considerations

Security analysis. The parameter choices described for OSIDH [5] were presented with two competing considerations: the K-orientation must remain hidden and a large logarithmic proportion λ of total supersingular curves should be reached by the class group:

$$\lambda = \log_p\left(|\mathscr{C}\ell(\mathscr{O}_n)|\right),$$

so that $\mathscr{C}\ell(\mathscr{O}_n) \to \mathrm{SS}(p)$ has large image. Secondly, the authors considered the image of the map:

$$I = \prod_{j=1}^{t} [-r_j, r_j] \longrightarrow \mathscr{C}\ell(\mathscr{O}_n),$$

noting that it should be computationally difficult to find a nontrivial element of the kernel

$$L_n = \ker\left(\mathbb{Z}^t \to \mathscr{C}\ell(\mathscr{O}_n)\right).$$

in the intersection with I.

A first necessary bound for the injectivity of $I \to \mathscr{C}\ell(\mathscr{O})$ is that the cardinalities are bounded. For constant $r_j = r$, this gives

$$\log_\ell(|I|) = t\log_\ell(2r+1) < \log_\ell\left(|\mathscr{C}\ell(\mathscr{O}_n)|\right) \approx n.$$

In taking parameters $n = 256$, $t = 74$, and $r = 5$, used for CSIDH-512 [4] this bound is critically attained.

Unfortunately, to provably preclude collisions, one needs a stronger bound on the norms of elements. A collision in the map $I \to \mathscr{C}\ell(\mathscr{O}_n)$ implies the relation

$$\prod_{j=1}^{t} \mathfrak{q}_j^{c_j} \sim \prod_{j=1}^{t} \mathfrak{q}_j^{d_j} \text{ hence } (\alpha) = \prod_{j=1}^{t} \mathfrak{q}_j^{s_j},$$

where $c_j, d_j \in [-r_j, r_j]$ with $s_j = c_j - d_j$, implying $|s_j| \le 2r_j$.

$$\log_\ell(N(\alpha)) = \sum_{j=1}^{t} s_j \log_\ell(q_j) \ge \log_\ell\left(\ell^{2n}\frac{|D_K|}{4}\right) = 2n + \log_\ell(|D_K|/4).$$

To avoid such collisions, one needs the stronger norm bound:

$$\sum_{j=1}^{t} r_j \log_\ell(q_j) < n,$$

which clearly fails for $(t = 74, r_j = 5, n = 256)$. In fact these parameters permitted the full class group computation for the CSIDH-512, resulting in the CSI-FiSh protocol [2]. In the setting of CSIDH-512, the class group and its order was unknown and $O(t)$ cycles are needed to determine the class group. In the setting of OSIDH, the class group is known, and only one short vector is needed to determine the K-orientation.

Onuki [11, §6.3] recalls the design objectives from OSIDH [5, §5.1] that the knowledge of the K-orientation breaks the cryptosystem, giving an ascending walk up the ℓ-isogeny chain, and proposed an exponential meet-in-the-middle attack. Dartois and De Feo [6] carry out the class group attack to determine a cycle hence breaking this parameter set. Combining a class group analysis in $\mathscr{C}\ell(\mathscr{O}_{n-r})$ and meet-in-the-middle attack to depth r makes this approach even more significant.

Countermeasures. These prior attacks and analyses motivated the authors to revisit the security parameters and include the non-split primes in the protocol. The data for these primes augments the class group, but no effective data is transmitted for a third party to compute cycle relations.

Let $\mathscr{O}$ be an order of the form $\mathscr{O} = \mathscr{O}_n(M) = \mathscr{O}_K(\ell^n M) = \mathbb{Z} + M\mathscr{O}_n$. The size of the orbit of $\mathscr{C}\ell(\mathscr{O})$ is controlled by the chain length n and conductor M, and the number of curves attained by the private walks is further limited by the prime power data, up to exponent bounds, which we allow ourselves to transmit.

We note that $\mathscr{C}\ell(\mathscr{O}) \simeq \mathrm{SS}^{pr}_{\mathscr{O}}(\rho)$ and define I the exponents space $I_1 \times \ldots \times I_t \subseteq \mathbb{Z}^t$ where $I_j = [-r_j, r_j]$. The security of OSIDH depends on the injectivity and surjectivity of the following maps

$$I = \prod_{i=1}^{t} [-r_i, r_i] \longrightarrow \mathrm{SS}^{pr}_{\mathscr{O}}(\rho) \longrightarrow \mathrm{SS}(p)$$

In order to analyze the security of the protocol, we state the following bounds and study their related impact. We first consider the properties of the map $\mathrm{SS}^{pr}_{\mathcal{O}}(\rho) \to \mathrm{SS}(p)$.

Supersingular covering bound. We deal with the problem of covering a reasonable number of elliptic curves in $\mathrm{SS}(p)$. We say that the map $\mathcal{C}\ell(\mathcal{O}) \simeq \mathrm{SS}^{pr}_{\mathcal{O}}(\rho) \longrightarrow \mathrm{SS}(p)$ is λ-*surjective* if $\#\mathcal{C}\ell(\mathcal{O}) \geq p^{\lambda}$ where λ is the *logarithmic covering radius*. We get

$$\lambda \log_\ell(p) \leq n + \log_\ell(M) + \log_\ell(h(\mathcal{O}_K)) \tag{SCB}$$

Remark. In SIDH the intermediate cloud covers $O(\sqrt{p})$ supersingular elliptic curves, so the logarithmic covering radius is $1/2$. By varying the conductor $M\ell^n$, we determine the proportion of curves covered.

Supersingular injectivity bound. How can one insure the injectivity of the map $\mathrm{SS}^{pr}_{\mathcal{O}}(\rho) \to \mathrm{SS}(p)$? We set

$$n + \log_\ell(M) + \frac{1}{2}\log_\ell(|D_K|) \leq \frac{1}{2}\log_\ell(p) \tag{SIB}$$

Lemma 11. *If (SIB) holds, then the map* $\mathrm{SS}^{pr}_{\mathcal{O}}(\rho) \to \mathrm{SS}(p)$ *is injective.*

Proof. The failure of the injectivity implies the existence of two embeddings of $\mathcal{O}$ in the endomorphism ring $\mathrm{End}(E)$. This implies the existence of T such that $(D^2 - T^2)/4 = mp$, see [5, Lemma 12]. Hence $2|D| \geq |D| + T \geq p$ and, therefore, if injectivity fails we must have $p \leq |D|$. □

Remark. Comparing bounds (SCB) and (SIB), one sees that the transition from injectivity to non-injectivity happens around the logarithmic covering radius $\lambda = \frac{1}{2}$ while surjectivity is only possible for $\lambda \geq 1$. Further, (SCB) does not guarantee surjectivity but only provides an upper bound on the number of classes covered. This incompatibility however, is not problematic as injectivity is not really necessary for security.

We will now focus on the first map $I \to \mathcal{C}\ell(\mathcal{O})$.

Minkowski norm bound. The set of elements obtained by random walks should contain no cycle; thus,

$$\sum_{i=1}^{t} r_i \log_\ell(q_i) \leq n + \log_\ell(M) + \frac{1}{2}\log_\ell(|D_K|/4) \tag{MNB}$$

Lemma 12. *If (MNB) holds, then the map* $I \to \mathrm{SS}^{pr}_{\mathcal{O}}(\rho)$ *is injective.*

Proof. As before, a collision in the map $I \to \mathcal{C}\ell(\mathcal{O})$ implies the relation

$$\prod_{j=1}^{t} \mathfrak{q}_j^{c_j} \sim \prod_{j=1}^{t} \mathfrak{q}_j^{d_j} \text{ hence } (\alpha) = \prod_{j=1}^{t} \mathfrak{q}_j^{s_j}$$

hence

$$\log_\ell(Nr(\alpha)) = \sum_{j=1}^{t} |s_j| \log_\ell(q_j) \geq \log_\ell(|D|/4) = 2n + 2\log_\ell(M) + \log_\ell(|D_K|/4)$$ □

Class group covering bound. We now consider λ-surjectivity of the map $I \to \mathrm{SS}^{pr}_{\mathcal{O}}(\rho)$. In order to have a uniform element of $\mathcal{C}\ell(\mathcal{O})$, it is desirable to be able to reach all elements of $\mathcal{C}\ell(\mathcal{O})$. A necessary condition for surjectivity is that the cardinality of walks ending points is at least the class number of $\mathcal{O}$. Adding a parameter λ, we say that $I \to \mathcal{C}\ell(\mathcal{O})$ is λ-surjective if $\#I \geq h(\mathcal{O})^{\lambda}$. Taking the logarithm, this gives

$$\sum_{i=1}^{t} \log_\ell(2r_i + 1) \geq \lambda\left(n + \log_\ell(M) + \log_\ell(h(\mathcal{O}_K))\right) \tag{CGCB}$$

Remark. In adaptations of the OSIDH protocol, see Section 6, the class group surjectivity for $\mathcal{C}\ell(\mathcal{O}_n)$ is achieved with the weaker bound

$$\sum_{i=1}^{t} \log_\ell(2r_i + 1) \geq n + \log_\ell(h(\mathcal{O}_K))$$

and surjectivity for $\mathcal{C}\ell(\mathcal{O})$ is achieved by a random walk at conductor M, giving a random element of the kernel $\mathcal{C}\ell(\mathcal{O}) \to \mathcal{C}\ell(\mathcal{O}_n)$.

8. Parameter selection

In this section we make specific proposals for the sets $\mathscr{P}_s$, $\mathscr{P}_A$ and $\mathscr{P}_B$, with respect to the maximal order $\mathcal{O}_K$ of discriminant $D_K = -3$ and prime $\ell = 2$, and analyze the security consequences. We defer the question of size of p, which does not impact the size of the class group or attacks in that class group, noting only that n should be of the same order of magnitude as $\log_\ell(p)$ to ensure that $|\mathscr{C}\ell(\mathcal{O}_n)| \approx p$, hence that the initial curve $E = E_n$ can be a generic curve in the supersingular graph. The use of a 256 bit prime was specified for the original OSIDH protocol [5].

In order to have a uniform contribution for each prime, we choose the maximum exponent r of a split prime q in $\mathscr{P}_s$ such that q^r is bounded by a bit bound B_s per prime power, i.e. we set $r = \lfloor B_s/\log_2(q) \rfloor$. For primes $\mathfrak{q}_j$ over q_j, we recall that a collision in the class group of $\mathcal{O}_n(M)$,

$$\prod_{j=1}^{t} \mathfrak{q}_j^{a_j} \sim \prod_{j=1}^{t} \mathfrak{q}_j^{b_j},$$

with $|a_j|, |b_j| \leq r_j$, gives rise to an endomorphism α, with exponents $s_j = |a_j - b_j|$ bounded by $2r_j$, up to replacing $\mathfrak{q}_j$ with $\bar{\mathfrak{q}}_j$, and with norm satisfying

$$\mathrm{N}(\alpha) = \prod_{j=1}^{t} q_j^{s_j} > \frac{|D|}{4} = M^2 \ell^{2n} \frac{|D_K|}{4}.$$

Hence, for $\ell = 2$ and $|D_K|/4$ negligible, if we choose t, B_s, M and n such that

$$\sum_{j=1}^{t} r_j \log_2(q_j) \leq tB_s \leq \log_2(M) + n$$

holds, no such endomorphism exists in $\mathcal{O}_n(M)$, and the lattice-based class group attack of Dartois-DeFeo [6] does not apply.

Security considerations. As a consequence of the analysis of security against lattice-based class group attack, the number t of split primes, and the product tB_s in particular, should be strictly controlled, as well as the size of the conductors M (equal to M_A or M_B). In the example below we begin with $t = 10$, and a bit bound $B_s = 32$ (or 24), giving $tB_s = 320$ bits (or 240 bits). We also use eddies at larger split primes which contribute equally a term $\log_2(q)$ to both sides of the inequality, but allow for an increase in the number of vertices reachable in the supersingular isogeny graph.

Parameter sets. In what follows we set $D_K = -3$ and $\ell = 2$, and propose parameter sets $\mathscr{P}_s$, $\mathscr{P}_A$ and $\mathscr{P}_B$ out of the primes up to 163 with an analysis of their security constraints. We subsequently consider the security implications of modifying B_s and moving split primes into the sets $\mathscr{P}_A$ or $\mathscr{P}_B$ to use for descents (in the eddy).

Ten split primes. For a parameter set, we consider $\mathscr{P}_s$ consisting of the first 10 split primes q in the maximal order $\mathcal{O}_K$ of discriminant -3, with a bound $r = \lfloor B_s/\log_2(p) \rfloor$, with $B_s = 32$, such that we take a walk whose length is in the interval $[-r, r]$ — a positive value is with respect to a prime $\mathfrak{q}$ over q and a negative value is with respect to $\bar{\mathfrak{q}}$.

	q:	7	13	19	31	37	43	61	67	73	79
$\mathscr{P}_s$:	r:	11	8	7	6	6	6	5	5	5	5
	#:	23	17	15	13	13	13	11	11	11	11

The third line is the number $2r + 1$ of curves reached by a uniformly random length walk. This gives a logarithmic contribution of

$$\sum_{j=1}^{10} \log_2(2r_j + 1) = 37.4569...$$

to the entropy of the random walk. On the other hand, the logarithmic norm, which we must bound is:

$$\sum_{j=1}^{10} r_j \log_2(q_j) = 306.2115... (< 320 = 32 \cdot 10).$$

We partition the remaining primes up to 163 into sets $\mathscr{P}_A$ and $\mathscr{P}_B$, with a radius for the cloud (or eddy), as follows:

	q:	2	11	17	41	47	59	83	101	103	109	131	149	151	157
$\mathscr{P}_A$:	r:	7	2	1	1	1	1	1	1	1	1	1	1	1	1
	#:	128	132	18	42	48	60	84	102	102	108	132	150	150	156

including the split primes 103, 109, 151 and 157, and

	q:	3	5	23	29	53	71	89	97	107	113	127	137	139	163
$\mathscr{P}_B$:	r:	4	3	1	1	1	1	1	1	1	1	1	1	1	1
	#:	81	150	24	30	54	72	90	96	108	114	126	138	138	162

including the split primes 97, 127, 139 and 163. The specification of the descending isogenies in $\mathscr{P}_A$ (by B) at the split primes leaks the horizontal directions for these primes, giving an additional contribution of

$$\log_2(103)+\log_2(109)+\log_2(151)+\log_2(157) = 27.9877..$$

bits to the logarithmic norm, and symmetrically, the descending isogenies given in $\mathscr{P}_B$ (by A) leak an additional

$$\log_2(97)+\log_2(127)+\log_2(139)+\log_2(163) = 28.0562...$$

bits to the logarithmic norm. In both cases the maximal logarithmic norm which can be attained is less than 335 bits. These prime sets each contribute a $\log_2(M)$ of 90 bits, such that n must be at least 244 to defeat the lattice-based class group attack.

The third line of the tables for $\mathscr{P}_A$ and $\mathscr{P}_B$ are the numbers $m(q,r)$ of curves at distance r:

$$m(q,r) = \left(q - \left(\frac{D_K}{q}\right)\right) q^{r-1},$$

and for each set we have $\sum \log_2(m(q,r)) \approx \log_2(M) \approx 90$ bits contributed to the number of curves reached by the random isogeny walk. Together with the contribution of split primes, this gives approximately 128 bits (out of $\log_2(p)$ bits in $|\mathrm{SS}(p)|$).

Remark. While the lattice-based class group attack is rendered ineffective for $n = 256$ by the logarithmic norm bound (including split primes in $\mathscr{P}_A$ or $\mathscr{P}_B$ with $r = 1$),

$$\sum_{j=1}^{t} r_j \log_2(q_j) \approx 344 \leq \log_2(M) + n \approx 90 + 256 = 356,$$

with respect to the class group $\mathscr{C}\ell(\mathscr{O}_n(M))$, the margin of security of 12 bits is insufficient. It suffices to construct a putative cycle, with respect to the class group $\mathscr{C}\ell(\mathscr{O}_m(M))$ for $m + 12 < n$, and construct the corresponding isogeny in the class group $\mathscr{C}\ell(\mathscr{O}_n(M))$, and carry out an exhaustive search up to radius 12 in the ℓ-isogeny graph to find a collision. This suggests a more moderate bound B_s like 24, which gives the split prime table:

	q:	7	13	19	31	37	43	61	67	73	79
$\mathscr{P}_s$:	r:	8	6	5	4	4	5	4	3	3	3
	#:	17	13	11	9	9	9	9	7	7	7

which results in a more secure 120-bit margin of security:

$$\sum_{j=1}^{t} r_j \log_2(q_j) \approx 236 \leq \log_2(M) + n \approx 356.$$

On the other hand the random walk at split primes contributes only 32 bits to the number of curves (modular invariants) attainable in the graph, for a total of 122 bits.

Two split primes. For comparison we consider the use of just two shared split primes with a bit bound $B_s = 64$ which allows for longer walks at the split primes.

	q:	7	13
$\mathscr{P}_s$:	m:	22	17
	#:	45	35

The longer walks, with $B_s = 64$ contribute 125 bits to the maximal norm of a product, while giving $35 \cdot 45 = 1575$ possible isogenies.

The remaining 36 primes up to 163 can be partitioned into sets $\mathscr{P}_A$ and $\mathscr{P}_B$ as follows:

	q:	2	11	17	31	37	41	47	59	67	73	83	101	103	109	131	149	151	157
$\mathscr{P}_A$:	r:	7	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	#:	128	132	18	30	36	42	48	60	66	72	84	102	102	108	132	150	150	156

including the split primes $31, 37, 67, 73, 103, 109, 151, 157$, and

	q:	3	5	19	23	29	43	53	61	71	79	89	97	107	113	127	137	139	163
$\mathscr{P}_B$:	r:	4	3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	#:	81	150	18	24	30	42	54	60	72	78	90	96	108	114	126	138	138	162

including the split primes $19, 43, 61, 79, 97, 127, 139, 163$. These split primes each contribute 50 bits to the maximal norm of a product ideal, and each set $\mathscr{P}_A$ and $\mathscr{P}_B$ contributes 112 bits to the conductor M ($= M_A$ or M_B). As a result we have a bound of $125 + 50 = 175$ bits on logarithmic norms, compared to $\log_2(M) = n = 112 + 256 = 368$ coming from the discriminant bounds.

Remark. We emphasize the disparity between 1575 curves reachable by 7- and 13-isogenies, a contribution of just 10 bits, compared to their maximal degree (norm of a product ideal) of 125 bits. These products will be complemented with eight other split primes in $\mathscr{P}_A$ or $\mathscr{P}_B$ (see below), but it is unlikely the set of $1575 \cdot 3^8$ possible products (a 23-bit number) will contain a principal ideal in a class group of 360 bits. This improbability of a collision with the principal class is ignored in favor of the stronger absolute bound on norms of endomorphisms, which gives a provable zone of exclusion.

Conclusion. By the same logic one could consider reducing to zero split primes. However, the number of curves which can be reached will be fewer, since the 45 or 35 isogeny neighbors which can be used by both A and B, which will be replaced by an eddy at 7 or 13 (in $\mathscr{P}_A$ or $\mathscr{P}_B$) of fewer than 1575 neighbors, since to depth 2 this contribute 42 and 156 neighbors respectively. Extending to depth 3 would explode the size of the cloud beyond reasonable limits (passing at most 162 neighbors per prime up to 163).

9. Conclusion

In view of the effectiveness of the class group attack of Dartois and De Feo [6], in addition to developing the approach to traversing isogeny graphs through Weber moduli, we give a more precise analysis and revision of the suitable parameters for the modular OSIDH. In particular, this proposal is based on the use of eddies at primes outside the shared list of small split primes $\mathscr{P}_s$. This also highlights a limitation of the shared primes, as observed by Dartois and De Feo, namely that, however corrected to avoid cycles in the class group $\mathscr{C}\ell(\mathscr{O}_n)$, they cannot cover all ideal classes, and they do not give rise to an effective group action in the sense of Alamati et al [1].

While the action by isogenies gives rise to commuting operators, acting through $\mathscr{C}\ell(\mathscr{O}_n)$, this action is restricted to exponents in a box $I \subset \mathbb{Z}^t$ of the form

$$I = \prod_{j=1}^{r} [-r_j, r_j]$$

such that a product ideal $\mathfrak{a}$ in the image satisfies $N(\mathfrak{a}) \leq \ell^n \sqrt{|D_K|/4}$. This set contributes to the entropy of the random walks of A and B, but does not determine a cycle. The main contribution to the entropy, however, comes from the descents of A and B in the eddies dividing M_A and M_B, which are acted on faithfully and transitively by the groups of the form $U(\mathscr{O}_n, q^r)$. By the Chinese remainder theorem, each group $U(\mathscr{O}_n, M)$ is a product of the groups $U(\mathscr{O}, q^r)$ for each prime power divisor q^r of M, acting on the eddies at q. Consequently, the data exchanged by A and B are images of random elements in the sets

$$I \times U(\mathscr{O}_n, M_A) \text{ and } I \times U(\mathscr{O}_n, M_B),$$

and the shared secret is then in $(I + I) \times U(\mathscr{O}_n, M_A M_B)$, where

$$I + I = \prod_{j=1}^{t} [-2r_j, 2r_j],$$

with a non-uniform distribution.[2] While $(I+I)\times U(\mathcal{O}_n, M_A M_B)$ is not a principal homogeneous space, it does admit a structure of homogeneous space over the group $U(\mathcal{O}_n, M_A M_B)$, with orbits over each element of $I+I$.

The framework described here, traversing the isogeny graph and composing isogenies via their moduli, leaves open various questions for future research, particularly in terms of efficiency and data compression. In particular, can the group action on moduli be made effective, given the Galois action of $\mathrm{SL}_2(\mathbb{Z}/q^r\mathbb{Z})$ on $X(q^r)$ over $X(1)$ and how to efficiently integrate use of torsion points with moduli. This would potentially compress a cloud at q from $O(q^r)$ points to a basis (P,Q) of q^r-torsion, a pair of points on $X_1(q^r)$, or a moduli point on $X(q^r)$. For small q, in which curves $X_0(\ell q)$, $X_0(\ell^2 q)$ or $X_0(\ell q^2)$ have genus 0, rational parametrizations can be exploited for algorithmic efficiency.

REFERENCES

[1] N. Alamati, L. De Feo, H. Montgomery, S. Patranabis. Cryptographic group actions and applications, In *Advances in Cryptology — ASIACRYPT 2020 (Daejeon, South Korea)*, 7–11, 2020.

[2] W. Beullens, T. Kleinjung, F. Vercauteren. Efficient isogeny based signatures through class group computations, *Advances in Cryptology — ASIACRYPT 2019, LNCS*, vol. **11921**, Springer, 227–247, 2019.

[3] A. Bostan, F. Morain, B. Salvy and É. Schost. Fast algorithms for computing isogenies between elliptic curves. *Mathematics of Computation*, **77**(263), 1755–1778, 2008.

[4] W. Castryck, T. Lange, C. Martindale, L. Panny, and J. Renes. CSIDH: an efficient post-quantum commutative group action, In *Advances in Cryptology - ASIACRYPT 2018, LNCS*, vol. **11274**, Springer, 395–427, 2018.

[5] L. Colò and D. Kohel. Orienting supersingular isogeny graphs, *Journal of Mathematical Cryptology*, **14**, 414–437, 2020.

[6] P. Dartois and L. De Feo. On the security of OSIDH, 2021. `https://ia.cr/2021/1681`.

[7] D. Jao and L. De Feo, Towards quantum-resistant cryptosystems from supersingular elliptic curve isogenies, In Bo-Yin Yang (ed.), *Post-Quantum Cryptography — PQCrypto 2011, LNCS*, **7071**, Springer, 19–34, 2011.

[8] L. De Feo, J. Kieffer and B. Smith. Towards practical key exchange from ordinary isogeny graphs, *Cryptology ePrint Archive, Report 2018/485*, 2018. `https://ia.cr/2018/485`.

[9] L. De Feo, A. Leroux, C. Petit, B. Wesolowski, SQISign: Compact Post-quantum Signatures from Quaternions and Isogenies, *Advances in Cryptology — Asiacrypt 2020, LNCS*, vol. **12491**, Springer, 64–93, 2020.

[10] A. Gee. Class invariants by Shimura's reciprocity law, *Journal de Théorie des Nombres de Bordeaux*, **11**, no. 1, 45–72, 1999.

[11] H. Onuki. On oriented supersingular elliptic curves, *Finite Fields and Their Applications*, **69**, 2021.

[12] C. Petit. Faster algorithms for isogeny problems using torsion point images, *In: T. Takagi, T. Peyrin (eds.) ASIACRYPT 2017, Part II. LNCS*, vol. **10625**, Springer, 330–353, 2017.

[13] J. Vélu. Isogénies entre courbes elliptiques, *Comptes-rendus de l'Académie des Sciences* **273**, 238–241, 1971. `https://gallica.bnf.fr`

[14] B. Wesolowski. The supersingular isogeny path and endomorphism ring problems are equivalent, *Cryptology ePrint Archive, Report 2021/919*, 2021. `https://ia.cr/2021/919`

[2] Precisely one can show that the probability of $(a_1,\dots,a_t)$ is $\prod_j P(a_j)$ where $P(a_j) = (2r_j+1-|a_j|)/(2r_j+1)^2$.