More Efficient Isogeny Proofs of Knowledge via Canonical Modular Polynomials

Thomas den Hollander* Sören Kleine† Marzio Mula‡ Daniel Slamanig§
Sebastian A. Spindler¶

Research Institute CODE, Universität der Bundeswehr München

Abstract

Proving knowledge of a secret isogeny has recently been proposed as a means to generate supersingular elliptic curves of unknown endomorphism ring, but is equally important for cryptographic protocol design as well as for real world deployments. Recently, Cong, Lai and Levin (ACNS'23) have investigated the use of general-purpose (non-interactive) zero-knowledge proof systems for proving the knowledge of an isogeny of degree 2^k between supersingular elliptic curves. In particular, their approach is to model this relation via a sequence of k successive steps of a walk in the supersingular isogeny graph and to show that the respective j-invariants are roots of the second modular polynomial. They then arithmetize this relation and show that this approach, when compared to state-of-the-art tailor-made proofs of knowledge by Basso et al. (EUROCRYPT'23), gives a 3-10$\times$ improvement in proof and verification times, with comparable proof sizes.

In this paper we ask whether we can further improve the modular polynomial-based approach and generalize its application to primes $\ell > 2$, as used in some recent isogeny-based constructions. We will answer these questions affirmatively, by designing efficient arithmetizations for each $\ell \in \{2, 3, 5, 7, 13\}$ that achieve an improvement over Cong, Lai and Levin of up to 48%.

Our main technical tool and source of efficiency gains is to switch from classical modular polynomials to canonical modular polynomials. Adapting the well-known results on the former to the latter polynomials, however, is not straight-forward and requires some technical effort. We prove various interesting connections via novel use of resultant theory, and advance the understanding of canonical modular polynomials, which might be of independent interest.

1 Introduction

More than twenty years have passed since the seminal works by Couveignes [Cou06], Rostovstev, and Stolbunov [RS06] have introduced the idea of using maps between elliptic curves, called *isogenies*, for cryptographic purposes. Although their original attempts seemed too inefficient to compare with concurrent cryptosystems, later efforts in this direction [JD11, CLM+18] gave birth to a rich, and still lively, branch of cryptography. A strong reason for researchers to push into in this field is that the main problem on which it is based – namely, recovering a secret isogeny between two given elliptic curves – is considered hard even for quantum computers. Moreover, compared with other proposals for post-quantum cryptography, isogenies enjoy shorter parameters which though come at the price of slower performance. Since its proposal, isogeny-based cryptography has evolved into a very active and dynamic field and many different cryptographic applications have been proposed so far.

In this work we are particularly interested in (non-interactive) zero-knowledge proofs of knowledge of secret isogenies. They are a central tool to enforce honest behaviors in multi-party protocols.

*thomasdh@unibw.de
†soeren.kleine@unibw.de
‡marzio.mula@unibw.de
§daniel.slamanig@unibw.de
¶s.spindler@unibw.de

More specifically, in many cryptographic applications it is required that when a party presents a public key, it also needs to provide a proof-of-possession (PoP) of the corresponding secret key, to prove that the public key is well-formed. Such PoPs are for instance an important measure to prevent rogue key attacks in multi-party signature protocols [RY07]. More practically, this is required in public-key infrastructures (PKIs) when requesting the issuance of a certificate. In current PKIs based on X.509 certificates [CSF+08], these so-called certificate signing requests realize the PoPs via signatures, i.e. the knowledge of the corresponding signing key is demonstrated by computing a signature. This only works for certifying signing keys. However, in a post-quantum setting this gets more relevant, e.g. when certifying KEM keys for KEMTLS [SSW20], one explicitly requires zero-knowledge proofs [GHL+22]. Another immediate application of such proofs is the design of signature schemes obtained via the the Fiat-Shamir heuristic, e.g. GPS signatures [GPS17], CSI-FiSh [BKV19] or SQISign [DKL+20].

More recently such proofs have been studied for settings where one wants to avoid a trusted setup to generate supersingular curves of unknown endomorphism ring [BCC+23, CLL23]. Such curves are needed for several isogeny-based protocols ranging from hash functions [CLG09] to VDFs [BBBF18, DMPS19], delay encryption schemes [BD21], and public-key cryptosystems [Mor23, FMP23]. In all these applications it is central to the security that the trapdoor is discarded after the trusted setup. This is a requirement that is hard to enforce in practice. Basso et al. [BCC+23] propose to implement a sequential multi-party ceremony to replace the trusted setup. Loosely speaking, they consider a walk in the isogeny graph $E_0 \to E_1 \to \cdots \to E_k$ which starts from some (well-known) curve E_0 and then each party i takes the previous curve E_{i-1}, generates a random isogeny to a new curve E_i and provides a proof that they know the isogeny from E_{i-1} to E_i. This is a well-known technique that is often used to avoid a trusted setup for generating the structured reference string (SRS) for succinct non-interactive argument of knowledge systems (zk-SNARKs) [GKM+18], and can be seen as a variant where one uses explicit zero-knowledge proofs for the updates [AGRS24] instead of knowledge assumptions as done in [GKM+18]. Such a protocol can be used to replace a trusted setup as long as one of the parties in the chain can be assumed to be honest (i.e. discards its secret isogeny).

1.1 Previous Work

In general one can distinguish between *tailor-made* approaches and *generic* (or *general-purpose*) approaches. Subsequently, we are only focusing on work directly relevant to our approach, and we refer the reader to a recent comprehensive survey of proofs of knowledge of isogenies by Beullens et al. [BFGP23] for a complete overview.

For tailor-made approaches, the most recent work is the one of Basso et al. [BCC+23], which builds on the SIDH proof of knowledge from [DFJP14, DDGZ22] and achieves statistical zero-knowledge. One main limitation of this (and most previous approaches with the exception of [DKL+20]) is that the small challenge space requires numerous parallel executions of the protocol in order to reduce the soundness error. A more concerning problem is that the knowledge soundness achieved in [BCC+23] is not exact but only relaxed, i.e., while the relation is intended to prove knowledge of a d-isogeny, one can only extract an $\ell^{2i}d$-isogeny for some small prime ℓ and $0 \leq i \leq n$.

The second approach is to take a general-purpose (non-interactive) zero-knowledge proof system that is capable of proving any language in NP, such as a zk-SNARK, and prove the respective isogeny relation using this proof system. While tailor-made approaches might intuitively seem to be more efficient than such a generic approach, there has been enormous progress in the field of zk-SNARKs over the last decade (cf. [Tha22] for a good overview). This has led Cong, Lai and Levin [CLL23] (CLL henceforth) to look into how well such an approach can perform when concretely instantiated with various recent general-purpose zero-knowledge proof systems. As a starting point CLL take the work by Chavez-Saab, Rodríguez-Henríquez and Tibouchi [CSRT22], which constructs isogeny-based verifiable delay functions (VDFs) [BBBF18] using a succinct non-interactive argument (SNARG) system. For the evaluation of their VDF they require to prove isogeny walks over supersingular elliptic curves. In brief, for a small prime ℓ they consider the supersingular isogeny graph of ℓ-isogenous supersingular elliptic curves (represented by their j-invariants) and want to prove a walk in this graph. Their idea now is to consider the ℓ^{th} modular polynomial $\Phi_\ell(X, Y) \in \mathbb{Z}[X, Y]$, for which it holds that two curves E and E' are ℓ-isogenous if and only if their j-invariants satisfy $\Phi_\ell(j(E), j(E')) = 0$. Consequently, when aiming to prove a walk in the ℓ-isogeny graph from some starting curve E to some curve E', we can efficiently represent it as a sequence of successive steps, i.e. a sequence of j-invariants $j_0, j_1, \ldots, j_k$ such that $\Phi_\ell(j_i, j_{i+1}) = 0$ for all $i \in \{0, \ldots, k-1\}$, and $j_0 = j(E)$ and $j_k = j(E')$. This means that the relation

$$\mathcal{R}_{\ell^k\text{-IsoPath}} = \{((E, E'), \phi) \mid \phi\colon E \to E' \text{ is an isogeny of degree } \ell^k\}$$

can be equivalently represented by means of the relation

$$\mathcal{R}_{\ell^k\text{-ModPoly}} = \left\{ \left((E, E'), (j_i)_{i\in[k-1]}\right) \middle| \begin{array}{c} \Phi_\ell(j(E), j_1) = 0 \wedge \\ \Phi_\ell(j_{k-1}, j(E')) = 0 \\ \bigwedge_{i\in[k-2]} \Phi_\ell(j_i, j_{i+1}) = 0 \end{array} \right\}. \quad (1)$$

We note that [CSRT22] do not require the zero-knowledge and knowledge soundness properties for the VDF application and thus a SNARG suffices. CLL then use this relation for the specific case of a degree 2^k isogeny to construct a rank-1 constraint system (R1CS), which is a very popular arithmetization method in state-of-the-art zk-SNARKs. They then take a number of existing plausibly post-quantum zero-knowledge argument systems and in particular Aurora [BCR+19], Ligero [AHIV17] and Limbo [DOT21] which do not need to make additional structured cryptographic assumptions (e.g., such as lattice-based proof systems for R1CS [NS22, BS23]).

Although CLL focus on $\ell = 2$, it is not uncommon for isogeny-based protocols to involve, or at least allow for, other small primes. For instance, the KEM presented in [Mor23] makes use of a 3-smooth isogeny as a public key, and a 5- or 7-smooth isogeny for encapsulation. Similarly, a 3-smooth isogeny is used for the encryption in the updatable PKE scheme from [DFV24] and for the encapsulation in the KEM from [Bas24, Protocol 2]. More generally, while the choice $\ell = 2$ is usually done for simplicity, considering different small primes can provide greater flexibility and also allows for trade-offs in the efficiency between the building blocks and the isogeny proofs of knowledge.

The results in [CLL23] show that the efficiency of this general-purpose approach when compared with the recent tailor-made approach in [BCC+23] achieves an order of magnitude improvement over proof and verification times, with slightly worse but still comparable proof sizes. Moreover, compared to existing tailor-made solutions, this approach provides a stronger notion of soundness, i.e. an exact instead of a relaxed one.

In this work we ask whether this is the best we can do when targeting R1CS and whether the approach can be generalized to prove the knowledge of isogenies of degree ℓ^k for primes ℓ greater than 2.

1.2 Our Contributions

The goal of this paper is to improve on the state-of-the-art results of [CLL23] for proving the knowledge of an isogeny, and we make the following contributions.

Use of canonical modular polynomials. We consider canonical modular polynomials in place of the classical modular polynomials used in [CSRT22, CLL23] and we show that constructing a proof of knowledge of the corresponding relation $\mathcal{R}_{\ell^k\text{-ModRoot}}$ is computationally equivalent to proving the relations $\mathcal{R}_{\ell^k\text{-IsoPath}}$ and $\mathcal{R}_{\ell^k\text{-ModPoly}}$ mentioned above. While the approach via classical modular polynomials stems directly from well-known theoretical results, the same results are not as readily available for the case of canonical polynomials. Therefore we incorporate them and prove connections to the classical modular polynomials via novel use of resultant theory. We also spot a few gaps in the relevant literature and provide new proofs of some basic properties of the modular polynomials over finite fields. For example, we analyze the existence of edges of multiplicity at least three in the supersingular ℓ-isogeny graph for small ℓ – which seems to be known to experts of the field, at least in a weaker form – and we study the relationship between multiple edges in the ℓ-isogeny graph on the one hand and multiple roots of the ℓ^{th} canonical modular polynomial on the other hand. Therefore this part might also be of interest beyond the concrete application in this paper.

Improved and generalized isogeny proofs of knowledge. By moving to canonical modular polynomials we obtain a more efficient arithmetization for the equivalent relation $\mathcal{R}_{\ell^k\text{-ModPoly}}$. Moreover, while [CLL23] only consider isogenies of 2-power degrees, we generalize the approach to cover isogenies of degree ℓ^k, where $\ell \in \{2, 3, 5, 7, 13\}$. This is of interest not only because such primes are used in some recent isogeny-based constructions [Mor23, DFV24, Bas24], but also because here we reduce the number of constraints further, potentially yielding even more efficient proof systems.

We first encode our new relation into an R1CS over the field $\mathbb{F}_{p^2}$ with p^2 elements, and subsequently lift these arithmetizations to $\mathbb{F}_p \times \mathbb{F}_p$ in order to obtain a formulation which works over $\mathbb{F}_p$. We describe several techniques to minimize the non-zero entries in the constraint matrices when lifting, such as a basis change for product relations and a change of variables for linear relations, which may be applicable more broadly. Additionally, we are able to further exploit the structured nature of the canonical modular polynomials to maximize the use of the more efficient squaring relations over $\mathbb{F}_p$, minimizing the number of R1CS-constraints. We expect our system to be up to twice as efficient as [CLL23] in concrete proof systems.

1.3 Concurrent Work

In [LP24], Levin and Pedersen examine radical isogenies and develop a verifiable random function (VRF) from an efficient proof of knowledge of an isogeny. Although the theory behind the two approaches is quite different, they also obtain an arithmetization that uses the same number of constraints and number of variables asymptotically for $\ell = 2$. An advantage of their approach is that it prevents backtracking for free, whereas ours would require an additional check as in [CLL23, Appendix A]. On the other hand, ours is more general in two respects: We do not put restrictions on the prime p while they need $p = 3 \mod 4$, and we also generalize to isogeny degrees beyond $\ell = 2$, where we are able to obtain systems with fewer constraints and variables.

2 Preliminaries

2.1 Isogeny Graphs and Classical Modular Polynomials

Let K be a perfect field, and let E_0, E_1 be elliptic curves over K. An *isogeny* is a morphism of curves $\phi\colon E_0 \longrightarrow E_1$ which induces also a surjective group homomorphism on the sets of $\overline{K}$-rational points. An isogeny of degree n is also called an *n-isogeny*, and two elliptic curves E_0 and E_1 over K are called *n-isogenous* if there exists an n-isogeny $\phi\colon E_0 \longrightarrow E_1$. An isogeny of degree 1 is called an *isomorphism*, an isogeny $\phi\colon E \longrightarrow E$ is called an *endomorphism*, and an endomorphism of degree 1 is called an *automorphism*. We refer the interested reader to Appendix A.1 for other standard definitions, properties and references on elliptic curves and isogenies.

We will say that two n-isogenies $\phi_1\colon E \longrightarrow E_1$ and $\phi_2\colon E \longrightarrow E_2$ are *equivalent* if they are the same up to post-composition with an isomorphism, i.e. if there exists an isomorphism $\sigma\colon E_1 \longrightarrow E_2$ such that $\phi_2 = \sigma \circ \phi_1$. The kernel of an isogeny ϕ can be represented by its *kernel polynomial* [Koh96, §2.4], which is the square-free monic polynomial whose roots are precisely the x-coordinates of the points in the kernel of ϕ. We say that ϕ is *defined over K* if the coefficients of its kernel polynomial lie in K. Closely related to kernel polynomials is the n^{th} *division polynomial* ψ_n of an elliptic curve E (which we scale by $2y$ for even n compared to the usual definition [Was08, p. 81], so that it is always a polynomial in x only): It is the kernel polynomial of the multiplication-by-n endomorphism scaled by the factor n (resp. $2n$) for odd n (resp. even n).

To each elliptic curve E defined over a field K one can attach a *j-invariant* $j(E) \in K$ which can be computed efficiently from the coefficients of E. Two elliptic curves E_0 and E_1 are isomorphic if and only if $j(E_0) = j(E_1)$, and any $j_0 \in \overline{K}$ is the j-invariant of an elliptic curve defined over $K(j_0)$ [Sil09, Proposition III.1.4]. By a slight abuse of terminology, we will often refer to the *number of non-equivalent n-isogenies* $j_0 \to j_1$ to indicate the number of equivalence classes of n-isogenies starting from a fixed elliptic curve of j-invariant j_0 and landing on some elliptic curve of j-invariant j_1; note that the number of such equivalence classes does not depend on the choice of representative of j_0.

Now fix a prime p and let $q = p^k$ for some positive integer k. An elliptic curve E over $\mathbb{F}_q$ is called *supersingular* if the cardinality of $E(\mathbb{F}_q)$ is congruent to 1 modulo p [Was08, Proposition 4.31]. Given a prime $\ell \neq p$, the *supersingular ℓ-isogeny graph* $\mathcal{G}_\ell(p)$ is defined as follows: The vertices of $\mathcal{G}_\ell(p)$ shall be the isomorphism classes of supersingular elliptic curves, which we parametrize by their j-invariants in $\mathbb{F}_{p^2}$ [Sil09, Theorem V.3.1], and the number of edges $j_0 \to j_1$ is precisely the number of non-equivalent ℓ-isogenies $j_0 \to j_1$. The graph $\mathcal{G}_\ell(p)$ is connected, $(\ell+1)$-regular, and it is Ramanujan [Piz90, CL24]. Furthermore, since every isogeny admits a dual isogeny [Sil09, Theorem III.6.1-2], it can almost be considered as an undirected graph; however, curves of j-invariants 0 or 1728 have special automorphisms [Sil09, Theorem III.10.1], which cause asymmetries in the graph for $p \geq 5$: If we write $\mu(0) := 3$, $\mu(1728) := 2$ and $\mu(j) := 1$ for $j \notin \{0, 1728\}$, then there are $\mu(j_0)/\mu(j_1)$ times as many non-equivalent ℓ-isogenies $j_0 \to j_1$ as there are non-equivalent ℓ-isogenies $j_1 \to j_0$ (cf. [AAM19, Formula (11)]).

In this paper we will consider random walks on $\mathcal{G}_\ell(p)$, i.e. sequences

$$j(E_0) \longrightarrow j(E_1) \longrightarrow \ldots \longrightarrow j(E_k)$$

of adjacent j-invariants in $\mathcal{G}_\ell(p)$. It is easy to check whether two given j-invariants belong to a pair of ℓ-isogenous elliptic curves. To this end, one can use the so-called *classical modular polynomials* (see [Mü95, §4.3], [AAM19, §2.4] and [Sut12, §2.3]): The ℓ^{th} classical modular polynomial $\Phi_\ell(X, Y)$ is a two-variable polynomial with integer coefficients whose roots are given by the pairs of j-invariants of ℓ-isogenous elliptic curves – more precisely, given the prime ℓ and any two elliptic curves E, E' over a field K with $\mathrm{char}(K) \neq \ell$, the number of non-equivalent ℓ-isogenies $E \to E'$ is equal to the multiplicity of $j(E')$ as a root of $\Phi_\ell(j(E), Y)$.

2.2 Resultants

Let R be an integral domain and let $f, g \in R[X]$ be non-zero polynomials. The *Sylvester matrix* of f and g and especially its determinant, the *resultant* of f and g, are important algebraic tools to detect common divisors between two polynomials. We only state the necessary properties here and give the theoretical background together with proofs for the below results in Appendix B.

Proposition 1. Let R be an integral domain, let $f, g \in R[X]$ be non-zero polynomials and let $\varphi \colon R \to S$ be a ring homomorphism of integral domains, extended to a ring homomorphism $\varphi \colon R[X] \to S[X]$ via coefficient-wise application. Then the following holds:

(a) If φ preserves the degress of f and g, then

$$\mathrm{res}(\varphi(f), \varphi(g)) = \varphi(\mathrm{res}(f, g)).$$

(b) $\mathrm{res}(f, g) = 0$ if and only if f and g share a common divisor of positive degree.

Proof. Lemma 17 and Corollary 18. □

Remark 1. We note that, as the *discriminant* of f is defined as the resultant of f and $\frac{\partial}{\partial X} f$ up to scaling, the above results also translate to the discriminant; see Corollary 23 for a precise formulation of the first claim for discriminants.

Proposition 2. Let $R = K[Y]$ be a polynomial ring over a field K and fix an element $y_0 \in K$. Additionally let $f, g \in R[X]$ be non-zero polynomials and extend the K-linear evaluation homomorphism $\varphi \colon R \to K$ given by $Y \mapsto y_0$ to a ring homomorphism $\varphi \colon R[X] \to K[X]$ via coefficient-wise application. Further suppose that φ preserves the X-degrees of f and g, and write

$$m := \deg \gcd(\varphi(f), \varphi(g)).$$

Then

$$\left.\frac{\partial^k}{\partial Y^k}\right|_{Y=y_0} \mathrm{res}(f, g) = 0 \text{ for } k \in \{0, \ldots, m-1\},$$

i.e. $\mathrm{res}(f, g) \in K[Y]$ has a root of multiplicity at least m at y_0.

Proof. Corollary 20. □

2.3 Zero-Knowledge Argument Systems and R1CS

As stated before, in this work we will use generic techniques to prove the knowledge of isogenies, improving and expanding on the previous results of [CLL23]. Since our arithmetization is broadly applicable across different argument systems, and we do not need any formal properties of an argument system throughout this work, we will omit a full formal treatment of zk-SNARKS. For a comprehensive formal treatment, readers are referred to the respective proof systems [BCR+19, AHIV17, XZZ+19].

A *zk-SNARK* is a non-interactive argument system that is complete, knowledge-sound, zero-knowledge and succinct. The proving algorithm takes a statement and witness pair (s, w) for some NP-relation and generates a proof π. There is a verification algorithm to check the validity of a proof. Completeness indicates that a valid proof can be generated form any pair (s, w) in the relation. Knowledge soundness means that any prover that can generate a valid proof for a statement s needs to know a corresponding witness w. Zero-knowledge means that the proof does not reveal any information about the witness. An argument system is succinct if the proof size is small and the proof can be verified efficiently. Usually, both proof size and verifier time are required to be polynomial in $|x|$ and polylogarithmic in $|w|$.

In this paper we design an efficient arithmetization in the form of a rank-1 constraint system (R1CS). This represents a popular choice and this allows us to cover many different proof systems. An R1CS is defined as follows:

Definition 1 ([BCR+19])**.** The relation $\mathcal{R}_{\text{R1CS}}$ is the set of pairs $((\mathbb{F}, k, n, m, A, B, C, v), w)$ where $\mathbb{F}$ is a finite field, $k, n, m \in \mathbb{N}$ denote the numbers of inputs, variables and constraints respectively ($k \leq n$), A, B, C are $m \times (1+n)$-matrices over $\mathbb{F}$, $v \in \mathbb{F}^k$, and $w \in \mathbb{F}^{n-k}$, such that for all $i \in [m]$

$$\left(\sum_{j=0}^{n} A_{ij} z_j\right) \cdot \left(\sum_{j=0}^{n} B_{ij} z_j\right) = \left(\sum_{j=0}^{n} C_{ij} z_j\right),$$

where $(1, v, w) =: z = (z_j)_j \in \mathbb{F}^{n+1}$.

It is worth noting that the efficiency of proving and verifying knowledge of a witness may depend on different aspects of the R1CS, depending on the proof system that is used. For example, the prover time of pairing-based SNARKs is usually $O(n)$, the proof size is constant, and the verifier time is $O(k)$ [Gro16, Lip22]. On the other hand, [BCR+19, AHIV17, XZS22] have prover time proportional to the circuit size, which corresponds to the number of non-zero entries in the constraint matrices A, B and C, which we will denote by nnz. Lastly, for [DOT21] the proof size and prover and verifier times seem to be determined by the number of multiplications, corresponding to the number of R1CS constraints m. The efficiency of the arithmetization in [CLL23] is only quantified through the number of constraints m and optimized using this metric.

In this work, we will provide all of n, m and nnz. When optimizing, we will focus on the latter two, since the number of variables is mostly relevant for the non-post-quantum secure pairing-based SNARKs. Often optimizing for one metric also improves another, such as when a linear constraint can be removed to eliminate a variable, but this is not always the case. As we will see in Section 5, we achieve very efficient constraint systems in terms of all three metrics.

3 Canonical Modular Polynomials

The classical modular polynomials tend to have many non-vanishing coefficients, which makes these polynomials quite expensive to handle in an R1CS. To be more precise, the polynomial $\Phi_\ell(X, Y)$ is symmetric in X and Y, of degree $\ell + 1$ in both variables [Lan87, Theorem 5.2.3], and typically most of the possible mixed monomials $X^i Y^j$ with $i, j \leq \ell$ occur.

For example, the third classical modular polynomial is given by (see [CFA+06, Example 17.18])

$$\begin{aligned}\Phi_3(X,Y) = &-X^3Y^3 + X^4 + Y^4 + 2232(X^3Y^2 + X^2Y^3)\\ &- 1069956(X^3Y + XY^3) + 36864000(X^3 + Y^3)\\ &+ 2587918086X^2Y^2 + 8900222976000(X^2Y + XY^2)\\ &+ 452984832000000(X^2 + Y^2) - 770845966336000000XY\\ &+ 1855425871872000000000(X + Y).\end{aligned}$$

Luckily, there exists a related class of polynomials called *canonical modular polynomials* Φ_ℓ^c, which are asymmetric, and have a smaller degree in the second variable. To contrast our previous example, the third canonical modular polynomial is given by

$$\Phi_3^c(X,j) = X^4 + 36X^3 + 270X^2 + 756X + 729 - X \cdot j.$$

One actually has a concrete formula for the degree κ in the second variable j (see [Mü95, Equation (5.1) and Lemma 5.7]): It is given by $\kappa = \frac{s\cdot(\ell-1)}{12}$, where s is the smallest non-zero natural number such that the fraction defining κ is an integer; in other words,

$$s = \frac{12}{\gcd(12,\ell-1)} \quad \text{and} \quad \kappa = \frac{\ell-1}{\gcd(12,\ell-1)}.$$

The ℓ^{th} canonical modular polynomial thus has the form

$$\Phi_\ell^c(X,j) = \sum_{i=0}^{\ell+1}\sum_{m=0}^{\kappa} c_{i,m}X^i j^m$$

for suitable coefficients $c_{i,m} \in \mathbb{Z}$. In the case $\kappa = 1$, i.e. $\ell \in \{2,3,5,7,13\}$, we can be more precise; we give all canonical modular polynomials for this case in Appendix A.2.

Lemma 3. Let $\ell \in \{2,3,5,7,13\}$. Then the ℓ^{th} canonical modular polynomial $\Phi_\ell^c(X,j)$ is of the form

$$\Phi_\ell^c(X,j) = X^{\ell+1} + \sum_{i=1}^{\ell} c_i X^i + \ell^s - X\cdot j = \Phi_\ell^c(X,0) - X\cdot j$$

with integer coefficients $c_i := c_{i,0} \in \mathbb{Z}$. In particular, the rational function

$$\Theta_\ell^c(X,j) := \Phi_\ell^c(\ell^s/X,j)\cdot X^{\ell+1}/\ell^s = X^{\ell+1} + \sum_{i=1}^{\ell} c_i\ell^{s\cdot(i-1)}X^{\ell+1-i} + \ell^{s\cdot\ell} - X^\ell \cdot j$$

is a monic polynomial with integer coefficients. Setting $\mathcal{J}_\ell(X) := \Phi_\ell^c(X,0)/X$, we furthermore have that for any $f \in K^\times$ and $j_0 \in K$:

$$j_0 = \mathcal{J}_\ell(f) \text{ if and only if } \Phi_\ell^c(f,j_0) = 0.$$

Proof. The claim on the form of $\Phi_\ell^c(X,j)$ follows from direct inspection, and the claim on the form of $\Theta_\ell^c(X,j)$ is an immediate consequence. Furthermore we have $\Phi_\ell^c(X,j) = X\cdot\mathcal{J}_\ell(X) - X\cdot j$, which implies the claim on $\mathcal{J}_\ell(X)$. □

Since the constant coefficients of $\Phi_\ell^c(X,j)$ and $\Theta_\ell^c(X,j)$ are powers of ℓ, we obtain:

Corollary 4. Let $\ell \in \{2,3,5,7,13\}$, assume $\mathrm{char}(K) \neq \ell$ and let $j,j' \in K$ with $j \neq j'$. Then

$$\gcd(\Phi_\ell^c(X,j),\Phi_\ell^c(X,j')) = 1 = \gcd(\Theta_\ell^c(X,j),\Theta_\ell^c(X,j')).$$

3.1 The Multiplicity Theorem

Unfortunately, it is no longer true that $\Phi_\ell^c(j(E_0), j(E_1)) = 0$ if E_0 and E_1 are ℓ-isogenous over $\overline{\mathbb{F}_p}$. Instead, taking inspiration from [CFA+06, §17.2.3], we will see that we need to find a common root X of the two functions $\Phi_\ell^c(X, j(E_0))$ and $\Phi_\ell^c(\ell^s/X, j(E_1))$, where s is defined as above. To prove a more precise version of this claim, we first relate the classical modular polynomial to the canonical modular polynomial via resultant theory.

However, we have to restrict ourselves to the case $\kappa = 1$. The corresponding primes – namely, $\ell \in \{2, 3, 5, 7, 13\}$ – are in fact exactly those for which the modular curve $X_0(\ell)$ has genus 0. This fact gives a high-level intuition of why these primes provide a compact representation of ℓ-isogenies: The elements of $X_0(\ell)$, that are the edges in $\mathcal{G}_\ell(p)$, can be parametrized (up to issues at the 'ramified' points $j \in \{0, 1728\}$) by $f \in \overline{\mathbb{F}_p}^\times$. This parametrization, which we will later analyze and exploit, has already been studied in the works of Fricke [Fri11, Section 2, Chapters 4-5], Mestre [Mes86, §5], and Elkies [Elk98, §4].

Corollary 5. For any $\ell \in \{2, 3, 5, 7, 13\}$ we have

$$\operatorname{res}(\Theta_\ell^c(X, J_1), \Phi_\ell^c(X, J_0)) = \ell^{s\cdot\ell} \cdot \Phi_\ell(J_0, J_1),$$

where the resultant is computed over the coefficient ring $R = \mathbb{Z}[J_0, J_1]$ of bivariate polynomials. In particular, suppose that we have a field K of characteristic $\operatorname{char}(K) \neq \ell$ as well as $j_0, j_1 \in K$ with $m := \deg \gcd(\Phi_\ell^c(X, j_0), \Theta_\ell^c(X, j_1))$. Then

$$\frac{\partial^k}{\partial J_1^k}\bigg|_{J_1 = j_1} \Phi_\ell(j_0, J_1) = 0 \text{ for } k \in \{0, \ldots, m-1\},$$

i.e. there are at least m non-equivalent ℓ-isogenies from j_0 to j_1.

Proof. The first claim can be checked by direct computation. For the second claim we first apply the ring homomorphism $\mathbb{Z}[J_0] \to K$ defined by $J_0 \mapsto j_0$ in view of Proposition 1(a) to obtain

$$\operatorname{res}(\Theta_\ell^c(X, J_1), \Phi_\ell^c(X, j_0)) = \ell^{s\cdot\ell} \cdot \Phi_\ell(j_0, J_1)$$

as an equality in $K[J_1]$. Now we consider the K-linear evaluation homomorphism $\varphi\colon K[J_1] \to K$ given by $J_1 \mapsto j_1$, extended via coefficient-wise application to a homomorphism $\varphi\colon K[J_1][X] \to K[X]$. As this homomorphism preserves the X-degrees of $f := \Theta_\ell^c(X, J_1)$ and $g := \Phi_\ell^c(X, j_0)$, we are exactly in the situation of Proposition 2. Therefore we deduce that $\operatorname{res}(f, g) = \ell^{s\cdot\ell} \cdot \Phi_\ell(j_0, J_1)$ has a root of multiplicity at least m at j_1, and $\operatorname{char}(K) \neq \ell$ yields the claim. □

The previous relation will be the main tool in establishing the desired connection between the classical and the canonical modular polynomial. For the proof we need to analyze root multiplicities of the canonical modular polynomial in the next two results, which can also be found in [Tsu13, §4.3] in the language of modular curves.

Lemma 6. Let $\ell \in \{2, 3, 5, 7, 13\}$, assume $\operatorname{char}(K) \neq \ell$ and let $j_0 \in K$. Then $\Phi_\ell^c(X, j_0)$ has a double root in $\overline{K}$ if and only if $j_0 = 0$ or $j_0 = 1728$.

Proof. To simplify notation we may assume that K is algebraically closed. In view of Lemma 3 we obtain the univariate polynomial

$$\mathcal{D}_\ell(X) := j + \tfrac{\partial}{\partial X}\Phi_\ell^c(X, j).$$

Now an element $f \in K$ is a double root of $\Phi_\ell^c(X, j_0)$ if and only if it is non-zero (due to $\operatorname{char}(K) \neq \ell$) and satisfies $\mathcal{J}_\ell(f) = j_0 = \mathcal{D}_\ell(f)$. From this we see that the double roots are precisely the common roots of $\Phi_\ell^c(X, j_0)$ and the polynomial

$$\mathcal{D}_\ell(X) \cdot X - \mathcal{J}_\ell(X) \cdot X = -\Phi_\ell^c(X, \mathcal{D}_\ell(X)),$$

which has leading coefficient ℓ. Via direct computation one now confirms that there are $e, m, n \in \mathbb{N}$ such that

$$\operatorname{res}(\Phi_\ell^c(X, J), -\Phi_\ell^c(X, \mathcal{D}_\ell(X))) = (-1)^\ell \cdot \ell^e \cdot (J - 0)^m \cdot (J - 1728)^n,$$

computed over the coefficient ring $\mathbb{Z}[J]$. Therefore $\mathrm{char}(K) \neq \ell$ allows us to apply Proposition 1 (with the homomorphism $\mathbb{Z}[J] \to K$ given by the evaluation $J \mapsto j_0$) to deduce the claimed equivalence. □

The second result discusses the special j-invariants 0 and 1728; for an explicit list of the factors given below we refer the reader to Appendix A.3.

Lemma 7. Let $\ell \in \{2,3,5,7,13\}$. Then there are monic polynomials $g_{\ell,0}$ and $g_{\ell,1728}$ in $\mathbb{Z}[X]$ of degree at most 2 and monic non-constant polynomials $h_{\ell,0,\pm}$ and $h_{\ell,1728,\pm}$ in $\mathbb{Z}[X]$ such that

$$\Phi_\ell^c(X,0) = g_{\ell,0} \cdot h_{\ell,0,+}^3, \quad \Theta_\ell^c(X,0) = g_{\ell,0} \cdot h_{\ell,0,-}^3$$

and

$$\Phi_\ell^c(X,1728) = g_{\ell,1728} \cdot h_{\ell,1728,+}^2, \quad \Theta_\ell^c(X,1728) = g_{\ell,1728} \cdot h_{\ell,1728,-}^2.$$

Moreover, if K is a field with $\mathrm{char}(K) \notin \{2,3,\ell\}$ and $j^* \in \{0,1728\}$, then each $h_{\ell,j^*,\pm}$ does neither have a double root nor share a root with g_{ℓ,j^*} in $\overline{K}$.

Proof. The factorizations follow from direct computations. For $j^* \in \{0,1728\}$ one further verifies that the prime factors of $\deg(h_{\ell,j^*,\pm})$ lie in $\{2,3\}$, and that the prime factors of $\mathrm{disc}(h_{\ell,j^*,\pm})$ and $\mathrm{res}(g_{\ell,j^*}, h_{\ell,j^*,\pm})$ lie in $\{2,3,\ell\}$. To prove the additional claim, we now make use of resultant theory once more by considering the unique homomorphism $\varphi\colon \mathbb{Z} \to K$. This homomorphism preserves the degree of $h_{\ell,j^*,\pm}$ and, as the prime factors of $\deg(h_{\ell,j^*,\pm})$ are contained in $\{2,3\}$, the degree of $\frac{\partial}{\partial X} h_{\ell,j^*,\pm}$ due to our assumption on $\mathrm{char}(K)$. Therefore this assumption and Proposition 1(a) (cf. Remark 1) yield

$$\mathrm{disc}(\varphi(h_{\ell,j^*,\pm})) = \varphi(\mathrm{disc}(h_{\ell,j^*,\pm})) \neq 0,$$

so $h_{\ell,j^*,\pm}$ cannot have a double root in $\overline{K}$ by Proposition 1(b). Similarly we obtain

$$\mathrm{res}(\varphi(g_{\ell,j^*}), \varphi(h_{\ell,j^*,\pm})) = \varphi(\mathrm{res}(g_{\ell,j^*}, h_{\ell,j^*,\pm})) \neq 0,$$

i.e. g_{ℓ,j^*} and $h_{\ell,j^*,\pm}$ cannot have a common root in $\overline{K}$ by Proposition 1(b). □

With the above preparations we are finally ready to state and prove the following crucial relation between the classical and canonical modular polynomial:

Multiplicity Theorem. *Let $\ell \in \{2,3,5,7,13\}$, let K be a field of characteristic $\mathrm{char}(K) \neq \ell$ and let $j_0, j_1 \in K$. Then there are exactly as many non-equivalent ℓ-isogenies $j_0 \to j_1$ as there are roots $f \in \overline{K}^\times$ of $\Phi_\ell^c(X, j_0)$ (counted with multiplicity) such that $\Phi_\ell^c(\ell^s/f, j_1) = 0$. In particular, j_0 and j_1 are ℓ-isogenous if and only if there exists an $f \in \overline{K}^\times$ such that*

$$\Phi_\ell^c(f, j_0) = 0 = \Phi_\ell^c(\ell^s/f, j_1). \tag{2}$$

Proof. To simplify notation we assume that K is algebraically closed. As before we consider the polynomial $\Theta_\ell^c(X, j_1) = \Phi_\ell^c(\ell^s/X, j_1) \cdot X^{\ell+1}/\ell^s$ instead of the rational function $\Phi_\ell^c(\ell^s/X, j_1)$, noting that it has the same roots (with the same multiplicities) since $\mathrm{char}(K) \neq \ell$. For any $j_0, j_1 \in K$ we write $\nu_\ell(j_0, j_1)$ for the number of roots of $\Phi_\ell^c(X, j_0)$ (counted with multiplicity) that are roots of $\Theta_\ell^c(X, j_1)$, and we write $\iota_\ell(j_0, j_1)$ for the number of non-equivalent ℓ-isogenies from j_0 to j_1.

Hence our goal is to show that $\nu_\ell(j_0, j_1) = \iota_\ell(j_0, j_1)$; however, it suffices to prove the inequality $\nu_\ell(j_0, j_1) \leq \iota_\ell(j_0, j_1)$ for all $j_0, j_1 \in K$. Indeed, summing both quantities over all possible j_1 for a fixed $j_0 \in K$ then yields

$$\begin{aligned} \ell + 1 &= \deg_X(\Phi_\ell^c(X, j_0)) = \sum_{j_1 \in K} \nu_\ell(j_0, j_1) \leq \sum_{j_1 \in K} \iota_\ell(j_0, j_1) \\ &= \deg_Y(\Phi_\ell(j_0, Y)) = \ell + 1 \end{aligned}$$

in view of Lemma 3, and thus all inequalities have to be equalities.

To prove the inequalities, we first note that Corollary 5 yields

$$\deg \gcd(\Phi_\ell^c(X, j_0), \Theta_\ell^c(X, j_1)) \leq \iota_\ell(j_0, j_1). \tag{3}$$

We will use this inequality in the following, but we need to work through a slightly tedious case distinction: First, if $j_0 \notin \{0, 1728\}$, then any root of $\Phi_\ell^c(X, j_0)$ is a simple root by Lemma 6, so

$$\nu_\ell(j_0, j_1) = \deg\gcd(\Phi_\ell^c(X, j_0), \Theta_\ell^c(X, j_1)) \leq \iota_\ell(j_0, j_1).$$

Next we consider an edge case: Let $\mathrm{char}(K) \in \{2, 3\}$ and let $j_0 = j^* = 0 = 1728$ be the unique supersingular j-invariant in K (cf. [Sil09, §V.4]). Then all $\ell + 1$ non-equivalent ℓ-isogenies starting from j^* are loops, and one can directly verify that we have $\Theta_\ell^c(X, 0) = \Phi_\ell^c(X, 0)$ over K. Therefore, for $j_1 \neq j^*$, Corollary 4 yields

$$\iota_\ell(j^*, j^*) = \ell + 1 = \nu_\ell(j^*, j^*) \quad \text{and} \quad \iota_\ell(j^*, j_1) = 0 = \nu_\ell(j^*, j_1);$$

thus we may now assume $\mathrm{char}(K) \notin \{2, 3, \ell\}$. To proceed we recall the multiplicity factors

$$\mu(0) = 3, \ \ \mu(1728) = 2 \ \text{ and } \ \mu(j) = 1$$

for $j \in K \setminus \{0, 1728\}$, and consider a special j-invariant $j_0 = j^* \in \{0, 1728\}$. For $j^* \neq j_1$ any (distinct) root f of $\Phi_\ell^c(X, j^*)$ such that $\Theta_\ell^c(f, j_1) = 0$ then has multiplicity $\mu(j^*)$ as a root of $\Phi_\ell^c(X, j^*)$ and multiplicity $\mu(j_1)$ as a root of $\Theta_\ell^c(X, j_1)$. Indeed, if $j_1 \notin \{0, 1728\}$ the second claim follows from Lemma 6, and otherwise all multiplicities are derived from Lemma 7 since Corollary 4 and the assumption $j^* \neq j_1$ force f to be a root of $h_{\ell,j^*,+}$ and $h_{\ell,j_1,-}$ (if $j_1 \in \{0, 1728\}$), which both only have simple roots by our restriction on the characteristic.

With inequality (3) and $\mu(1728) < \mu(0)$ we hence see that the roots of the greatest common divisor of $\Phi_\ell^c(X, 1728)$ and $\Theta_\ell^c(X, 0)$ all have multiplicity $\mu(1728)$, so we obtain

$$\nu_\ell(1728, 0) = \deg\gcd(\Phi_\ell^c(X, 1728), \Theta_\ell^c(X, 0)) \leq \iota_\ell(1728, 0).$$

Using the multiplicity-preserving correspondence $f \mapsto \ell^s/f$ between roots of $\Phi_\ell^c(X, j)$ and roots of $\Theta_\ell^c(X, j)$, we further deduce

$$\nu_\ell(j^*, j_1) = \tfrac{\mu(j^*)}{\mu(j_1)} \cdot \nu_\ell(j_1, j^*) \leq \tfrac{\mu(j^*)}{\mu(j_1)} \cdot \iota_\ell(j_1, j^*) = \iota_\ell(j^*, j_1)$$

in all other cases with $j^* \neq j_1$, where the middle inequality has been derived in previous cases and the last equality is due to the larger automorphism groups at the special j-invariants 0 and 1728, as explained in Subsection 2.1.

The final case to consider is $j^* = j_1$, where Lemma 7 also does the heavy lifting: Here any root f of $\Phi_\ell^c(X, j^*)$ is either a root of g_{ℓ,j^*} – then with the same multiplicity for both $\Phi_\ell^c(X, j^*)$ and $\Theta_\ell^c(X, j^*)$ – or it is a root of both $h_{\ell,j^*,\pm}$, in which case its multiplicity for both $\Phi_\ell^c(X, j^*)$ and $\Theta_\ell^c(X, j^*)$ is $\mu(j^*)$ since each $h_{\ell,j^*,\pm}$ only has simple rots. Therefore we also deduce

$$\nu_\ell(j^*, j^*) = \deg\gcd(\Phi_\ell^c(X, j^*), \Theta_\ell^c(X, j^*)) \leq \iota_\ell(j^*, j^*)$$

from inequality (3), and this finishes the proof. □

Remark 2. In fact, it is true for all primes $\ell \in \mathbb{N}$ and ℓ-isogenous j-invariants $j_0, j_1 \in \overline{\mathbb{F}_p}$ that we can always find a common root of system (2). Indeed, we can first view j_0 and j_1 as reductions modulo p of CM j-invariants $J_0, J_1 \in \mathbb{C}$ ([Lan87, Theorem 13.5.14]) that are integral by [Cox13, Theorem 11.1]. Now equations (5.2-4) in [Mü95, §5] show that there is a common solution $\tilde{f} \in \mathbb{C}$ that has to be integral as it satisfies the polynomial $\Phi_\ell^c(X, J_0)$; therefore it can be reduced to a solution $f \in \overline{\mathbb{F}_p}$ of system (2).

However, the restriction to $\kappa = 1$ is crucial for the other direction: For example, for $\ell = 11$, $p = 61$ the j-invariants $j_0 = 41$ and $j_1 = 37$ are not ℓ-isogenous over $\overline{\mathbb{F}}_p$; in fact, they are not even isogenous since j_0 is supersingular, whereas j_1 is ordinary. Nonetheless, either root of the polynomial $X^2 + 3X - 27 \in \mathbb{F}_{p^2}[X]$ gives a solution to system (2).

In spite of that, experiments seem to suggest that there are still *at most* as many ℓ-isogenies $j_0 \to j_1$ as there are roots from j_0 to j_1 (counted as in the Multiplicity Theorem). Note that this does not contradict our previous findings since Corollary 4 fails for $\kappa > 1$: For instance, in our example we have

$$\gcd(\Phi_{11}^c(X, 41), \Phi_{11}^c(X, 50)) = X^2 - 30X - 1 \in \mathbb{F}_p[X].$$

3.2 Isogeny Relations and Root Computation

Recall that our goal is to build an efficient proof of knowledge for isogenies of degree ℓ^k (where $\ell \in \{2,3,5,7,13\}$), or, equivalently, for the relation $\mathcal{R}_{\ell^k\text{-ModPoly}}$ (Eq. (1)). However, to apply the canonical modular polynomials we instead need to consider the relation

$$\mathcal{R}_{\ell^k\text{-ModRoot}} := \left\{ \begin{array}{c} ((E,E'), \\ \left((j_i)_{i\in[k-1]}, (f_i)_{i\in[k]}\right)) \end{array} \;\middle|\; \begin{array}{c} \Phi^c_\ell(f_1, j(E)) = 0 \wedge \\ \Theta^c_\ell(f_k, j(E')) = 0 \\ \bigwedge_{i\in[k-1]} \Theta^c_\ell(f_i, j_i) = 0 = \Phi^c_\ell(f_{i+1}, j_i) \end{array} \right\}.$$

With the Multiplicity Theorem we see that simply omitting the roots $(f_i)_{i\in[k]}$ from the witness brings us back to the previous relation, so this new relation is harder to prove. To gauge how much harder it is, we need to investigate two questions: *How many roots can system* (2) *have? And in which field are those roots contained?*

To answer the first question in view of the Multiplicity Theorem, we investigate the number of ℓ-isogenies between j-invariants more closely. The following result is a consequence of the well-known structure of *ordinary isogeny volcanoes*:

Proposition 8. Let ℓ be a prime and suppose that we have two j-invariants $j_0, j_1 \in \overline{\mathbb{F}_p}$ for some prime $p \neq \ell$. If j_0 is ordinary, then the following holds:

(a) If $j_0 \neq 0$ or $j_1 = 0$, then there are at most two non-equivalent ℓ-isogenies $j_0 \to j_1$.

(b) If $j_0 = 0 \neq j_1$, then there are at most three non-equivalent ℓ-isogenies $j_0 \to j_1$ and at most one non-equivalent ℓ-isogeny $j_1 \to j_0$.

(c) If $j_0 = 1728 \neq j_1$, then there are at most two non-equivalent ℓ-isogenies $j_0 \to j_1$ and at most one non-equivalent ℓ-isogeny $j_1 \to j_0$.

Proof. This follows immediately from [Sut12, Theorem 7 & Remark 8] (note, however, that for $\ell = 2$ there is exactly 1 vertex at the first level of the ordinary isogeny graph component containing 1728 – the second formula given in Remark 8 only holds for odd ℓ). □

Remark 3. Müller claims in [Mü95, Lemma 4.14] that the $\ell + 1$ non-equivalent ℓ-isogenies defined on an ordinary curve over $\mathbb{F}_p$ with j-invariant not in $\{0, 1728\}$ map to $\ell + 1$ distinct j-invariants, i.e. to $\ell + 1$ non-isomorphic elliptic curves. However, if we consider $p = 29$, $\ell = 7$ and the two j-invariants $j_0 = 23$ and $j_1 = 12$ (noting that $1728 \equiv 17 \bmod 29$), then the curve

$$E_0\colon y^2 = x^3 + 21x + 26$$

satisfies $j(E_0) = 23 = j_0$ and admits two 7-isogenies α_1 and α_2 (defined over $\mathbb{F}_{29}$) to the elliptic curve

$$E_1\colon y^2 = x^3 + 6x + 9$$

of j-invariant $j(E_1) = 12 = j_1$. Importantly, the kernels of these two 7-isogenies are distinct (and hence the isogenies are not equivalent) since their kernel polynomials

$$x^3 + 2x^2 + 21x + 16 \;\text{ and }\; x^3 + 14x^2 + 13x + 23$$

are distinct. The issue is that the degree 7^2 endomorphism $\hat{\alpha}_2 \circ \alpha_1$ is not equivalent to the multiplication-by-7 isogeny $[7]$ on E_0. Indeed, this can be checked by comparing the kernel polynomial of $\hat{\alpha}_2 \circ \alpha_1$ to the kernel polynomial of $[7]$, the latter of which is equal to the 7^{th} division polynomial on E_0 up to the factor 7.

In general, Proposition 8 does not extend to supersingular j-invariants – the following example can easily be checked with either modular polynomial according to the Multiplicity Theorem:

Example 1. For $\ell = 7$ and $p = 71$ there are 6 non-equivalent ℓ-isogenies $0 \to 48$, 2 non-equivalent ℓ-isogenies $48 \to 0$ and 4 non-equivalent ℓ-isogenies $40 \to 40$.

Luckily, we can strictly limit when the claims of Proposition 8 do not transfer to supersingular j-invariants in our setting:

Theorem 9. *Let $\ell \leq 13$ be a prime. Then there is a prime $p_\ell < 4\ell^3$ (given in Table 1) such that for any prime $p > p_\ell$ and any two supersingular j-invariants $j_0, j_1 \in \mathbb{F}_{p^2}$ the following holds:*

(a) *If $j_0 \neq 0$ or $j_1 = 0$, then there are at most two non-equivalent ℓ-isogenies $j_0 \to j_1$.*

(b) *If $j_0 = 0 \neq j_1$, then there are at most three non-equivalent ℓ-isogenies $j_0 \to j_1$ and at most one non-equivalent ℓ-isogeny $j_1 \to j_0$.*

(c) *If $j_0 = 1728 \neq j_1$, then there are at most two non-equivalent ℓ-isogenies $j_0 \to j_1$ and at most one non-equivalent ℓ-isogeny $j_1 \to j_0$.*

Proof. We first focus on the case $j_0 \neq 0$ in claim (a), where we have to prove that there are at most two non-equivalent ℓ-isogenies $j_0 \to j_1$. If this is not the case, then $\Phi_\ell(j_0, Y)$ has a triple root at $Y = j_1$, and in Appendix B – specifically Proposition 22 – we argue computationally that this can only happen up to a prime $p_\ell < 4\ell^3$ due to $j_0 \neq 0$.

Next, by factoring $\Phi_\ell(0, 1728) \in \mathbb{Z}$, we see that 0 and 1728 cannot be ℓ-isogenous for $p > p_\ell$. Now consider $j_0 = 0$, which forces $p \equiv 2 \mod 3$ (cf. [Sil09, Example V.4.4]). For $\ell \in \{2, 3\}$ one directly confirms that, due to $p > p_\ell$, $\Phi_\ell(0, Y)$ has a triple root $j_1 \in \mathbb{F}_p \setminus \{0, 1728\}$, and for $\ell = 3$ a single root at 0. Since the number of non-equivalent ℓ-isogenies $0 \to j_1$ is three times the number of non-equivalent ℓ-isogenies $j_1 \to 0$, both claims hence hold for $j_0 = 0$ here.

For $\ell > 3$ we can use $p > p_\ell > 3\ell^2$ to apply [LOX20, Theorem 2(2)], which directly yields claim (b) and further shows that there are exactly

$$(\ell + 1) - 3 \cdot \tfrac{1}{3}\left(\ell - \left(\tfrac{\ell}{3}\right)\right) = 1 + \left(\tfrac{\ell}{3}\right) \leq 2$$

non-equivalent ℓ-isogenies $0 \to 0$, thus also finishing the proof of claim (a).

Finally, for $j_0 = 1728 \neq j_1$ we have already shown that there are at most two non-equivalent ℓ-isogenies $j_0 \to j_1$ in (a), and that j_0 is not ℓ-isogenous to 0. Thus the number of non-equivalent ℓ-isogenies $1728 \to j_1$ is exactly twice the number of non-equivalent ℓ-isogenies $j_1 \to 1728$, and claim (c) follows. □

ℓ	2	3	5	7	11	13
p_ℓ	13	53	379	1217	5101	8387
j_0	5	$6a + 28$	$117a + 322$	$379a + 173$	$977a + 4220$	$326a + 4482$
j_1	5	$47a + 28$	$262a + 322$	$838a + 173$	$4124a + 4220$	$8061a + 4482$

Table 1: Maximal primes p_ℓ for which a pair (j_0, j_1) of non-zero j-invariants with at least three ℓ-isogenies between them exists (a is a square root of 349 modulo p_ℓ).

Remark 4. For any prime $\ell \in \mathbb{N}$ one can find a prime p_ℓ as in Theorem 9, and one has the bound $p_\ell < 4\ell^4$. Indeed, due to [LOX20, Theorem 2] it suffices to consider the situation where we have at least three non-equivalent ℓ-isogenies $j_0 \to j_1$ for $j_0 \notin \{0, 1728\}$ or $(j_0, j_1) = (0, 1728)$. In view of [BCNE+18, Theorem 4.10] we can then construct two non-commuting endomorphisms of degree ℓ^2 on a curve E with $j(E) = j_0$ by composing two non-equivalent ℓ-isogenies $j_0 \to j_1$ with a suitable ℓ-isogeny $j_1 \to j_0$. Hence we obtain two embeddings of quadratic orders into the endomorphism ring of E with distinct images, and these embeddings can be extended to optimal embeddings of (possibly larger) quadratic orders that still have distinct images as the endomorphisms do not commute. Thus Kaneko's bound [Kan89, Theorem 2'] yields $4p_\ell \leq (-4\ell^2)^2$, i.e. $p_\ell < 4\ell^4$. In particular, by the discussion in Subsection 2.1 this proves that 0 and 1728 cannot be ℓ-isogenous for $p > 4\ell^4$.

Returning to our relation, let us suppose that we have two j-invariants $j_0, j_1 \in \overline{\mathbb{F}_p}$. As before we consider the polynomials

$$\Phi_\ell^c(X, j_0) \quad \text{and} \quad \Theta_\ell^c(X, j_1) = \Phi_\ell^c(\ell^s / X, j_1) \cdot X^{\ell+1} / \ell^s,$$

which have the same set of common solutions as the system (2) due to $p \neq \ell$; these common solutions are, moreover, precisely the roots of the polynomial

$$\Gamma_\ell(j_0, j_1) := \gcd\left(\Phi_\ell^c(X, j_0), \Theta_\ell^c(X, j_1)\right) \in \mathbb{F}_p(j_0, j_1)[X].$$

The previous results now show that the degree of $\Gamma_\ell(j_0, j_1)$ is low in most cases:

Corollary 10. Let $\ell \in \{2,3,5,7,13\}$, let $j_0, j_1 \in \overline{\mathbb{F}_p}$ for a prime $p \neq \ell$, and consider as above the gcd-polynomial $\Gamma_\ell(j_0, j_1) \in \overline{\mathbb{F}_p}[X]$. Then

$$\deg \Gamma_\ell(j_0, j_1) = \min\{\iota_\ell(j_0, j_1), \iota_\ell(j_1, j_0)\}$$

where $\iota_\ell(j, j')$ denotes the number of non-equivalent ℓ-isogenies $j \to j'$. In particular:

(a) $\deg \Gamma_\ell(j_0, j_1) \geq 1$ if and only if j_0 and j_1 are ℓ-isogenous.

(b) If j_0 is ordinary or $p > p_\ell$ (e.g. if $p \geq 4\ell^3$), then $\deg \Gamma_\ell(j_0, j_1) \leq 2$.

Proof. The main claim is a direct consequence of the Multiplicity Theorem and the multiplicity analysis that was performed in its proof, noting that the minimum is necessary to account for higher root multiplicities at j-invariants 0 and 1728 (see also Lemma 7); hence claim (a) follows immediately. Moreover, Proposition 8 and Theorem 9 imply claim (b). □

Remark 5. Since we can always factor a degree ℓ^2-isogeny into two degree ℓ-isogenies (cf. [Sil09, Corollary III.4.11]), the j-invariants j_0 for which there is some j-invariant j_1 with $\deg \Gamma_\ell(j_0, j_1) \geq 2$ correspond precisely to the j-invariants of ℓ^2-*small* curves as defined in [LB20], where the authors also prove that these j-invariants form a vanishingly small, but generally non-empty subset of $\overline{\mathbb{F}_p}$ for large p. For $\ell \in \{2,3,5,7,13\}$ we can bound the size of this set more precisely: By factoring the J_1-resultant

$$\mathrm{res}(\Phi_\ell(J_0, J_1), \tfrac{\partial}{\partial J_1}\Phi_\ell(J_0, J_1))$$

over the coefficient ring $R = \mathbb{Z}[J_0]$, we see that the degree sum of all distinct irreducible factors is at most $\ell^2 + 1$. Hence Proposition 1(a) shows that there are at most $\ell^2 + 1$ invariants $j_0 \in \overline{\mathbb{F}_p}$ that can belong to an ℓ^2-small curve for $p > \ell$.

Computing roots. Corollary 10 suggests the following efficient strategy to find a root f of the system (2) for two ℓ-isogenous j-invariants $j_0, j_1 \in \overline{\mathbb{F}_p}$: If the j-invariants do not lie in $\mathbb{F}_{p^2}$, and are hence necessarily ordinary, or if we have $p \geq 4\ell^3$ (which is guaranteed for cryptographically large primes), we can simply compute the gcd-polynomial $\Gamma_\ell(j_0, j_1)$ and obtain a root either by directly reading it off (in the degree 1 case) or by using the quadratic formula (in the rare degree 2 case). Otherwise we will see in the next section (Theorem 12) that $\Gamma_\ell(j_0, j_1) \in \mathbb{F}_{p^2}[X]$ splits into linear factors over $\mathbb{F}_{p^2}$ – this allows us to factor the polynomial over $\mathbb{F}_{p^2}$, e.g. using Berlekamp factorization. Note, however, that we do not even need a full factorization, as we are only interested in one root. Therefore we may, starting with $\Gamma_\ell(j_0, j_1)$, compute a partial factorization of our current polynomial and then only keep the factor of smallest degree for the next step, until we reach a low enough degree to solve for a root directly.

In view of the above results, we conclude that our new relation is practically equivalent to the relation $\mathcal{R}_{\ell^k\text{-ModPoly}}$, i.e. a user with knowledge of an ℓ-isogeny j-invariant chain of length k can efficiently compute the additional roots $(f_i)_{i \in [k]}$ needed to prove their knowledge with respect to the relation $\mathcal{R}_{\ell^k\text{-ModRoot}}$.

3.3 Isogeny Reconstruction and Splitting Behavior

As we have now given an essentially optimal bound on the number of roots of system (2) in large characteristic, we next investigate where these roots lie. Due to Remark 5, this question is easily answered in the overwhelming majority of cases: For a j-invariant j_0 that does not admit two non-equivalent ℓ-isogenies to the same target, the Multiplicity Theorem shows that each root f of $\Phi_\ell^c(X, j_0)$ is the unique root of the system given by $\Phi_\ell^c(X, j_0) = 0 = \Theta_\ell^c(X, j_1)$ for some $j_1 \in \overline{\mathbb{F}_p}$, and hence lies in the field extension $\mathbb{F}_p(j_0, j_1)$. In fact, as $\mathcal{J}_\ell(f) = j_0$ and $\mathcal{J}_\ell(\ell^s/f) = j_1$ by Lemma 3, we have $\mathbb{F}_p(f) = \mathbb{F}_p(j_0, j_1)$.

To analyze the splitting behavior in general, however, some additional work is required. As the following example shows, the field extension generated by the two j-invariants is not guaranteed to contain any root of system (2):

Example 2. Let $\ell = 3$, $p = 61$ and $j_0 = 9$. Then we have the factorizations

$$\Phi_3(9, j_1) = (j_1 - 9)^2 \cdot (j_1 - 41)^2 \text{ and } \Phi_3^c(X, 9) = (X^2 - 15X - 3) \cdot (X^2 - 10X + 1)$$

over $\mathbb{F}_{61}$. Further we have $\Gamma_3(9, 9) = X^2 - 15X - 3$ and $\Gamma_3(9, 41) = X^2 - 10X + 1$; these polynomials are irreducible over $\mathbb{F}_{61}$, i.e. their roots lie in $\mathbb{F}_{61^2} \setminus \mathbb{F}_{61}$.

As a remedy to this issue, the following result describes how we can reconstruct an ℓ-isogeny from a common root f of system (2) in most cases:

Reconstruction Theorem. *Let $\ell \in \{2, 3, 5, 7, 13\}$, let K be a field of characteristic $\mathrm{char}(K) \notin \{2, 3, \ell\}$, let $j_0 \in K \setminus \{0, 1728\}$ be a j-invariant and let $f \in \overline{K}^\times$ be a root of $\Phi_\ell^c(X, j_0)$. Define the parameters*

$$A = -3j_0(j_0 - 1728) \text{ and } B = -2j_0(j_0 - 1728)^2$$

and the elliptic curve

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

Then $j(E) = j_0$, and we can find a kernel polynomial $\phi_\ell(f)$ of degree $\left\lceil \frac{\ell-1}{2} \right\rceil$ that defines over $K(f)$ an ℓ-isogeny from E to a curve with j-invariant $\mathcal{J}_\ell(\ell^s/f)$. Moreover, f can be expressed as a K-rational function in the coefficients $1 = s_0, s_1, \ldots, s_n$ of $\phi_\ell(f)$, i.e.

$$K(f) = K(s_1, \ldots, s_n).$$

Proof. To limit confusion of variable names in this proof, we will write the canonical modular polynomial in the variables T and j (instead of the usual variables X and j). This proof will be highly computational – the observational claims used along the way can be verified via the Sage script `kernel_polynomials.sage`, which can be found in the accompanying GitHub repository[1].

The discriminant of the curve E is $2^{11} \cdot 3^5 \cdot j_0 \cdot (j_0 - 1728)$, which is non-zero by our assumptions. Hence E is an elliptic curve and one easily verifies that $j(E) = j_0$.

To work computationally, we will consider the coefficient ring $\mathbb{Z}[T, T^{-1}]$ of Laurent polynomials over $\mathbb{Z}$; since the root f is non-zero due to $\mathrm{char}(K) \neq \ell$, we can then apply the ring homomorphism $\mathbb{Z}[T, T^{-1}] \to K(f)$ given by evaluating T at f. By Lemma 3 we can write

$$j_0(T) = \frac{\Phi_\ell^c(T, 0)}{T} \text{ and } j_0(T) - 1728 = \frac{\Phi_\ell^c(T, 1728)}{T}.$$

Now considering the coefficients of E as elements of $\mathbb{Z}[T, T^{-1}]$, and hence considering $E = E(T)$ as a curve over $\mathbb{Q}(T)$, we can apply the following deciding trick, which is based on the ideas in [CW05] and [Tsu13, §3-4]: The ℓ^{th} division polynomial $\psi_\ell \in \mathbb{Q}(T)[x]$ of $E(T)$ has coefficients in $\mathbb{Z}[T, T^{-1}]$ and admits in $\mathbb{Z}[T, T^{-1}][x]$ a monic factor $\phi_\ell \in \mathbb{Z}[T, T^{-1}][x]$ of degree $\left\lceil \frac{\ell-1}{2} \right\rceil =: n$.

Evaluating at f hence yields the polynomial $\phi_\ell(f) \in K(f)[x]$ of degree n, and we have to show that this is a kernel polynomial. For $\ell = 2$ we do this directly: Writing $\phi_2 = x - \rho$ and plugging the root ρ into the x-coordinate of $E(T)$ yields $y^2 = 0$, so $(\rho(f), 0)$ is a 2-torsion point of $E(f)$ as desired.

To prove that $\phi_\ell(f)$ is a kernel polynomial for $\ell \geq 3$, we want to apply the *Kernel polynomial criterion* given in [Tsu13, §3.3]. First we note that $a = 2$ is a *semi-primitive root modulo* ℓ as stated on [Tsu13, p. 34], so we have to compute the action of the [2]-endomorphism of E on the x-coordinate. By the point doubling formula [Sil09, Group Law Algorithm III.2.3(d)] we have

$$[2]^*(x) = \frac{x^4 - 2A(T)x^2 - 8B(T)x^2 + A(T)^2}{4x^3 + 4A(T)x + 4B(T)} =: \frac{t_1(T)}{t_2(T)},$$

and due to our assumptions that $\mathrm{char}(K) \notin \{2, 3\}$ and $j_0 \notin \{0, 1728\}$ one can easily check with the Euclidean algorithm that $\gcd(t_1(f), t_2(f)) = 1$ in $K(f)[x]$.

Next we evaluate ϕ_ℓ (in x) at the rational function $[2]^*(x)$ to obtain

$$\phi_\ell\left(\frac{t_1(T)}{t_2(T)}\right) = \frac{1}{t_2(T)^n} \cdot \left[\phi_\ell\left(\frac{t_1(T)}{t_2(T)}\right) \cdot t_2(T)^n\right] =: \frac{1}{t_2(T)^n} \cdot h_\ell(T)$$

[1] `https://github.com/QuSAC/IsogenyPoKviaCanonicalModPolys`

where $h_\ell(f)$ is coprime to $t_2(f)^n$ in $K(f)[x]$ since $t_1(f)$ and $t_2(f)$ are coprime. Now we define the monic polynomial

$$\tau_2(\phi_\ell(f)) := \gcd(\psi_\ell(f), h_\ell(f));$$

as $\psi_\ell(f)$ is the ℓ^{th} division polynomial of $E(f)$, we see with [Tsu13, Corollary 3.3.2] that $\phi_\ell(f)$ is a kernel polynomial of an ℓ-isogeny if and only if $\tau_2(\phi_\ell(f)) = \phi_\ell(f)$.

However, by [Tsu13, Proposition 3.3.1] we also see that $\tau_2(\phi_\ell(f))$ is a monic polynomial of degree $n = \frac{\ell-1}{2}$, so it suffices to show that $\phi_\ell(f)$ divides $\tau_2(\phi_\ell(f))$. To this end, first note that $\phi_\ell(T)$ divides $\psi_\ell(T)$ over $\mathbb{Z}[T, T^{-1}]$ by construction, which allows us to deduce that $\phi_\ell(f)$ divides $\psi_\ell(f)$.

Furthermore we can check computationally that $\phi_\ell(T)$ divides $h_\ell(T)$ over $\mathbb{Z}[T, T^{-1}]$, and hence $\phi_\ell(f)$ also divides $h_\ell(f)$. Therefore the definition of $\tau_2(\phi_\ell(f))$ forces it to be divisible by $\phi_\ell(f)$, and we conclude that $\phi_\ell(f)$ is a kernel polynomial of an ℓ-isogeny defined on $E(f)$.

Penultimately, we want to show that the isogeny defined by $\phi_\ell(f)$ maps to a curve of j-invariant $\mathcal{J}_\ell(\ell^s/f)$. This can be deduced directly from *Vélu's formulas* – more precisely, in [Koh96, §2.4] Kohel describes the target curve in terms of A, B and the coefficients of $\phi_\ell(f)$. Applying these formulas, we see that the target curve is in short Weierstrass form with discriminant $\Delta = 2^{11} \cdot 3^5 \cdot f^{\ell-1} \cdot A \cdot B \neq 0$ and j-invariant $\mathcal{J}_\ell(\ell^s/f)$.

Finally, we refer to Appendix A.4 for the (ℓ-dependent) expressions of f as a K-rational function in the coefficients of $\phi_\ell(f)$; here we only note that, as a kernel polynomial, $\phi_\ell(f)$ will never have any double roots, so $\mathrm{disc}(\phi_\ell(f)) \neq 0$ by Proposition 1(b) (in view of Remark 1). □

Remark 6. Note that we can also reconstruct ℓ-isogenies from a root $f \in \overline{K}^\times$ of $\Phi_\ell^c(X, j_0)$ if we have $j_1 = \mathcal{J}_\ell(\ell^s/f) \notin \{0, 1728\}$. Indeed, in this case we first compute the dual isogeny (up to equivalence) by applying the above techniques to ℓ^s/f, and then take its dual and precompose with the different automorphisms at j_0 to obtain the non-equivalent ℓ-isogenies $j_0 \to j_1$ corresponding to f.

To now analyze the splitting behavior of $\Phi_\ell^c(X, j_0)$ for supersingular j-invariants, we still need to handle the special j-invariants:

Proposition 11. Let $\ell \in \{2, 3, 5, 7, 13\}$, let $p \neq \ell$ be a prime and let moreover $j^* \in \{0, 1728\} \subseteq \mathbb{F}_{p^2}$ be supersingular. Then $\Phi_\ell^c(X, j^*)$ splits over $\mathbb{F}_{p^2}$.

Proof. We assume $p > p_\ell$, all other cases can be checked directly. By Theorem 9(b-c) and Corollary 10, we see that $\Gamma_\ell(j^*, j_1) \in \mathbb{F}_{p^2}[X]$ has degree at most 1 for any supersingular $j_1 \neq j^*$, so the only roots of $\Phi_\ell^c(X, j^*)$ that may not lie in $\mathbb{F}_{p^2}$ are the roots of $\Gamma_\ell(j^*, j^*)$. However, by Corollary 10(b) this is a polynomial of degree at most 2 over $\mathbb{F}_p$. □

The previous results give us the splitting behavior of $\Phi_\ell^c(X, j_0)$ for a supersingular j-invariant j_0:

Theorem 12. *Let $\ell \in \{2, 3, 5, 7, 13\}$, let $p \neq \ell$ be a prime and let $j_0 \in \mathbb{F}_{p^2}$ be a supersingular j-invariant. Then $\Phi_\ell^c(X, j_0)$ splits over $\mathbb{F}_{p^2}$.*

Proof. Due to Proposition 11 we may assume $j_0 \notin \{0, 1728\}$ and, in particular, $p \geq 5$. Now let $f \in \overline{\mathbb{F}_p}$ be a root of $\Phi_\ell^c(X, j_0)$. With the Reconstruction Theorem we can then associate to f the kernel polynomial $\phi_\ell(f)$ of an ℓ-isogeny defined on a curve E over $\mathbb{F}_{p^2}$.

Moreover, the kernel of this isogeny is invariant under the action of the p^2-Frobenius isomorphism of $\overline{\mathbb{F}_p}$ on E. Indeed, this action is given by evaluation of the p^2-Frobenius endomorphism π of E; due to [AAM19, §4] we further see that π has trace $\pm 2p$ since $j_0 \notin \{0, 1728\}$, so it acts on E via scalar multiplication by $\pm p$ (cf. [AAM19, §5]) and we conclude that any subgroup of $E(\overline{\mathbb{F}_p})$ is invariant under the action of π.

Therefore the coefficients of the kernel polynomial $\phi_\ell(f)$ lie in $\mathbb{F}_{p^2}$ as well, and with the second part of the Reconstruction Theorem we deduce $f \in \mathbb{F}_{p^2}$ as desired. □

We give an application, which can alternatively be proven by showing that any Legendre parameter of a supersingular j-invariant lies in $\mathbb{F}_{p^2}$ [AT02, Proposition 2.2].

Corollary 13. Let p be a prime and let $j_0 \in \mathbb{F}_{p^2}$ be a supersingular j-invariant. Then $j_0 - 1728$ is a square in $\mathbb{F}_{p^2}$.

Proof. As all elements of $\mathbb{F}_p$ are squares in $\mathbb{F}_{p^2}$, we may assume $j_0 \notin \{0, 1728\}$ and, in particular, $p \geq 5$. Due to Proposition 1(a) we can thus compute the discriminant of $\Phi_2^c(X, j_0)$ by first computing the discriminant of $\Phi_2^c(X, J_0)$ over $\mathbb{Z}[J_0]$ and then reducing modulo p and evaluating at j_0, which yields

$$\mathrm{disc}(\Phi_2^c(X, j_0)) = 2^2 \cdot j_0^2 \cdot (j_0 - 1728).$$

Now $\Phi_2^c(X, j_0)$ does not have multiple roots by Lemma 6 and splits into three linear factors over $\mathbb{F}_{p^2}$ by Theorem 12, so [Gow90, Theorem 1.8] shows that $\mathrm{disc}(\Phi_2^c(X, j_0))$ must be a square in $\mathbb{F}_{p^2}$, and the claim follows. □

4 Proving Isogeny Knowledge via R1CS

In the previous section we have laid all the theoretical foundations for describing the R1CS that will enable us to build an efficient proof of knowledge for an ℓ-isogeny walk with k steps. Before we describe our approach based on canonical modular polynomials, we briefly revisit the strategy pursued in [CLL23] for the prime $\ell = 2$.

4.1 Revisiting the Approach in [CLL23]

The authors of [CLL23] use the classical modular polynomial $\Phi_2(X, Y)$ to construct an R1CS to prove knowledge of a degree 2^k isogeny with respect to the relation $\mathcal{R}_{\ell^k\text{-ModPoly}}$ (Eq. (1)). They do this by finding an efficient arithmetization to prove that $\Phi_2(j_i, j_{i+1}) = 0$ for a chain of $k+1$ successive j-invariants. Here $j_0 = j(E_0)$ and $j_k = j(E_k)$ are part of the statement, and j_i for $0 < i < k$ are part of the R1CS witness. We can recover the original isogeny by searching at each step for ℓ-isogenous elliptic curves E_i, E_{i+1} where $j(E_i) = j_i$ and $j(E_{i+1}) = j_{i+1}$. On the other hand, such a chain of j-invariants can be found for any degree 2^k isogeny by iteratively computing degree 2 isogenies using kernel points. This means that the problem of finding such a chain of j-invariants is equivalent to finding an explicit isogeny $E_0 \to E_k$.

To arithmetize the authors express each step of the isogeny walk as an R1CS gadget, which is then employed for each link in the chain. Two tricks are used to optimize:

- The values j_i, j_i^2 and j_i^3 are computed for all $i \in \{0, \ldots, k\}$, as well as $j_i j_{i+1}$ for each $i < k$. The condition that $\Phi_2(j_i, j_{i+1}) = 0$ can then be expressed as a single R1CS constraint.
- To express the gadget over $\mathbb{F}_p$ as well as over $\mathbb{F}_{p^2}$, the authors use arithmetizations for products and squares that are more efficient than the naive approach of computing each cross term individually. In particular, the product $(x_1 + x_2\alpha)(y_1 + y_1\alpha)$, with $x_1, x_2, y_1, y_2 \in \mathbb{F}_p$ and $\alpha^2 = d$ some non-square residue in $\mathbb{F}_p$, can be expressed in three products over $\mathbb{F}_p$. Squarings can be expressed in two products.

Our goal is now twofold: First, to further optimize the arithmetization for $\ell = 2$. Second, to construct efficient R1CS for more primes $\ell > 2$, more specifically for the primes $\ell \in \{2, 3, 5, 7, 13\}$, for which we have developed a good understanding of canonical modular polynomials in the previous section.

By Lemma 3, for $\ell \in \{2, 3, 5, 7, 13\}$ the ℓ^{th} canonical modular polynomial has the form

$$\Phi_\ell^c(X, j) = X^{\ell+1} + \textstyle\sum_{i=1}^{\ell} c_i X^i + \ell^s - X \cdot j.$$

In what follows we will write $c_0 = \ell^s$ and $c_{\ell+1} = 1$. In view of the Multiplicity Theorem the proof of knowledge with respect to relation $\mathcal{R}_{\ell^k\text{-ModRoot}}$ can be encoded step-wise via the system of equations (2). Multiplying the second equation $\Phi_\ell^c(\ell^s/X, j_1) = 0$ by $X^{\ell+1}/\ell^s$ to obtain $\Theta_\ell^c(X, j_1)$ as before, we obtain the equivalent system (where $c'_i = c_{\ell+1-i} \cdot \ell^{s(\ell-i)}$):

$$\textstyle\sum_{i=0}^{\ell+1} c_i X^i - j_0 \cdot X = 0 \quad \wedge \quad \sum_{i=0}^{\ell+1} c'_i X^i - j_1 \cdot X^\ell = 0. \tag{4}$$

We will reformulate these equations as an R1CS in the upcoming subsection. In our applications we consider supersingular j-invariants, which are known to be contained in $\mathbb{F}_{p^2}$. It is crucial for the effectivity of our method that in this situation the roots of the above equations still lie in the quadratic extension $\mathbb{F}_{p^2}$ of $\mathbb{F}_p$, rather than in a larger extension, as proven in Theorem 12.

4.2 Reformulation as an R1CS

Compared to the classical modular polynomials, the canonical modular polynomials have a structure that lends itself better to an R1CS for three reasons: First, the total degree of the polynomials is lower, going from a single polynomial of total degree 2ℓ to two polynomials of degree $\ell+1$. Second, whereas Φ_ℓ is very dense, Φ_ℓ^c and Θ_ℓ^c are both polynomials in just X in addition to a single term containing j. Hence there are fewer terms to produce in the R1CS. Lastly, the structure described in Lemma 7 allows us to factor part of this polynomial as a square, improving arithmetization over $\mathbb{F}_p$.

We can compute the powers $1, X, X^2, \ldots, X^\ell$ together with the j-invariants j and j' and rewrite the equations as

$$X \cdot \left(\sum_{i=0}^{\ell} c_{i+1} X^i - j \right) + c_0 = 0, \tag{5}$$

$$X^\ell \cdot \left(\sum_{i=0}^{1} c'_{\ell+i} X^i - j' \right) + \sum_{i=0}^{\ell-1} c'_i X^i = 0. \tag{6}$$

To reduce the amount of non-zero entries, we employ a change of variables and have the prover supply $y = j - c_1$ instead of j and $y' = j' - c_1$ instead of j'. This eliminates the term X from the first equation and the term X^ℓ from the second equation, since $c_1 = c'_\ell$. Clearly knowledge of a chain of j-invariants is equivalent to knowledge of a chain of y's.

These equations are expressed as an R1CS as follows. The assignment vector z has the form $z = (1\ X\ X^2\ \ldots\ X^\ell\ y\ y')^T$, and the corresponding constraint matrices are given by

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & & & \vdots & & & \vdots \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & & & \vdots & & & \vdots \\ 0 & c_2 & c_3 & \cdots & c_{\ell+1} & -1 & 0 \\ 0 & c'_{\ell+1} & 0 & \cdots & 0 & 0 & -1 \end{pmatrix},$$

and

$$C = \begin{pmatrix} 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & \\ \vdots & & & & \vdots & & & \vdots & \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 & 0 & \\ -c_0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ -c'_0 & -c'_1 & -c'_2 & -c'_3 & \cdots & -c'_{\ell-1} & 0 & 0 & 0 \end{pmatrix}.$$

For an isogeny path of length k, each new step introduces a new value y_{i+1} together with the ℓ powers of the current solution X_i. Moreover, we need $\ell-1$ more constraints for checking the powers of the new variable X_i, together with two more constraints which encode the two equations (4). There are $\ell+1$, $2\ell+2$ and 2ℓ non-zero entries in A, B and C, respectively. This means that we can arithmetize a degree ℓ^k isogeny for $\ell \in \{2,3,5,7,13\}$ in $(\ell+1)k+1$ variables, $(\ell+1)k$ constraints and $(5\ell+3)k$ non-zero entries in the R1CS.

In fact, we can do better for $\ell \in \{7,13\}$, in the following way. We let $t = \frac{\ell+1}{2}$ and rewrite the two equations as

$$X^t \cdot \left(\sum_{i=0}^{t} c_{i+t} X^i \right) + \sum_{i=0, i\neq 1}^{t-1} c_i X^i - yX \quad = \quad 0, \tag{7}$$

$$X^t \cdot \left(\sum_{i=0, i\neq t-1}^{t} c'_{i+t} X^i - y' X^{t-1} \right) + \sum_{i=0}^{t-1} c'_i X^i \quad = \quad 0. \tag{8}$$

This way we need to compute the variables $X^2, \ldots, X^t, yX, y'X^{t-1}$, and we have $t+1+2$ constraints ($t+1$ consistency checks, and the two above equations). We have t variables for the powers of X and two for yX and $y'X^{t-1}$, as well as one per j-invariant, through y. This gives a total of $(t+3)k+1$ variables. The $t+1$ consistency checks can be computed in $3t+3$, while the two other equations require $4t+4$ non-zero entries in the constraint matrices. In total, this means $(t+3)k+1$ variables, $(t+3)k$ constraints and $(7t+7)k$ non-zero entries in the R1CS to arithmetize a walk of length k for $\ell \in \{3,5,7,13\}$.

4.3 Lifting to $\mathbb{F}_p \times \mathbb{F}_p$

So far, we have described efficient arithmetizations for proving isogenies that work over $\mathbb{F}_{p^2}$. While it is perfectly valid to use an R1CS over this field, this is not supported by all proof systems. Fortunately, we can lift any arithmetization over $\mathbb{F}_{p^2}$ to $\mathbb{F}_p \times \mathbb{F}_p$.

We represent elements in $\mathbb{F}_{p^2}$ as $x + y\alpha$ with $\alpha^2 = d$ for some quadratic non-residue $d \in \mathbb{F}_p$. Linear operations over $\mathbb{F}_{p^2}$ translate directly to linear operations over $\mathbb{F}_p$, whereas a multiplication $(x_1 + x_2\alpha)(y_1 + y_2\alpha) = z_1 + z_2\alpha$ would naively induce four multiplications for all cross terms x_1y_1, $x_1y_2\alpha$, $x_2y_1\alpha$ and x_2y_2d. As noted in [CLL23], there exist well-known techniques to do this more efficiently. Multiplication can be performed using one auxiliary variable and three constraints:

$$\begin{aligned} u &= x_2y_2, \\ z_1 - du &= x_1y_1, \\ z_1 + z_2 + (1-d)u &= (x_1+x_2)(y_1+y_2). \end{aligned}$$

This immediately implies an upper bound for the cost of lifting: We can substitute this relation for every row of the original R1CS to obtain a new system with $m' = 3m$. The amount of variables becomes $n' = 2n + m$, since variables now have two components and we add one intermediate variable for every constraint. The number of non-zero entries for this new system is $\mathrm{nnz}' \le 4\,\mathrm{nnz}$ due to the doubling of the number of variables, and because all variables are used at most twice in the above system of equations.

Squaring a variable can be performed more efficiently. For

$$z_1 + z_2\alpha = (x_1 + x_2\alpha)^2,$$

the following system of equations suffices:

$$\begin{aligned} z_2 &= 2x_1x_2, \\ z_1 + (d+1)/2z_2 &= (x_1+x_2)(x_1+dx_2). \end{aligned}$$

This uses just two constraints and nine non-zero entries, requiring no auxiliary variables.

A more efficient basis. Given these gadgets for multiplying and squaring, there is an optimization we can perform to decrease the number of non-zero entries in the constraint matrices. The basis change $(x_1, x_2) \to (x_1, x_1 + x_2) = (x_1, x_s)$ saves one non-zero entry in the last constraint for the general product. This might seem minor, but since both coefficients of x can itself be a linear combination of many variables, this is significant for our system. Unfortunately the same technique does not give a saving for squaring. The right-hand side does not incur a cost by expressing $x \in \mathbb{F}_{p^2}$ as x_1 and x_s, but the left hand side does incur a cost of expressing $z \in \mathbb{F}_{p^2}$ as z_1 and z_s. Hence we have two options: Incur this cost here and potentially save it if the result is used as the result of a multiplication, or not incur this cost. Although this sum basis is almost always at least as efficient as the regular basis, one can simply choose the most efficient basis for each variable individually.

Exploiting Φ_ℓ^c's structure. The above gadgets and the sum basis help us to express the canonical modular polynomials efficiently in an R1CS over $\mathbb{F}_p \times \mathbb{F}_p$. Furthermore, for some ℓ, the canonical polynomial can be rewritten using the structure described in Lemma 7 such that it utilizes more squares. This in turn minimizes the number of necessary constraints. For $\ell = 2$ over $\mathbb{F}_p \times \mathbb{F}_p$, we can write the entire system using two squares and one product. First, define $y = c_1 - c_2^2 \cdot (4c_3)^{-1} - j$ (and y' accordingly). Then

$$\begin{aligned} c_3\left(c_2 \cdot (2c_3)^{-1} + X\right)^2 + y + c_0X^{-1} &= 0, \\ c_3\ell^{3s}\left(c_2\ell^s \cdot (2c_3)^{-1} + X^{-1}\right)^2 + \ell^s y' + c_0X &= 0, \end{aligned}$$

captures both polynomials in two squaring relations. We only need to compute the inverse X^{-1} with a single additional multiplication. Over $\mathbb{F}_p \times \mathbb{F}_p$, the resulting system has $7k$ constraints and $7k+2$ variables. In a similar fashion, for $\ell = 7$, we can use the fact that Equation (7) can be written as a square plus the term $\tilde{y} = 1728 - j$. For $\ell = 13$, we can express 9 coefficients correctly using the square of a degree 7 polynomial, after which we only need to correct for the lowest degree terms.

Computing powers of X. To compute even powers, we can square directly. It is however also possible to use a squaring for odd powers, through the relation

$$b\left(a \cdot (2b)^{-1} X^i + X^{i+1}\right)^2 = a^2 \cdot (4b)^{-1} X^{2i} + aX^{2i+1} + bX^{2i+2}. \tag{9}$$

We see that we obtain a linear combination of three powers, where we can freely choose $a \in \mathbb{F}_p$ and $b \in \mathbb{F}_p^\times$. By subtracting away $a^2 \cdot (4b)^{-1} X^{2i}$ and bX^{2i+2}, we obtain a constraint for the odd power X^{2i+1}. Unfortunately, this method cannot be used for the products with y or the highest power of X, i.e. X^ℓ or X^t, since here we cannot compensate for the even powers appearing on the right-hand side.

Change of variables. One disadvantage of the above method is that it increases the number of non-zero entries in the constraint matrices. To remedy this, we note that some powers of X are only used in linear combinations with other powers, i.e. in a polynomial where all three powers are already present. As such, we make a change of variables and store the right-hand side of Eq. (9) directly instead of X^{2i+1}. We can then choose a and b appropriately such that they agree with the coefficients of one of the polynomials. For example, for $\ell = 5$ we define $Z = c_5(c_4 \cdot (2c_5)^{-1} X + X^2)^2$ and rewrite Eq. (5) as

$$X \cdot \left[c_2 X + \left(c_3 - c_4^2 \cdot (4c_5)^{-1}\right) X^2 + Z + X^5 - y\right] + c_0 = 0. \tag{10}$$

The advantage is twofold: we do not need to subtract powers from Eq. (9) when computing X^3, and we do not have to add an X^4 term to Equation (10). We should only use this substitution for powers that are not necessary to compute higher powers: the above would be inefficient if we also required the value of X^3 to express $(X^3)^2 = X^6$. Concretely, we use this trick to replace X^3 for $\ell \in \{5, 7\}$ and X^5 for $\ell = 13$.

More generally, the number of non-zero entries can often be minimized through a change of variables. For example, since jX and $j'X^{t-1}$ are both only used once, we can instead already add the terms of the linear combinations in which they will be used later. This is advantageous, since a linear combination uses fewer nnz in the outcome of a square than in its input, and is cheaper still in the outcome of a multiplication. This way, the intermediate variables that are unavoidable can be used as efficiently as possible.

5 Evaluation

We provide `constraints.sage` in the accompanying GitHub repository[2] which expresses and verifies all arithmetizations and automatically counts the number of constraints, variables and nnz. These can be found in Table 2. To additionally compare results for distinct ℓ, we normalize by considering a security level λ, such that $\ell^k > 2^\lambda$. By increasing ℓ we can decrease k, reducing the number of constraints necessary for the relation $\mathcal{R}_{\text{R1CS}}$. These results can be found in Table 3. We achieve significant improvements everywhere, ranging between $25\% - 45\%$ for the number of constraints and $27\% - 48\%$ for nnz, making our arithmetizations suitable for ℓ-power isogenies for each $\ell \in \{2, 3, 5, 7, 13\}$, as well as mixed power isogenies.

In particular, the number of non-zero entries in our system is 38% smaller for $\ell = 2$ over $\mathbb{F}_{p^2}$ and 48% over $\mathbb{F}_p$. This is relevant for proof systems such as Aurora, where the prover runs in time $O(\text{nnz} \log(\text{nnz}))$, the verifier in time $O(\text{nnz})$ and the proof size is $O(\log^2(\text{nnz}))$. Over $\mathbb{F}_p$ we therefore expect the prover and verifier time to be roughly halved. For Ligero, the proof size is $O(\sqrt{\text{nnz}})$, which is thus expected to shrink by 30%. The proof size, verifier and prover times of Limbo are linear in the number of constraints m, which decreases in terms of λ as we move to higher primes. In particular, for $\ell = 7$ we achieve a 38% improvement in m over $\mathbb{F}_{p^2}$ and a 45% improvement over $\mathbb{F}_p$. We expect this improvement to translate directly to its concrete efficiency.

[2] `https://github.com/QuSAC/IsogenyPoKviaCanonicalModPolys`

ℓ	Field	m		n		nnz	
		[CLL23]	Ours	[CLL23]	Ours	[CLL23]	Ours
2	$\mathbb{F}_{p^2}$	$4k+2$	$3k$	$4k+3$	$3k+1$	$21k+6$	$13k$
3			$4k$		$4k+1$		$18k$
5			$6k$		$6k+1$		$28k$
7			$7k$		$7k+1$		$35k$
13			$10k$		$10k+1$		$56k$
2	$\mathbb{F}_p$	$11k+5$	$7k$	$11k+7$	$7k+2$	$79k+23$	$41k$
3			$11k$		$11k+2$		$65k$
5			$15k$		$15k+2$		$97k$
7			$17k$		$17k+2$		$123k$
13			$24k$		$24k+2$		$194k$

Table 2: Our results compared to [CLL23]. We consider the number of constraints m, the number of variables n and the number of non-zero entries in the constraint matrices nnz.

ℓ	Field	m		n		nnz	
		[CLL23]	Ours	[CLL23]	Ours	[CLL23]	Ours
2	$\mathbb{F}_{p^2}$	$4\lambda+2$	3λ	$4\lambda+3$	$3\lambda+1$	$21\lambda+6$	13λ
3			2.524λ		$2.524\lambda+1$		11.357λ
5			2.584λ		$2.584\lambda+1$		12.059λ
7			2.493λ		$2.493\lambda+1$		12.467λ
13			2.702λ		$2.702\lambda+1$		15.133λ
2	$\mathbb{F}_p$	$11\lambda+5$	7λ	$11\lambda+7$	$7\lambda+2$	$79\lambda+23$	41λ
3			6.940λ		$6.940\lambda+2$		41.010λ
5			6.460λ		$6.460\lambda+2$		41.776λ
7			6.056λ		$6.056\lambda+2$		43.813λ
13			6.486λ		$6.486\lambda+2$		52.426λ

Table 3: Our results compared to those of [CLL23], normalized for security parameter λ.

6 Conclusion and Open Problems

In this paper we improved on the state-of-the-art of using general-purpose zero-knowledge proof systems for proving knowledge of an isogeny via R1CS. We were able to generalize the approach of Cong, Lai and Levin [CLL23] beyond $\ell = 2$ to prime numbers $\ell \in \{3, 5, 7, 13\}$ via the use of canonical modular polynomials. Moreover, we optimized the arithmetizations for the corresponding relation both over $\mathbb{F}_{p^2}$ and over $\mathbb{F}_p \times \mathbb{F}_p$.

In the course of our work we encountered interesting mathematical questions, some of which might hold in greater generality. For example, while Remark 4 argues that one can generalize Theorem 9 to any prime ℓ, the growth trend displayed in Table 1 suggests that even tighter bounds on the prime p_ℓ, such as $2\log_2(\ell)\ell^3$, could be achievable.

It might be even more interesting to study the canonical modular polynomials (or different, equivalent polynomials) for primes ℓ such that $\kappa > 1$. In that case we do not know whether the Multiplicity Theorem still holds true. More precisely, we expect one inequality to still hold, but the other to fail generally – see Remark 2.

Therefore the mathematical contributions in this paper might motivate deeper studies in the future.

Acknowledgements

The last author would like to thank Jonathan Love, Eli Orvis and Travis Morrison for interesting discussions on the supersingular isogeny graph, which led to Remark 4, at the Leuven Isogeny Days 5.

References

[AAM19] Gora Adj, Omran Ahmadi, and Alfred Menezes. On isogeny graphs of supersingular elliptic curves over finite fields. *Finite Fields Their Appl.*, 55:268–283, 2019.

[AGRS24] Behzad Abdolmaleki, Noemi Glaeser, Sebastian Ramacher, and Daniel Slamanig. Circuit-succinct universally composable nizks with updatable crs. In *2024 IEEE 37th Computer Security Foundations Symposium (CSF)*, pages 527–542, Los Alamitos, CA, USA, jul 2024. IEEE Computer Society.

[AHIV17] Scott Ames, Carmit Hazay, Yuval Ishai, and Muthuramakrishnan Venkitasubramaniam. Ligero: Lightweight sublinear arguments without a trusted setup. In Bhavani M. Thuraisingham, David Evans, Tal Malkin, and Dongyan Xu, editors, *ACM CCS 2017: 24th Conference on Computer and Communications Security*, pages 2087–2104, Dallas, TX, USA, October 31 – November 2, 2017. ACM Press.

[AT02] Roland Auer and Jaap Top. Legendre elliptic curves over finite fields. *Journal of Number Theory*, 95(2):303–312, 2002.

[Bas24] Andrea Basso. POKE: A framework for efficient PKEs, split KEMs, and OPRFs from higher-dimensional isogenies. Cryptology ePrint Archive, Paper 2024/624, 2024.

[BBBF18] Dan Boneh, Joseph Bonneau, Benedikt Bünz, and Ben Fisch. Verifiable delay functions. In Hovav Shacham and Alexandra Boldyreva, editors, *Advances in Cryptology – CRYPTO 2018, Part I*, volume 10991 of *Lecture Notes in Computer Science*, pages 757–788, Santa Barbara, CA, USA, August 19–23, 2018. Springer, Cham, Switzerland.

[BCC+23] Andrea Basso, Giulio Codogni, Deirdre Connolly, Luca De Feo, Tako Boris Fouotsa, Guido Maria Lido, Travis Morrison, Lorenz Panny, Sikhar Patranabis, and Benjamin Wesolowski. Supersingular curves you can trust. In Carmit Hazay and Martijn Stam, editors, *Advances in Cryptology – EUROCRYPT 2023, Part II*, volume 14005 of *Lecture Notes in Computer Science*, pages 405–437, Lyon, France, April 23–27, 2023. Springer, Cham, Switzerland.

[BCNE+18] Efrat Bank, Catalina Camacho-Navarro, Kirsten Eisenträger, Travis Morrison, and Jennifer Park. Cycles in the supersingular ℓ-isogeny graph and corresponding endomorphisms, September 2018.

[BCR+19] Eli Ben-Sasson, Alessandro Chiesa, Michael Riabzev, Nicholas Spooner, Madars Virza, and Nicholas P. Ward. Aurora: Transparent succinct arguments for R1CS. In Yuval Ishai and Vincent Rijmen, editors, *Advances in Cryptology – EUROCRYPT 2019, Part I*, volume 11476 of *Lecture Notes in Computer Science*, pages 103–128, Darmstadt, Germany, May 19–23, 2019. Springer, Cham, Switzerland.

[BD21] Jeffrey Burdges and Luca De Feo. Delay encryption. In Anne Canteaut and François-Xavier Standaert, editors, *Advances in Cryptology – EUROCRYPT 2021, Part I*, volume 12696 of *Lecture Notes in Computer Science*, pages 302–326, Zagreb, Croatia, October 17–21, 2021. Springer, Cham, Switzerland.

[BFGP23] Ward Beullens, Luca De Feo, Steven D. Galbraith, and Christophe Petit. Proving knowledge of isogenies: a survey. *Des. Codes Cryptogr.*, 91(11):3425–3456, 2023.

[BKV19] Ward Beullens, Thorsten Kleinjung, and Frederik Vercauteren. CSI-FiSh: Efficient isogeny based signatures through class group computations. In Steven D. Galbraith and Shiho Moriai, editors, *Advances in Cryptology – ASIACRYPT 2019, Part I*, volume 11921 of *Lecture Notes in Computer Science*, pages 227–247, Kobe, Japan, December 8–12, 2019. Springer, Cham, Switzerland.

[Bos18] Siegfried Bosch. *Algebra. From the Viewpoint of Galois Theory*. Birkhäuser Advanced Texts Basler Lehrbücher. Birkhäuser Cham, 2018.

[BPR06] Saugata Basu, Richard Pollack, and Marie-Françoise Roy. *Algorithms in Real Algebraic Geometry*, volume 10 of *Algorithms and Computation in Mathematics*. Springer Berlin Heidelberg, 2nd ed. edition, 2006.

[BS23] Ward Beullens and Gregor Seiler. LaBRADOR: Compact proofs for R1CS from module-SIS. In Helena Handschuh and Anna Lysyanskaya, editors, *Advances in Cryptology – CRYPTO 2023, Part V*, volume 14085 of *Lecture Notes in Computer Science*, pages 518–548, Santa Barbara, CA, USA, August 20–24, 2023. Springer, Cham, Switzerland.

[CFA+06] Henri Cohen, Gerhard Frey, Robero M. Avanzi, Christophe Doche, Tanja Lange, Kim Nguyen, and Frederik Vercauteren. *Handbook of elliptic and hyperelliptic curve cryptography.* Discrete Math. Appl. (Boca Raton). Boca Raton, FL: Chapman & Hall/CRC, 2006.

[CL24] Giulio Codogni and Guido Lido. Spectral theory of isogeny graphs. 2024.

[CLG09] Denis X. Charles, Kristin E. Lauter, and Eyal Z. Goren. Cryptographic hash functions from expander graphs. *Journal of Cryptology*, 22:93–113, 2009.

[CLL23] Kelong Cong, Yi-Fu Lai, and Shai Levin. Efficient isogeny proofs using generic techniques. In Mehdi Tibouchi and Xiaofeng Wang, editors, *ACNS 23: 21st International Conference on Applied Cryptography and Network Security, Part II*, volume 13906 of *Lecture Notes in Computer Science*, pages 248–275, Kyoto, Japan, June 19–22, 2023. Springer, Cham, Switzerland.

[CLM+18] Wouter Castryck, Tanja Lange, Chloe Martindale, Lorenz Panny, and Joost Renes. CSIDH: An efficient post-quantum commutative group action. In Thomas Peyrin and Steven Galbraith, editors, *Advances in Cryptology – ASIACRYPT 2018, Part III*, volume 11274 of *Lecture Notes in Computer Science*, pages 395–427, Brisbane, Queensland, Australia, December 2–6, 2018. Springer, Cham, Switzerland.

[Cou06] Jean-Marc Couveignes. Hard homogeneous spaces. Cryptology ePrint Archive, Report 2006/291, 2006.

[Cox13] David A. Cox. *Primes of the Form $x^2 + ny^2$. Fermat, Class Field Theory, and Complex Multiplication*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. John Wiley & Sons, 2nd edition, 2013.

[CSF+08] D. Cooper, S. Santesson, S. Farrell, S. Boeyen, R. Housley, and W. Polk. Internet x.509 public key infrastructure certificate and certificate revocation list (crl) profile. RFC 5280, RFC Editor, May 2008. http://www.rfc-editor.org/rfc/rfc5280.txt.

[CSRT22] Jorge Chávez-Saab, Francisco Rodríguez-Henríquez, and Mehdi Tibouchi. Verifiable isogeny walks: Towards an isogeny-based postquantum VDF. In Riham AlTawy and Andreas Hülsing, editors, *SAC 2021: 28th Annual International Workshop on Selected Areas in Cryptography*, volume 13203 of *Lecture Notes in Computer Science*, pages 441–460, Virtual Event, September 29 – October 1, 2022. Springer, Cham, Switzerland.

[CW05] John Cremona and Mark Watkins. Computing isogenies of elliptic curves. http://magma.maths.usyd.edu.au/users/watkins/papers/isogs.ps, 2005. preprint.

[DDGZ22] Luca De Feo, Samuel Dobson, Steven D. Galbraith, and Lukas Zobernig. SIDH proof of knowledge. In Shweta Agrawal and Dongdai Lin, editors, *Advances in Cryptology – ASIACRYPT 2022, Part II*, volume 13792 of *Lecture Notes in Computer Science*, pages 310–339, Taipei, Taiwan, December 5–9, 2022. Springer, Cham, Switzerland.

[DFJP14] Luca De Feo, David Jao, and Jérôme Plût. Towards quantum-resistant cryptosystems from supersingular elliptic curve isogenies. *J. Math. Cryptol.*, 8(3):209–247, 2014.

[DFV24] Max Duparc, Tako Boris Fouotsa, and Serge Vaudenay. SILBE: an updatable public key encryption scheme from lollipop attacks. Cryptology ePrint Archive, Paper 2024/400, 2024. Accepted at SAC 2024.

[DKL+20] Luca De Feo, David Kohel, Antonin Leroux, Christophe Petit, and Benjamin Wesolowski. SQISign: Compact post-quantum signatures from quaternions and isogenies. In Shiho Moriai and Huaxiong Wang, editors, *Advances in Cryptology – ASIACRYPT 2020, Part I*, volume 12491 of *Lecture Notes in Computer Science*, pages 64–93, Daejeon, South Korea, December 7–11, 2020. Springer, Cham, Switzerland.

[DMPS19] Luca De Feo, Simon Masson, Christophe Petit, and Antonio Sanso. Verifiable delay functions from supersingular isogenies and pairings. In Steven D. Galbraith and Shiho Moriai, editors, *Advances in Cryptology – ASIACRYPT 2019, Part I*, volume 11921 of *Lecture Notes in Computer Science*, pages 248–277, Kobe, Japan, December 8–12, 2019. Springer, Cham, Switzerland.

[DOT21] Cyprien Delpech de Saint Guilhem, Emmanuela Orsini, and Titouan Tanguy. Limbo: Efficient zero-knowledge MPCitH-based arguments. In Giovanni Vigna and Elaine Shi, editors, *ACM CCS 2021: 28th Conference on Computer and Communications Security*, pages 3022–3036, Virtual Event, Republic of Korea, November 15–19, 2021. ACM Press.

[Elk98] Noam D. Elkies. Elliptic and modular curves over finite fields and related computational issues. In D. A. Buell and J. T. Teitelbaum, editors, *Computational Perspectives on Number Theory*, volume 7 of *Studies in Advanced Mathematics*, pages 21–76. American Mathematical Society (AMS), 1998.

[FMP23] Tako Boris Fouotsa, Tomoki Moriya, and Christophe Petit. M-SIDH and MD-SIDH: Countering SIDH attacks by masking information. In Carmit Hazay and Martijn Stam, editors, *Advances in Cryptology – EUROCRYPT 2023, Part V*, volume 14008 of *Lecture Notes in Computer Science*, pages 282–309, Lyon, France, April 23–27, 2023. Springer, Cham, Switzerland.

[Fri11] Robert Fricke. *Die elliptischen Funktionen und ihre Anwendungen: Zweiter Teil: Die algebraischen Ausführungen*. Life Science and Basic Disciplines (German Language). Springer Berlin, Heidelberg, 1st edition, 2011. Originally published by B.G. Teubner-Verlag. (1922).

[GHL+22] Tim Güneysu, Philip W. Hodges, Georg Land, Mike Ounsworth, Douglas Stebila, and Greg Zaverucha. Proof-of-possession for KEM certificates using verifiable generation. In Heng Yin, Angelos Stavrou, Cas Cremers, and Elaine Shi, editors, *ACM CCS 2022: 29th Conference on Computer and Communications Security*, pages 1337–1351, Los Angeles, CA, USA, November 7–11, 2022. ACM Press.

[GKM+18] Jens Groth, Markulf Kohlweiss, Mary Maller, Sarah Meiklejohn, and Ian Miers. Updatable and universal common reference strings with applications to zk-SNARKs. In Hovav Shacham and Alexandra Boldyreva, editors, *Advances in Cryptology – CRYPTO 2018, Part III*, volume 10993 of *Lecture Notes in Computer Science*, pages 698–728, Santa Barbara, CA, USA, August 19–23, 2018. Springer, Cham, Switzerland.

[Gow90] Roderick Gow. Some properties and uses of the discriminant of a polynomial. In Ray Ryan, Ted Hurley, and Phil Rippon, editors, *Irish Mathematical Society Bulletin*, volume 24, pages 12–19. Irish Mathematical Society, March 1990.

[GPS17] Steven D. Galbraith, Christophe Petit, and Javier Silva. Identification protocols and signature schemes based on supersingular isogeny problems. In Tsuyoshi Takagi and Thomas Peyrin, editors, *Advances in Cryptology – ASIACRYPT 2017, Part I*, volume 10624 of *Lecture Notes in Computer Science*, pages 3–33, Hong Kong, China, December 3–7, 2017. Springer, Cham, Switzerland.

[Gro16] Jens Groth. On the size of pairing-based non-interactive arguments. In Marc Fischlin and Jean-Sébastien Coron, editors, *Advances in Cryptology – EUROCRYPT 2016, Part II*, volume 9666 of *Lecture Notes in Computer Science*, pages 305–326, Vienna, Austria, May 8–12, 2016. Springer, Berlin, Heidelberg, Germany.

[JD11] David Jao and Luca De Feo. Towards quantum-resistant cryptosystems from supersingular elliptic curve isogenies. In Bo-Yin Yang, editor, *Post-Quantum Cryptography - 4th International Workshop, PQCrypto 2011*, pages 19–34, Tapei, Taiwan, November 29 – December 2 2011. Springer, Berlin, Heidelberg, Germany.

[Kan89] Masanobu Kaneko. Supersingular j-invariants as singular moduli mod p. *Osaka Journal of Mathematics*, 26(4):849–855, December 1989.

[Koh96] David Kohel. *Endomorphism rings of elliptic curves over finite fields*. PhD thesis, 1996.

[Lan87] Serge Lang. *Elliptic Functions*, volume 112 of *Graduate Texts in Mathematics*. Springer, 1987.

[LB20] Jonathan Love and Dan Boneh. Supersingular curves with small noninteger endomorphisms. In *ANTS XIV: Proceedings of the Fourteenth Algorithmic Number Theory Symposium*, volume 4 of *Open Book Series*, pages 7–22. Mathematical Sciences Publishers, 2020.

[Lip22] Helger Lipmaa. A unified framework for non-universal SNARKs. In Goichiro Hanaoka, Junji Shikata, and Yohei Watanabe, editors, *PKC 2022: 25th International Conference on Theory and Practice of Public Key Cryptography, Part I*, volume 13177 of *Lecture Notes in Computer Science*, pages 553–583, Virtual Event, March 8–11, 2022. Springer, Cham, Switzerland.

[LOX20] Songsong Li, Yi Ouyang, and Zheng Xu. Neighborhood of the supersingular elliptic curve isogeny graph at $j = 0$ and 1728. *Finite Fields and Their Applications*, 61:101600, 2020.

[LP24] Shai Levin and Robi Pedersen. Faster proofs and VRFs from isogenies. Cryptology ePrint Archive, Paper 2024/1626, 2024.

[Mes86] Jean-Francois Mestre. La méthode des graphes. exemples et applications. In *Proceedings of the international conference on class numbers and fundamental units of algebraic number fields (Katata)*, pages 217–242, 1986. English version (tr. Andrei Jorza): `https://wstein.org/papers/rank4/mestre-en.pdf`.

[Mor23] Tomoki Moriya. IS-CUBE: An isogeny-based compact KEM using a boxed SIDH diagram. Cryptology ePrint Archive, Paper 2023/1506, 2023.

[Mü95] Volker Müller. *Ein Algorithmus zur Bestimmung der Punktanzahl elliptischer Kurven über endlichen Körpern der Charakteristik größer drei*. PhD thesis, Universität des Saarlandes, 1995.

[NS22] Ngoc Khanh Nguyen and Gregor Seiler. Practical sublinear proofs for R1CS from lattices. In Yevgeniy Dodis and Thomas Shrimpton, editors, *Advances in Cryptology – CRYPTO 2022, Part II*, volume 13508 of *Lecture Notes in Computer Science*, pages 133–162, Santa Barbara, CA, USA, August 15–18, 2022. Springer, Cham, Switzerland.

[Piz90] Arnold K. Pizer. Ramanujan graphs and Hecke operators. *Bull. Am. Math. Soc., New Ser.*, 23(1):127–137, 1990.

[RS06] Alexander Rostovtsev and Anton Stolbunov. Public-key cryptosystem based on isogenies. Cryptology ePrint Archive, Report 2006/145, 2006. `https://eprint.iacr.org/2006/145`.

[RY07] Thomas Ristenpart and Scott Yilek. The power of proofs-of-possession: Securing multiparty signatures against rogue-key attacks. In Moni Naor, editor, *Advances in Cryptology – EUROCRYPT 2007*, volume 4515 of *Lecture Notes in Computer Science*, pages 228–245, Barcelona, Spain, May 20–24, 2007. Springer, Berlin, Heidelberg, Germany.

[Sil09] Joseph H. Silverman. *The arithmetic of elliptic curves*, volume 106 of *Grad. Texts Math.* New York, NY: Springer, 2nd ed. edition, 2009.

[SSW20] Peter Schwabe, Douglas Stebila, and Thom Wiggers. Post-quantum TLS without handshake signatures. In Jay Ligatti, Xinming Ou, Jonathan Katz, and Giovanni Vigna, editors, *ACM CCS 2020: 27th Conference on Computer and Communications Security*, pages 1461–1480, Virtual Event, USA, November 9–13, 2020. ACM Press.

[Sut12] Andrew V. Sutherland. Isogeny volcanoes. *CoRR*, abs/1208.5370, 2012.

[Tha22] Justin Thaler. Proofs, arguments, and zero-knowledge. *Foundations and Trends® in Privacy and Security*, 4(2–4):117–660, 2022.

[Tsu13] Kiminori Tsukazaki. *Explicit isogenies of elliptic curves*. PhD thesis, University of Warwick, July 2013.

[vzGG13] Joachim von zur Gathen and Jürgen Gerhard. *Modern Computer Algebra*. Cambridge University Press, 3rd edition, 2013.

[Was08] Lawrence C. Washington. *Elliptic curves. Number theory and cryptography*. Boca Raton, FL: Chapman and Hall/CRC, 2nd ed. edition, 2008.

[XZS22] Tiancheng Xie, Yupeng Zhang, and Dawn Song. Orion: Zero knowledge proof with linear prover time. In Yevgeniy Dodis and Thomas Shrimpton, editors, *Advances in Cryptology – CRYPTO 2022, Part IV*, volume 13510 of *Lecture Notes in Computer Science*, pages 299–328, Santa Barbara, CA, USA, August 15–18, 2022. Springer, Cham, Switzerland.

[XZZ+19] Tiancheng Xie, Jiaheng Zhang, Yupeng Zhang, Charalampos Papamanthou, and Dawn Song. Libra: Succinct zero-knowledge proofs with optimal prover computation. In Alexandra Boldyreva and Daniele Micciancio, editors, *Advances in Cryptology – CRYPTO 2019, Part III*, volume 11694 of *Lecture Notes in Computer Science*, pages 733–764, Santa Barbara, CA, USA, August 18–22, 2019. Springer, Cham, Switzerland.

A More on Elliptic Curves

A.1 Elliptic Curves and Isogenies

In this subsection we gather some well-known definitions and results on elliptic curves and isogenies.

Let K be a perfect field. An *elliptic curve* over K is a non-singular projective algebraic curve of genus 1 with a distinguished 'point at infinity', usually denoted by O. Given an elliptic curve E over K, the set of K-rational points $E(K)$ forms a group with neutral element O.

Let E_0 and E_1 be elliptic curves over K, and $K \subseteq L \subseteq \overline{K}$ a field extension of K. An *isogeny (defined) over* L, or L*-isogeny*, is a morphism of the curves $\phi\colon E_0 \longrightarrow E_1$ over L (in particular it can be expressed by rational maps with coefficients in L) which induces also a surjective group homomorphism on the sets of $\overline{K}$-rational points.

Any isogeny ϕ has a finite kernel, and the cardinality of this kernel equals the *degree* $\deg(\phi)$ of ϕ as a morphism if ϕ is separable [Sil09, Theorem III.4.10]. Further, if $\phi_1\colon E_0 \longrightarrow E_1$ and $\phi_2\colon E_1 \longrightarrow E_2$ are two isogenies, then

$$\deg(\phi_2 \circ \phi_1) = \deg(\phi_2) \cdot \deg(\phi_1).$$

Given an elliptic curve E_0 over K and any finite subgroup $G \subseteq E_0(\overline{K})$, there exist a unique (up to equivalence) elliptic curve E_1 and a separable isogeny $\phi_G\colon E_0 \longrightarrow E_1$ with kernel equal to G [Sil09, Proposition III.4.12].

A classic example of an n^2-isogeny is the *multiplication-by-n* endomorphism $[n]$ of E, which maps each $\overline{K}$-rational point of an elliptic curve to its n^{th} scalar multiple. The kernel of the induced map on the $\overline{K}$-rational points is called the *n-torsion* of E, denoted $E[n]$.

Notably, each n-isogeny $\phi\colon E_0 \longrightarrow E_1$ admits a *dual isogeny* of degree n, which is the unique isogeny $\hat{\phi}\colon E_1 \longrightarrow E_0$ such that $\phi \circ \hat{\phi} = \hat{\phi} \circ \phi = [n]$ [Sil09, Theorem III.6.1-2].

A.2 Canonical Modular Polynomials for $\kappa = 1$

Below we list the canonical modular polynomials for the primes $\ell \in \{2, 3, 5, 7, 13\}$, which are the primes that satisfy $\kappa = \frac{\ell-1}{\gcd(12,\ell-1)} = 1$.

$$\begin{aligned}
\Phi_2^c(X,j) &= X^3 + 48X^2 + 768X + 4096 - X \cdot j,\\
\Phi_3^c(X,j) &= X^4 + 36X^3 + 270X^2 + 756X + 729 - X \cdot j,\\
\Phi_5^c(X,j) &= X^6 + 30X^5 + 315X^4 + 1300X^3 + 1575X^2 + 750X + 125 - X \cdot j,\\
\Phi_7^c(X,j) &= X^8 + 28X^7 + 322X^6 + 1904X^5 + 5915X^4 + 8624X^3 + 4018X^2\\
&\quad + 748X + 49 - X \cdot j,\\
\Phi_{13}^c(X,j) &= X^{14} + 26X^{13} + 325X^{12} + 2548X^{11} + 13832X^{10} + 54340X^9\\
&\quad + 157118X^8 + 333580X^7 + 509366X^6 + 534820X^5 + 354536X^4\\
&\quad + 124852X^3 + 15145X^2 + 746X + 13 - X \cdot j.
\end{aligned}$$

A.3 The Polynomial Factors of Lemma 7

In this subsection we list the polynomials g_{ℓ,j^*} and $h_{\ell,j^*,\pm}$ described in Lemma 7 – we start with $j^* = 0$:

ℓ	$g_{\ell,0}$	$h_{\ell,0,+}$	$h_{\ell,0,-}$
2	1	$X + 16$	$X + 256$
3	$X + 27$	$X + 3$	$X + 243$
5	1	$X^2 + 10X + 5$	$X^2 + 250X + 3125$
7	$X^2 + 13X + 49$	$X^2 + 5X + 1$	$X^2 + 245X + 2401$
13	$X^2 + 5X + 13$	$X^4 + 7X^3 + 20X^2 + 19X + 1$	$X^4 + 247X^3 + 3380X^2 + 15379X + 28561$

Table 4: The polynomials $g_{\ell,0}$ and $h_{\ell,0,\pm}$.

Next we give the polynomials for $j^* = 1728$:

ℓ	$g_{\ell,1728}$	$h_{\ell,1728,+}$	$h_{\ell,1728,-}$
2	$X+64$	$X-8$	$X-512$
3	1	$X^2+18X-27$	$X^2-486X-19683$
5	$X^2+22X+125$	X^2+4X-1	$X^2-500X-15625$
7	1	$X^4+14X^3+63X^2+70X-7$	$X^4-490X^3-21609X^2$ $-235298X-823543$
13	$X^2+6X+13$	$X^6+10X^5+46X^4+108X^3$ $+122X^2+38X-1$	$X^6-494X^5-20618X^4$ $-237276X^3-1313806X^2$ $-3712930X-4826809$

Table 5: The polynomials $g_{\ell,1728}$ and $h_{\ell,1728,\pm}$.

A.4 Rational Formulas for the Reconstruction Theorem

In this subsection we give, for the Reconstruction Theorem, the ℓ-dependent expressions of the root f as a K-rational function in the coefficients of $\phi_\ell(f)$. In these formulas we index f on the left hand side by the prime ℓ for emphasis.

For $\ell = 3$ we write $\phi_3(f) = x + C$ to obtain

$$f_3 = -\frac{(2A^2 + 7ACj_0 + 3C^3j_0)^2}{2^8 \cdot 3 \cdot A^3 \cdot j_0}.$$

For $\ell \in \{5, 7, 13\}$ we can use $\mathrm{char}(K) \notin \{2, 3\}$ to compute $\mathrm{disc}(\phi_\ell(f))$ by evaluating $\mathrm{disc}(\phi_\ell(T))$ at $T = f$ according to Proposition 1(a). Thus computations in $\mathbb{Z}[T, T^{-1}]$ show

$$f_5 = \frac{(-48A)^3}{\mathrm{disc}(\phi_5(f))^3 \cdot j_0} \quad \text{and} \quad f_7 = \frac{(-48A)^3}{\mathrm{disc}(\phi_7(f)) \cdot j_0}.$$

Lastly, for $\ell \in \{2, 13\}$ we see that the discriminant $\Delta = 2^{11} \cdot 3^5 \cdot f^{\ell-1} \cdot A \cdot B \neq 0$ of the target curve lies in $K(s_1, \ldots, s_n)$ by Kohel's description of Vélu's formulas [Koh96, §2.4], and

$$f_2 = \frac{\Delta}{2^{11} \cdot 3^5 \cdot A \cdot B} \quad \text{and} \quad f_{13} = \frac{2^{297} \cdot 3^{135} \cdot A^{27} \cdot B^{27}}{\mathrm{disc}(\phi_{13}(f))^5 \cdot \Delta^2}.$$

B The Sylvester Matrix and the Resultant

In the following we discuss the theory of *resultants*, using [Bos18, §4.4] and [vzGG13, §6.3] as general references, together with novel applications that will be important for our proofs. Throughout this section we fix R to be a commutative (as well as unital and associative) ring.

We briefly recall the adjugate matrix: Let $n \in \mathbb{N}$ and suppose that we have a square matrix $M = (m_{ij}) \in R^{n \times n}$. For any $i, j \in [n] = \{1, \ldots, n\}$ we let M_{ij} denote the matrix obtained from M by removing the i^{th} row and the j^{th} column. Then the *adjugate matrix* $\mathrm{adj}(M)$ of M is defined as the square matrix $\mathrm{adj}(M) = (a_{ij})_{i,j=1}^n \in R^{n \times n}$ with entries

$$a_{ij} = (-1)^{i+j} \det(M_{ji}).$$

The following result gives the well known *Laplace expansion* formulas:

Lemma 14. We have

$$\mathrm{adj}(M) \cdot M = \det(M) I_n = M \cdot \mathrm{adj}(M)$$

where I_n denotes the n^{th} identity matrix. Equivalently, for any $i \in [n]$ we can compute $\det(M)$ via *Laplace expansion along the i^{th} row* given by

$$\det(M) = \sum_{j=1}^{n} (-1)^{i+j} \cdot m_{ij} \cdot \det(M_{ij}),$$

or via *Laplace expansion along the i^{th} column* given by

$$\det(M) = \sum_{j=1}^{n} (-1)^{j+i} \cdot m_{ji} \cdot \det(M_{ji}).$$

For $n \in \mathbb{N}_0$ we now consider the free R-module

$$\mathcal{P}_n := \{h \in R[X] \colon \deg(h) < n\}$$

of rank n, equipped with the *descending monomial basis* $(X^{n-1}, X^{n-2}, \ldots, X, 1)$.

Definition 2. Let $f, g \in R[X]$ be non-zero polynomials of respective degrees $d, e \in \mathbb{N}_0$ and define the R-linear map

$$\sigma \colon \mathcal{P}_e \times \mathcal{P}_d \to \mathcal{P}_{d+e}, \ (s, t) \mapsto s \cdot f + t \cdot g.$$

We call the representation matrix of σ with respect to the descending monomial bases of each $\mathcal{P}_n$ the *Sylvester matrix* $\mathrm{Syl}(f, g)$ of f and g.

Remark 7. In the literature the transpose of the representation matrix of σ is sometimes referred to as the Sylvester matrix instead – see, for example, [BPR06, Notation 4.12].

Example 3. For polynomials $f = a_0X^3 + a_1X^2 + a_2X + a_3$ and $g = b_0X^2 + b_1X + b_2$ with $a_0, b_0 \neq 0$ we have

$$\mathrm{Syl}(f, g) = \begin{pmatrix} a_0 & & b_0 & & \\ a_1 & a_0 & b_1 & b_0 & \\ a_2 & a_1 & b_2 & b_1 & b_0 \\ a_3 & a_2 & & b_2 & b_1 \\ & a_3 & & & b_2 \end{pmatrix}$$

where empty entries are understood to be zero.

The following observation is straightforward, but crucial:

Lemma 15. Let K be a field and let $f, g \in K[X]$ be non-zero polynomials. Then the rank of the Sylvester matrix of f and g satisfies

$$\mathrm{rk}\,\mathrm{Syl}(f, g) \leq \deg(f) + \deg(g) - \deg\gcd(f, g).$$

Proof. Let $h := \gcd(f, g)$ and set $m := \deg(h)$. By the rank-nullity theorem we have to prove that the kernel of σ has dimension at least m, which we can do by giving m kernel vectors that are K-linearly independent. Setting $s := g/h$ and $t := f/h$, we have the kernel vectors

$$(s \cdot X^{i-1}, -t \cdot X^{i-1}) \in \ker(\sigma) \text{ for } i \in [m].$$

Indeed, the degree of $s \cdot X^{i-1}$ (resp. $-t \cdot X^{i-1}$) is strictly smaller than $\deg(g)$ (resp. $\deg(f)$) for any $1 \leq i \leq m$, and

$$\sigma\left(s \cdot X^{i-1}, -t \cdot X^{i-1}\right) = X^{i-1} \cdot (sf - tg) = X^{i-1} \cdot h^{-1} \cdot (gf - fg) = 0.$$

Moreover, the above m vectors are K-linearly independent; in fact, even their first components are already K-linearly independent as they have pairwise distinct degrees. □

The determinant of the Sylvester matrix also plays an important role:

Definition 3. Let $f, g \in R[X]$ be non-zero polynomials. The *resultant* $\mathrm{res}(f, g)$ of f and g is defined as

$$\mathrm{res}(f, g) := \det\mathrm{Syl}(f, g).$$

If we consider two coprime polynomials, then the inequality in Lemma 15 becomes an equality:

Corollary 16. Let K be a field and let $f, g \in K[X]$ be non-zero polynomials. Then the following are equivalent:

(i) f and g are coprime, i.e. $\gcd(f, g) = 1$.

(ii) $\operatorname{rk} \operatorname{Syl}(f, g) = \deg(f) + \deg(g)$.

(iii) $\operatorname{res}(f, g) \neq 0$.

Proof. By definition $\operatorname{Syl}(f, g)$ is a square matrix with $\deg(f) + \deg(g)$ rows, so the equivalence of (ii) and (iii) follows from linear algebra. Moreover, Lemma 15 shows that (ii) implies (i); finally, suppose that f and g are coprime and let $(s, t) \in \ker(\sigma)$. Then the equality $sf = -tg$ shows that g divides s and f divides t, which forces $s = t = 0$ due to degree restrictions. Hence σ is an injection between $(\deg(f) + \deg(g))$-dimensional K-vector spaces, so it has to be an isomorphism. Thus its representation matrix $\operatorname{Syl}(f, g)$ has to be invertible, yielding (iii). □

Next we note an immediate but important compatibility result that we will use frequently in the sequel:

Lemma 17. Let $f, g \in R[X]$ be non-zero polynomials, and let $\varphi\colon R \to S$ be a ring homomorphism, extended to a ring homomorphism $\varphi\colon R[X] \to S[X]$ via coefficient-wise application. If φ preserves the degrees of f and g, then we have

$$\varphi(\operatorname{Syl}(f, g)) = \operatorname{Syl}(\varphi(f), \varphi(g)),$$

where on the left hand side φ is applied entry-wise. In particular, in this situation we have

$$\varphi(\operatorname{res}(f, g)) = \operatorname{res}(\varphi(f), \varphi(g)).$$

Proof. This follows from the definitions and the fact that the determinant is compatible with ring homomorphisms as it is defined as a multivariate polynomial in the entries of the matrix. □

This compatibility result also extends Corollary 16 to integral domains:

Corollary 18. Let R be an integral domain and let $f, g \in R[X]$ be non-zero polynomials. Then f and g share a common divisor of positive degree if and only if $\operatorname{res}(f, g) = 0$.

Proof. We consider the embedding $\varphi\colon R \to K$ of R into its field of fractions K, and its extension to $R[X] \to K[X]$ via coefficient-wise application. As this clearly preserves the degrees of f and g, we see with Lemma 17 that $\operatorname{res}(f, g)$ is non-zero if and only if $\operatorname{res}(\varphi(f), \varphi(g))$ is. Furthermore, f and g share no common divisor of positive degree if and only if $\varphi(f)$ and $\varphi(g)$ are coprime, as both statements are equivalent to the fact that $\varphi(f)$ and $\varphi(g)$ do not have a common root in an algebraic closure of K. Therefore the claim follows from Corollary 16. □

In our proofs we will consider the situation that $R = A[Y]$ is itself a polynomial ring, and we will be interested in deriving the resultant $\operatorname{res}(f, g)$ with respect to Y. As we want to connect the derivatives to *k-minors* of $\operatorname{Syl}(f, g)$, i.e. to determinants of $(k \times k)$-submatrices of $\operatorname{Syl}(f, g)$, we need *Jacobi's formula*:

Lemma 19 (Jacobi)**.** Let $R = A[Y]$ be a polynomial ring over a commutative ring A, let $n \in \mathbb{N}$ and let $M = (m_{ij}) \in R^{n \times n}$ be a square matrix. Furthermore let $\frac{\partial}{\partial Y} M$ denote the matrix obtained from M via entry-wise derivation. Then

$$\frac{\partial}{\partial Y} \det(M) = \operatorname{tr}\left(\operatorname{adj}(M) \cdot \frac{\partial}{\partial Y} M\right).$$

In particular, for any $k \in \{0, \ldots, n\}$ we have

$$\frac{\partial^k}{\partial Y^k} \det(M) \in R \cdot \{(n-k)\text{-minors of } M\},$$

i.e. the k^{th} derivative of $\det(M)$ with respect to Y is an R-linear combination of $(n-k)$-minors of M.

Proof. For ease of notation we index submatrices of M by indices in $[n] \times [n]$ where we disallow indices of removed rows as the first index respectively of removed columns as the second index. We proceed by induction on n: For $n = 1$ the matrix $\operatorname{adj}(M)$ has the single entry 1, so the formula clearly holds.

Assuming $n \geq 2$, we apply the product rule to the Laplace expansion of $\det(M)$ along the last column to obtain

$$\frac{\partial}{\partial Y} \det(M) = \sum_{i=1}^{n} (-1)^{i+n} \det(M_{in}) \frac{\partial}{\partial Y} m_{in} + (-1)^{i+n} m_{in} \frac{\partial}{\partial Y} \det(M_{in}). \tag{11}$$

By the induction hypothesis we furthermore have

$$\begin{aligned}\frac{\partial}{\partial Y} \det(M_{in}) &= \operatorname{tr}\left(\operatorname{adj}(M_{in}) \cdot \frac{\partial}{\partial Y} M_{in}\right) \\ &= \sum_{j=1}^{n-1} \sum_{l=1}^{i-1} (-1)^{j+l} \det((M_{in})_{lj}) \frac{\partial}{\partial Y} m_{lj} + \sum_{j=1}^{n-1} \sum_{l=i+1}^{n} (-1)^{j+l-1} \det((M_{in})_{lj}) \frac{\partial}{\partial Y} m_{lj}.\end{aligned}$$

Therefore swapping the summation order of l and i yields

$$\begin{aligned}&\sum_{i=1}^{n} (-1)^{i+n} m_{in} \frac{\partial}{\partial Y} \det(M_{in}) \\ &\quad= \sum_{j=1}^{n-1} \sum_{l=1}^{n} \sum_{i=l+1}^{n} (-1)^{l+j} (-1)^{(i-1)+(n-1)} m_{in} \det((M_{lj})_{in}) \frac{\partial}{\partial Y} m_{lj} \\ &\quad\quad+ \sum_{j=1}^{n-1} \sum_{l=1}^{n} \sum_{i=1}^{l-1} (-1)^{l+j} (-1)^{i+(n-1)} m_{in} \det((M_{lj})_{in}) \frac{\partial}{\partial Y} m_{lj} \\ &\quad= \sum_{j=1}^{n-1} \sum_{l=1}^{n} (-1)^{l+j} \det(M_{lj}) \frac{\partial}{\partial Y} m_{lj}\end{aligned}$$

where we used Laplace expansion of $\det(M_{lj})$ along the last column to get rid of the sum over i. Now we see that the first summands in Equation (11) give precisely the n^{th} outer sum above, so in total we obtain:

$$\frac{\partial}{\partial Y} \det(M) = \sum_{j=1}^{n} \sum_{i=1}^{n} (-1)^{i+j} \det(M_{ij}) \frac{\partial}{\partial Y} m_{ij} = \operatorname{tr}\left(\operatorname{adj}(M) \cdot \frac{\partial}{\partial Y} M\right) \tag{12}$$

Finally we argue why the second claim follows from this formula by induction on k. For $k = 0$ the claim is immediate as $\det(M)$ is the unique $(n-0)$-minor of M. Now expressing $\frac{\partial^{k-1}}{\partial Y^{k-1}} \det(M)$ as an R-linear combination of $(n-k+1)$-minors of M via the induction hypothesis, we see by Jacobi's formula (12) (applied to each $(n-k+1)$-minor of M) and the product rule that $\frac{\partial^k}{\partial Y^k} \det(M)$ is an R-linear combination of $(n-k+1)$-minors of M and their $(n-k+1-1)$-minors; the latter are $(n-k)$-minors of M, and the former are R-linear combinations of $(n-k)$-minors of M due to Laplace expansion, hence yielding the claim. □

The following consequence is tailored to our needs:

Corollary 20. Let $R = K[Y]$ be a polynomial ring over a field K and fix an element $y_0 \in K$. Additionally let $f, g \in R[X]$ be non-zero polynomials and extend the K-linear evaluation homomorphism $\varphi\colon R \to K$ given by $Y \mapsto y_0$ to a ring homomorphism $\varphi\colon R[X] \to K[X]$ via coefficient-wise application. Further suppose that φ preserves the X-degrees of f and g, and write

$$m := \deg \gcd(\varphi(f), \varphi(g)).$$

Then

$$\left.\frac{\partial^k}{\partial Y^k}\right|_{Y=y_0} \operatorname{res}(f, g) = 0 \text{ for } k \in \{0, \ldots, m-1\},$$

i.e. $\operatorname{res}(f, g) \in K[Y]$ has a root of multiplicity at least m at y_0.

Proof. Let $k \in \{0, \dots, m-1\}$ and set $n := \deg(f) + \deg(g)$. By Lemma 19 the k^{th} Y-derivative of $\operatorname{res}(f,g)$ is an R-linear combination of $(n-k)$-minors of $\operatorname{Syl}(f,g)$. Moreover, by Lemma 17 we have $\operatorname{Syl}(\varphi(f), \varphi(g)) = \varphi(\operatorname{Syl}(f,g))$, so the images of the $(n-k)$-minors of $\operatorname{Syl}(f,g)$ under φ are $(n-k)$-minors of $\operatorname{Syl}(\varphi(f), \varphi(g))$. Hence

$$\varphi\left(\frac{\partial^k}{\partial Y^k} \operatorname{res}(f,g)\right) = \frac{\partial^k}{\partial Y^k}\bigg|_{Y=y_0} \operatorname{res}(f,g)$$

is a K-linear combination of $(n-k)$-minors of $\operatorname{Syl}(\varphi(f), \varphi(g))$. Finally, by Lemma 15 we know that

$$\operatorname{rk}\operatorname{Syl}(\varphi(f), \varphi(g)) \leq \deg(\varphi(f)) + \deg(\varphi(g)) - m = n - m < n - k,$$

so all $(n-k)$-minors of $\operatorname{Syl}(\varphi(f), \varphi(g))$ are zero by linear algebra and the claim follows. □

To turn our attention to the second application of resultants in this paper, we relate the resultant $\operatorname{res}(f,g)$ back to f and g:

Lemma 21. Let R be an integral domain and let $f, g \in R[X]$ be non-zero polynomials of respective degrees $d, e \in \mathbb{N}_0$ such that $d + e \geq 1$. Then there are polynomials $(s,t) \in \mathcal{P}_e \times \mathcal{P}_d$ such that

$$\operatorname{res}(f,g) = sf + tg.$$

Proof. Recalling the definition of $\operatorname{res}(f,g)$ via the linear map $\sigma\colon \mathcal{P}_e \times \mathcal{P}_d \to \mathcal{P}_{d+e}$ and translating the existence of the adjugate matrix into linear maps, we obtain an R-linear map $\phi\colon \mathcal{P}_{d+e} \to \mathcal{P}_e \times \mathcal{P}_d$ such that

$$\sigma \circ \phi(h) = \det(\operatorname{Syl}(f,g)) \cdot h = \operatorname{res}(f,g) \cdot h$$

for all $h \in \mathcal{P}_{d+e}$. Applying this composition to $h = 1$, which is possible since $d + e \geq 1$, hence shows that $\phi(1) = (s,t)$ satisfies the required linear combination. □

With this we can finish arguing the missing part of the proof of Theorem 9:

Proposition 22. Let $\ell \leq 13$ be a prime and let p_ℓ be given according to

$$(p_2, p_3, p_5, p_7, p_{11}, p_{13}) = (13, 53, 379, 1217, 5101, 8387).$$

Additionally let K be a field of characteristic $\operatorname{char}(K) \notin [p_\ell]$ and let $j_0 \in K^\times$ be a non-zero j-invariant. Then $\Phi_\ell(j_0, Y)$ does not have a triple root in $\overline{K}$.

Proof. This proof is highly computational – the observational claims used along the way can be verified via the Sage script `maximal_primes.sage` found in the accompanying GitHub repository[3].

We first introduce some notation: For $a \in \{0, 1, 2\}$ we write

$$\Psi_a(X,Y) := \frac{\partial^a}{\partial Y^a} \Phi_\ell(X,Y) \in \mathbb{Z}[X][Y].$$

Furthermore we consider the resultants (with respect to the Y-variable)

$$g_1 := \operatorname{res}(\Psi_0, \Psi_1),\ g_2 := \operatorname{res}(\Psi_0, \Psi_2),\ g_3 := \operatorname{res}(\Psi_1, \Psi_2),$$

which are elements of $\mathbb{Z}[X]$. The content of each g_i, i.e. the greatest common divisor of its coefficients, does not have a prime factor larger than p_ℓ, so we divide each g_i by this scalar (for computational efficiency) as well as by the maximal power of X that divides g_i – to simplify notation, we will denote the polynomials we obtain through these divisions by g_i again. In the following we will apply the ring homomorphism

$$\eta\colon \mathbb{Z}[X][Y] \to K[Y],\ X \mapsto j_0,\ Y \mapsto Y$$

to all of these polynomials; we note that, since $p_\ell > \ell$, the Y-degree of each Ψ_a is preserved by this map.

Now suppose for a proof by contradiction that $\Phi_\ell(j_0, Y) \in K[Y]$ has a triple root. Then the polynomials $(\eta(\Psi_0), \eta(\Psi_1), \eta(\Psi_2))$ all share a common root, so due to Lemma 17 and Corollary 16

[3] https://github.com/QuSAC/IsogenyPoKviaCanonicalModPolys

we see that, as powers of j_0 and all primes not larger than p_ℓ are invertible in K, we have $\eta(g_i) = 0$ for each $i \in \{1,2,3\}$. Due to the previous divisions by the powers of X the polynomials (g_1, g_2, g_3) turn out to pairwise have no common factor of non-zero degree, i.e. the number

$$\gamma_\ell \coloneqq \gcd(\mathrm{res}(g_1, g_2), \mathrm{res}(g_1, g_3), \mathrm{res}(g_2, g_3)) \in \mathbb{Z}$$

is non-zero by Corollary 18, and we can see that p_ℓ is its largest prime factor. However, by Lemma 21 (noting that each g_i is non-constant) we find for any $i, j \in \{1,2,3\}$, $i < j$, polynomials $s_i, t_j \in \mathbb{Z}[X]$ such that $\mathrm{res}(g_i, g_j) = s_i g_i + t_j g_j$ and thus

$$\eta(\mathrm{res}(g_i, g_j)) = \eta(s_i)\eta(g_i) + \eta(t_j)\eta(g_j) = 0.$$

Hence each $\mathrm{res}(g_i, g_j)$ is zero in K, so $\mathrm{char}(K) = p > 0$ has to be a prime factor of γ_ℓ. As p_ℓ is the maximal prime factor of γ_ℓ, we obtain a contradiction to our assumption on $\mathrm{char}(K)$. □

An important special case of the resultant is the *discriminant*, which we will define now: Let R be an integral domain and let $f \in R[X]$ such that $\frac{\partial}{\partial X} f$ is non-zero. Then all entries in the first row of $\mathrm{Syl}(f, \frac{\partial}{\partial X} f)$ are divisible by the leading coefficient a_0 of f; therefore the resultant $\mathrm{res}(f, \frac{\partial}{\partial X} f)$ is also divisible by a_0 due to Laplace expansion along this first row, and one defines

$$\mathrm{disc}(f) \coloneqq (-1)^{\binom{\deg(f)}{2}} \cdot a_0^{-1} \cdot \mathrm{res}\left(f, \frac{\partial}{\partial X} f\right)$$

to be the *discriminant* of f. We directly obtain the following from Lemma 17:

Corollary 23. Let R be an integral domain and $f \in R[X]$ a polynomial such that $\frac{\partial}{\partial X} f$ is non-zero. Furthermore let $\varphi\colon R \to S$ be a ring homomorphism of integral domains, extended to a ring homomorphism $\varphi\colon R[X] \to S[X]$ via coefficient-wise application. If we have $\deg(\frac{\partial}{\partial X} f) = \deg(\frac{\partial}{\partial X} \varphi(f))$, then

$$\varphi(\mathrm{disc}(f)) = \mathrm{disc}(\varphi(f)).$$