Author: Derickx, M.
Title: Torsion points on elliptic curves over number fields of small degree
Issue Date: 2016-09-21

CHAPTER 3

Torsion points on elliptic curves over number fields of small degree

TORSION POINTS ON ELLIPTIC CURVES OVER NUMBER FIELDS OF SMALL DEGREE

MAARTEN DERICKX, SHELDON KAMIENNY, WILLIAM STEIN, AND MICHAEL STOLL

Abstract. We determine the set $S(d)$ of possible prime orders of K-rational points on elliptic curves over number fields K of degree d, for $d = 4, 5$ and 6.

1. Introduction

For an integer $d \geq 1$, we let $S(d)$ be the set of primes p such that there exists an elliptic curve E over a number field K of degree d with a K-rational point of order p in $E(K)$. The notation $\mathrm{Primes}(n)$ will be used to denote the set of all primes $\leq n$. Mazur [1977, 1978] has famously proved that

$$S(1) = \mathrm{Primes}(7).$$

Kamienny [1992b] showed that

$$S(2) = \mathrm{Primes}(13)$$

and Parent [2000, 2003], extending the techniques used by Mazur and Kamienny, proved that

$$S(3) = \mathrm{Primes}(13).$$

In fact $S(d)$ is finite for every d as proven in Merel [1996], and Merel even gave an explicit but super exponential bound on the largest element of $S(d)$. Shortly after Merel proved the finiteness of $S(d)$, Oesterlé managed to improve upon Merel's bound by showing $S(d) \subseteq \mathrm{Primes}((3^{d/2}+1)^2)$ if $d > 3$ and $S(3) \subseteq \mathrm{Primes}(37) \cup \{43\}$. The result of Parent mentioned earlier depends on Oesterlé's bound for $S(3)$ and a hypothesis Parent denoted by $(*)_p$ [Parent, 2000, p. 724] for the primes $p \leq 43$. The hypothesis $(*)_p$ is that the rank of the winding quotient $J_\mu^{\mathbf{e}}(p)$ is zero. Parent already mentioned that $(*)_p$ probably holds for all primes and that this result would follow from results announced by Kato, but these results were not yet published at the time that Parent wrote is article. These results have now indeed been published as Kato [2004]. Details on $J_\mu^{\mathbf{e}}(p)$ and how to do derive $(*)_p$ from the work of Kato are given in Section 4. Oesterlé never published his results, but was kind enough to give us his unpublished notes so that the gap in the literature could be filled. The Appendix of this article contains his arguments for showing that $S(d) \subseteq \mathrm{Primes}((3^{d/2} + 1)^2)$ for $d \geq 6$ and $S(d) \subseteq \mathrm{Primes}(410)$ for $d = 3, 4, 5$ as stated in Theorem A.2. His notes also included a section where he further improved the bound on $S(d)$ with

Date: September 7, 2016.

$d = 3, 4, 5$, but these are omitted since we have found it easier to deal with these cases using the techniques developed in the main text.

Theorem 1.1. *Suppose that* $S(d) \subseteq \text{Primes}(2281)$ *for* $3 \le d \le 7$*, then*

$$\begin{aligned} &S(3) = \text{Primes}\,(13),\\ &S(4) = \text{Primes}\,(17),\\ &S(5) = \text{Primes}\,(19),\\ &S(6) = \text{Primes}\,(19) \cup \{37\} \qquad \textit{and}\\ &S(7) \subseteq \text{Primes}\,(23) \cup \{37, 43, 59, 61, 67, 71, 73, 113, 127\}\,. \end{aligned}$$

The reason for including the condition $S(d) \subseteq \text{Primes}(2281)$ in the statement is to make it possible for us to give a proof that does not use Oesterlé's bound (Theorem A.1 of the appendix). Theorem A.2 of the appendix tells us that condition $S(d) \subseteq \text{Primes}(2281)$ is satisfied for $3 \le d \le 7$ so the conclusion of the above Theorem holds unconditionally. Theorem A.1 actually also implies $S(d) \subseteq \text{Primes}(2281)$ for $3 \le d \le 7$, but the proof given in the appendix depends on Theorem 1.1, so we need to use Theorem A.2 to avoid creating circular references. Additionally the reason for reproving the already known result on $S(3)$ is because the results of Parent [2000, 2003] depend on the unpublished results of Oesterlé. We cannot cite Parent in the appendix in order to prove $S(3) \subseteq \text{Primes}(43)$, since we want to give a proof of Oesterlé's unpublished results in the appendix.

From our computation it even follows that $S(7) \subseteq \text{Primes}\,(23) \cup \{37\}$ if the condition $(**)_{d,p,\ell}$ holds for $d = 7$, $p = 43, 59, 61, 67, 71, 73, 113, 127$ and $\ell = 2$.

> The effective divisors $D \subseteq X_1(p)^{(d)}(\mathbb{F}_\ell)$ such that the associated line bundle $\mathcal{O}_{X_1(p)_{\mathbb{F}_\ell}}(D)$ lifts to $\mathbb{Z}_{(\ell)}$ are exactly the effective divisors whose support consists of the cusps mapping to the cusp $0_{\mathbb{F}_\ell}$ of $X_0(p)(\mathbb{F}_\ell)$. $\quad (**)_{d,p,\ell}$

This condition is easily seen to be true if $p > (\ell^{d/2} + 1)^2$, see Section 5.3, and we managed even to verify it for many $p \le (2^{d/2} + 1)^2$ and $d \le 7$. However the verifying of the condition for the $p \le (2^{d/2} + 1)^2$ and $d \le 7$ was done using explicit calculations and careful case by case studies. Finding a theoretical argument that also works for $p \le (\ell^{d/2} + 1)^2$ is of interest though, since if there exists a function $P^{**} : \mathbb{N}_{>0} \to \mathbb{R}$ such that for every integer $d > 0$ and prime p with $p > P^{**}(d)$ one can find an $\ell > 2$ such that $(**)_{d,p,\ell}$ holds, then [Parent, 1999, Thm. 1] shows that $S(d) \subseteq \text{Primes}(\max(P^{**}(d), 65(2d)^6))$. So from the existence of a function $P^{**}(d) < (3^{d/2} + 1)^2$ as above one obtains an improvement upon Oesterlé's bound.

Let $S'(d)$ be the set of primes p such that there exist infinitely many elliptic curves E with a point of order p and pairwise distinct j-invariants over a number field K of degree d. One of course has $S'(d) \subseteq S(d)$. For $d = 1, 2$ or 3 one even has an equality $S'(d) = S(d)$ Mazur [1977], Kamienny [1992b], Jeon et al. [2011a]. There are a lot more $S'(d)$ known, indeed $S'(4) = \text{Primes}(17)$ Jeon et al. [2011b], $S'(5) = S'(6) = \text{Primes}(19)$ and $S'(7) = S'(8) = \text{Primes}(23)$ Derickx and van

Hoeij [2014]. These results, together with the fact that a twist of the elliptic curve $y^2 + xy + y = x^3 + x^2 - 8x + 6$ has a point of order 37 over the degree 6 number field $\mathbb{Q}(\sqrt{5}, \cos(2\pi/7))$ [Elkies, 1998, Eq. 108], show that we only need to prove $\subseteq$ instead of $=$ in Theorem 1.1.

The $\subseteq$ inclusions are obtained by studying the points on $X_1(p)$ over number fields of degree d. Indeed if E is an elliptic curve over a number field K of degree d and $P \in E(K)$ a point of order p, then the pair (E, P) gives rise to a point $s \in X_1(p)(K)$. If one lets $\sigma_1, \dots, \sigma_d : K \to \overline{\mathbb{Q}}$ be the d different embeddings of K in $\overline{\mathbb{Q}}$ then

$$s^{(d)} := \sum_{i=1}^{(d)} \sigma_i(s) \in X_1(p)^{(d)}(\mathbb{Q}) \tag{1}$$

is a $\mathbb{Q}$ rational point on the d-th symmetric power of $X_1(p)$. Conversely, every point in $X_1(p)^{(d)}(\mathbb{Q})$ can be written as $\sum_{i=1}^{m} n_i s_i^{(d_i)}$ with $s_i \in X_1(K_i)$, K_i a number field of degree d_i and $n_i \in \mathbb{N}_{>0}$. So the question whether $p \in S(d)$ can be answered if one can find all $\mathbb{Q}$ rational points on $X_1(p)^{(d)}$.

In Section 3 some general theory is developed that, if certain conditions are met, allows one to find all rational points on the symmetric powers of a curve. This theory is similar to the Chabauty for symmetric powers of curves in Siksek [2009], except for the fact that we use formal immersions, as done in Mazur [1978] and Kamienny [1992a], instead of the p-adic integration used by Siksek. As we will see later, this allows us to work over discrete valuation ring with smaller residue characteristic than Siksek. The discussion of Mazur and Kamienny is specific to modular curves, whereas in Section 3 we took care to write down how their arguments work out for arbitrary curves. The most essential part of Section 3 for obtaining Theorem 1.1 is the trick of Parent [2000] that allows one to also work over discrete valuation rings in characteristic 2: this trick is the use of assumption (3) instead of assumption (1) of Proposition 3.4 .

In Section 5 we spell out very explicitly what the results of Section 3 mean when applied to modular curves, giving several variations on the strategies of finding all rational points on symmetric powers of modular curves as a corollary of Section 3. We even work out the strategy explicitly enough so that it can be tested by a computer program written in Sage [2014]. Most cases were handled quite easily by this computer program, although the proof that $29, 31, 41 \notin S(d)$ for $d \leq 7$ and $73 \notin S(6)$ required some extra attention.

Acknoweldgements. We would like to thank Barry Mazur and Bas Edixhoven for their many valuable comments and suggestions, Pierre Parent for his suggestion to look at CM elliptic curves in order to show that $73 \notin S(6)$, and Tessa Schild for her proofreading.

2. Formal Immersions

Definition / Proposition 2.1 (Formal Immersion)**.** *Let $\phi\colon X \to Y$ be a morphism of Noetherian schemes and $x \in X$ be a point which maps to $y \in Y$. Then ϕ is a formal immersion at x if one of the two following equivalent conditions hold:*

- *the induced morphism of the complete local rings $\widehat{\phi}^*\colon \widehat{\mathcal{O}_{Y,y}} \to \widehat{\mathcal{O}_{X,x}}$ is surjective.*
- *The maps $\phi\colon k(y) \to k(x)$ and $\phi^*\colon \operatorname{Cot}_y(Y) \to \operatorname{Cot}_x(X)$ are both surjective.*

Proof. It is clear that the first condition implies the second. The other implication can be proved by using Nakayama's lemma to lift a basis of $\operatorname{Cot}_y(Y)$ to a set of generators $f_1, \ldots, f_n$ of m_y, the maximal ideal of $\widehat{\mathcal{O}_{Y,y}}$. The fact that $\widehat{\phi}^*(f_1), \ldots, \widehat{\phi}^*(f_n)$ generate m_x/m_x^2 implies that $\widehat{\phi}^*(f_1), \ldots, \widehat{\phi}^*(f_n)$ also generate m_x. As a consequence we get that for all i the map $m_y^i/m_y^{i+1} \to m_x^i/m_x^{i+1}$ is surjective, hence by the completeness of $\widehat{\mathcal{O}_{Y,y}}$ we also have that $\widehat{\phi}^*$ is surjective. □

There is one important property of formal immersions that we will use:

Lemma 2.2. *Let X, Y be Noetherian schemes. Let R be a Noetherian local ring, with maximal ideal m and residue field $k = R/m$. Suppose $f\colon X \to Y$ is a morphism of schemes that is a formal immersion at a point $x \in X(k)$ and suppose $P, Q \in X(R)$ are two points such that $x = P_k = Q_k$ and $f(P) = f(Q)$. Then $P = Q$.*

Proof. Let $y = f(x)$ and view P, Q as morphisms $\operatorname{Spec} R \to X$ and hence write $f \circ P$ instead of $f(P)$. The morphisms P, Q and f induce maps on the local rings, we will call these P_m^*, Q_m^* and f_x^* respectively:

$$\begin{array}{ccccc} R & \overset{P_m^*}{\underset{Q_m^*}{\leftleftarrows}} & \mathcal{O}_{X,x} & \xleftarrow{\ f_x^*\ } & \mathcal{O}_{Y,y} \\ \downarrow & & \downarrow & & \downarrow \\ \widehat{R} & \overset{\widehat{P_m^*}}{\underset{\widehat{Q_m^*}}{\leftleftarrows}} & \widehat{\mathcal{O}_{X,x}} & \xleftarrow{\ \widehat{f_x^*}\ } & \widehat{\mathcal{O}_{Y,y}} \end{array}$$

Since $f \circ P = f \circ Q$ we also know that $\widehat{P_m^*} \circ \widehat{f_x^*} = \widehat{Q_m^*} \circ \widehat{f_x^*}$. Now f is a formal immersion at x. This means $\widehat{f_x^*}$ is surjective and hence that $\widehat{P_m^*} = \widehat{Q_m^*}$. Because R is Noetherian local ring, the map $R \to \widehat{R}$ is injective and hence $P_m^* = Q_m^*$. The proposition now follows from the following commuting diagrams:

$$\begin{array}{ccc} & & X \\ & \overset{P}{\nearrow} & \uparrow \\ \operatorname{Spec} R & \xrightarrow{\ P_m\ } & \operatorname{Spec} \mathcal{O}_{X,x} \end{array} \qquad \begin{array}{ccc} & & X \\ & \overset{Q}{\nearrow} & \uparrow \\ \operatorname{Spec} R & \xrightarrow{\ Q_m\ } & \operatorname{Spec} \mathcal{O}_{X,x} \end{array}$$

□

3. Rational points on symmetric powers of curves

This section contains a very general discussion on rational points on symmetric powers of curves similar to [Siksek, 2009, §3]. There is a huge overlap where both the results of [Siksek, 2009, §3] and this section are applicable. However, both Siksek's and our own results are applicable in situations where the others result is not; both the overlap and differences will be discussed.

Throughout this section R will be a discrete valuation ring whose residue field k is perfect. Its fraction field will be denoted by K and its maximal ideal by $\mathfrak{m}$. If C is a smooth and projective curve over R, such that C_K is geometrically irreducible, then its Jacobian J exists. Let J^0 be the fiberwise connected component of 0, which is isomorphic to $\mathrm{Pic}^0_{C/R}$ and semi-Abelian [Bosch et al., 1990, §9.7 Cor. 2]. Since C is smooth over R, actually $J^0 = J$ and the special fiber of J is an Abelian variety, hence J is an Abelian scheme over R.

For any R-scheme S and any $x \in C^{(d)}(S)$, define

$$f_{d,x} : C_S^{(d)} \to J_S \tag{2}$$

as the map that for all S-schemes T and all $D \in C_S^{(d)}(T)$ sends D to the class of $\mathcal{O}_{C_T}(D - x_T)$ in $J_S(T)$, where we use [Bosch et al., 1990, §9.3 Prop. 3] to see the points in $C_S^{(d)}(T)$ as effective relative Cartier divisors of degree d on C_T over T.

The following Lemma is the key Lemma which will be used throughout this paper to study the rational points on $C^{(d)}$.

Lemma 3.1. *Let C be a smooth and projective curve over R with geometrically irreducible generic fiber and Jacobian J. Let $t : J \to A$ be a map of Abelian schemes* [1] *over R. Let $y \in C_k^{(d)}(k)$ and assume that the following conditions hold:*

(1) *$t(J^1(R)) = \{0\}$, where $J^1(R) := \ker\left(J(R) \overset{\text{red}}{\to} J(k)\right)$,*

(2) *the map $t \circ f_{d,y} : C_k^{(d)} \to A_k$ is a formal immersion at y.*

Then there is at most one point in $C^{(d)}(R)$ whose reduction is y.

Proof. If there is no point in $C^{(d)}(R)$ whose reduction is y, then there is nothing to prove, so let $x \in C^{(d)}(R)$ be a point whose reduction is y. Then condition 2 above ensures that $t \circ f_{d,x} : C^{(d)} \to A$ is a formal immersion at y. Indeed, both $\mathrm{Cot}_y\, C^{(d)} / \mathrm{Cot}_y\, C_k^{(d)}$ and $\mathrm{Cot}_y\, A / \mathrm{Cot}_y\, A_k$ are canonically isomorphic with $\mathfrak{m}/\mathfrak{m}^2 = \mathrm{Cot}_k\, R$, hence the surjectivity of $t \circ f_{d,x}{}^* : \mathrm{Cot}_0\, A \to \mathrm{Cot}_y\, C^{(d)}$ follows from the surjectivity of $(t \circ f_{d,y})^* : \mathrm{Cot}_0\, A_k \to \mathrm{Cot}_y\, C_k^{(d)}$. Now let $x' \in C^{(d)}(R)$ be a point

[1] one could even more generally take t to be a map from the formal group of J to a formal group F over R, and replace $f_{d,y}$ by $\widehat{f_{d,y}} : \mathrm{Spf}\, \widehat{\mathcal{O}_{C_k^{(d)},y}} \to \mathrm{Spf}\, \widehat{\mathcal{O}_{J_k,0}}$. But in the case where we want to apply this lemma the Abelian variety J_K is of GL_2 type and hence J had enough endomorphisms to not need to use the formal group version

whose reduction is y, then condition 1 together with $f_{d,x}(x')_k = 0_k = f_{d,x}(x)_k$ imply that $t \circ f_{d,x}(x') = 0_R = t \circ f_{d,x}(x)$. Finally, according to Lemma 2.2 the fact that $t \circ f_{d,x}$ is a formal immersion implies $x' = x$. □

The most straightforward way to turn the above Lemma into a way to determine all rational points in $C^{(d)}(R)$ is the following:

Theorem 3.2. *Let C be a curve that is smooth and projective over R such that C_K is geometrically irreducible, and let J denote its Jacobian over R. Let d be a positive integer and $S \subseteq C^{(d)}(R)$ be a finite set. Let $t : J \to A$ be a map of Abelian schemes over R, denote by red_k the reduction to k map and $\mu : C^{(d)} \to \mathrm{Pic}^d_{C/R}$ the map sending a divisor to its associated line bundle. Assume that the following conditions hold:*

1. $t(J^1(R)) = \{0\}$, *where* $J^1(R) := \ker\left(J(R) \overset{red}{\to} J(k)\right)$,
2. *the map* $t \circ f_{d,s} : C_k^{(d)} \to A_k$ *with* $f_{d,s}$ *as in Eq.* (2) *is a formal immersion at all* $s \in \mathrm{red}_k(S)$ *and*
3. $\mathrm{red}_k(S) = \mu^{-1}(\mathrm{red}_k(\mathrm{Pic}^d_{C/R}(R)))$.

Then $S = C^{(d)}(R)$.

Proof. Condition 3 ensures that $\mathrm{red}_k(C^{(d)}(R)) = \mathrm{red}_k(S)$, and the first two conditions together with Lemma 3.1 ensure that every point in $\mathrm{red}_k(S)$ has exactly one point in $C^{(d)}(R)$ reducing to it. □

In the above theorem however the set S might be huge, and it might get impractical to verify condition 2 explicitly in concrete examples. It turned out that this is the case in the situation where we want to apply it. However in our setup there will often exist a map of curves $\phi : C \to D$ such that the set S for which we want to prove $S = C^{(d)}(R)$ is the inverse image of a single point under $\phi^{(d)} : C^{(d)}(R) \to D^{(d)}(R)$. The following generalization of the above theorem whose proof is similar will be useful in these cases.

Theorem 3.3. *Let C and D be smooth and projective curves over R whose generic fibers are geometrically irreducible. Let $\phi : C \to D$ be a non constant map. Denote by J the Jacobian of D over R. Let d be a positive integer and $S \subseteq C^{(d)}(R)$ and $T \subseteq D^{(d)}(R)$ be finite sets such that $S = {\phi^{(d)}}^{-1}(T) \subseteq C^{(d)}(R)$. Let $t : J \to A$ be a map of Abelian schemes over R, denote by $\mu : C^{(d)} \to \mathrm{Pic}^d_{C/R}$ the map sending a divisor to its associated line bundle. Assume that the following conditions hold:*

1. $t(J^1(R)) = \{0\}$, *where* $J^1(R) := \ker\left(J(R) \overset{\mathrm{red}}{\to} J(k)\right)$,
2. *the map* $t \circ f_{d,s} : D_k^{(d)} \to A_k$ *with* $f_{d,s}$ *as in Eq.* (2) *is a formal immersion at all* $s \in \mathrm{red}_k(T)$ *and*
3. $\phi^{(d)}(\mu^{-1}(\mathrm{red}_k(\mathrm{Pic}^d_{C/R}(R)))) \subseteq \mathrm{red}_k(T)$.

Then $S = C^{(d)}(R)$.

$$\begin{array}{ccccc}
 & & S & \longrightarrow & T \\
 & & \big\downarrow & & \big\downarrow \\
\mathrm{Pic}^d_{C/R}(R) & \xleftarrow{\mu} & C^{(d)}(R) & \xrightarrow{\phi^{(d)}} & D^{(d)}(R) \\
\big\downarrow \mathrm{red}_k & & \big\downarrow \mathrm{red}_k & & \big\downarrow \mathrm{red}_k \\
\mathrm{Pic}^d_{C/R}(k) & \xleftarrow{\mu} & C^{(d)}(k) & \xrightarrow{\phi^{(d)}} & D^{(d)}(k)
\end{array}$$

Proof. Condition (3) ensures that

$$\mathrm{red}_k(\phi^{(d)}(C^{(d)}(R))) = \phi^{(d)}(\mathrm{red}_k(C^{(d)}(R))) \subseteq \mathrm{red}_k(T),$$

and the first two conditions together with Lemma 3.1 ensure that every point in $\mathrm{red}_k(T)$ has exactly one point in $D^{(d)}(R)$ reducing to it. So we can conclude that $\phi^{(d)}(C^{(d)}(R)) = T$ hence the theorem follows from the assumption $S = {\phi^{(d)}}^{-1}(T)$. $\square$

Remark. Theorems 3.2 and 3.3 are still true if one lets t depend on s. Theoretically this is not a huge gain since one can always take $t : J \to A$ to be the universal map of Abelian schemes such that (1) holds. However, if one wants to restrict the choice of t to $t \in \mathrm{End}_R J$, then the elements such that (1) holds form a two sided ideal $I \subseteq \mathrm{End}_R J$. If this ideal is not principal then it might pay to use a t that depends on s.

If condition (3) of Theorem 3.2 holds, then taking $T = \phi^{(d)}(S)$ ensures that (3) of Theorem 3.3 holds. However, even in the case that (3) of Theorem 3.2 fails to hold for $S = C^{(d)}(R)$, it might still be possible to find an $\phi : C \to D$ and a $T \subseteq D^{(d)}(R)$ such that (3) of Theorem 3.3 holds. The only case where we will make use of this is for showing $73 \notin S(6)$. There we found a $\mathbb{Q}$ rational point $x^{(6)} \in (X_1(73)/\langle 10\rangle)^{(6)}(\mathbb{Q})$ that was the only $\mathbb{Q}$-rational point in its residue class mod 2. We could show that none of the points $X_1(73)^{(6)}(\overline{\mathbb{Q}})$ mapping to $x^{(6)}$ were defined over $\mathbb{Q}$, hence we could show that the 4 points in $X_1(73)^{(6)}(\mathbb{F}_2)$ mapping to $x^{(6)}_{\mathbb{F}_2}$ had no $\mathbb{Q}$-rational points above them.

If the curve C is a smooth curve over some global field and one has generators for a finite index subgroup of the Mordell-Weil group of (a quotient of) J, then instead of using Theorem 3.2 or Theorem 3.3 for a single prime, one could even use the Mordell-Weil sieve as described in [Siksek, 2009, §5] to combine the information about the rational points of $C^{(d)}$ obtained by Lemma 3.1 for several primes. This however, was not necessary for our purposes.

In the setting where we want to apply Lemma 3.1, the ring R will be $\mathbb{Z}_{(\ell)}$. In this case $J^1(R)$ is a finite index subgroup of $J(R)$ and hence we need $t(J(R))$ to be finite in order for condition (1) to be satisfied. Conversely if $t(J(R))$ is finite, then there are some quite mild conditions on t, A and R that imply that condition(1) is satisfied.

Proposition 3.4. *Suppose that $R = \mathbb{Z}_{(\ell)}$ and $t(J(R))$ is finite and either*

(1) $\ell > 2$,
(2) $\ell = 2$ *and* $A(R)[2]$ *injects into* $A(\mathbb{F}_2)$, *or*
(3) $\ell = 2$ *and* $t = t_2 \circ t_1$ *where* $t_1 : J \to A', t_2 : A' \to A$ *are maps of Abelian schemes such that* $t_1(J(R))$ *is finite and* t_2 *kills all the elements in* $A'(R)[2]$ *that reduce to* 0 *mod* 2.

then condition (1) *of Lemma 3.1 is satisfied.*

Proof. If either $\ell > 2$, or $\ell = 2$ and $A(R)[2]$ injects into $A(k)$, then $t(J(R)) \to A(k)$ is injective, hence $t(J^1(R)) = \{0\}$ which deals with the first two cases. Alternatively one could see them as special cases of the third one with $t_2 = 1$. In the third case we know that $t_1(J^1(R))$ is finite and contained in the kernel of reduction. But a $\mathbb{Z}_{(2)}$ valued torsion point that specializes to the identity mod 2 on a group scheme must be a two torsion point [Parent, 2000, Lem 1.7]. This means that $t_1(J^1(R)) \subset A'(R)[2]$ and hence $t_2 \circ t_1(J^1(R)) = \{0\}$ by the definition of t_2. □

A more general statement of the above proposition over arbitrary discrete valuation rings of unequal characteristics also easily obtained by using [Parent, 2000, Prop 2.3] in the proof instead of Lemma 1.7 of loc. cit.. Bu

In the case that the map t of Lemma 3.1 is the identity map, condition (2) of that Lemma can be nicely restated in terms of $C^{(d)}_{2,k}$ where $C^{(d)}_{2,k} \subseteq C^{(d)}_k$ is defined as the closed sub-variety corresponding to the divisors D over $\bar{k}$ of degree d such that $H^0(C_{\bar{k}}, \mathcal{O}_{C_{\bar{k}}}(D))$ is a $\bar{k}$ vector space whose dimension is at least 2.

Proposition 3.5. *Let $y \in C^{(d)}_k(\bar{k})$ be a point then the map $f_{d,y} : C^{(d)}_{\bar{k}} \to J_{\bar{k}}$ is a formal immersion at y if and only if $y \notin C^{(d)}_{2,k}(\bar{k})$. In particular if $C(k) \neq \emptyset$, then $f_{d,y}$ is a formal immersion at all points in $C^{(d)}_k(k)$ if and only if $k(C_k)$ contains no non-constant functions of degree $\leq d$.*

Proof. Since the map $\mathcal{L} \mapsto \mathcal{L}(-y)$ induces an isomorphism $\operatorname{Pic}^d C_k \to J$, we see that $f_{d,y}$ is a formal immersion at y if and only if the canonical map $C^{(d)}_k \to \operatorname{Pic}^d_{C_k/k}$ is. The map $C^{(d)}_k \setminus C^{(d)}_{2,k} \to \operatorname{Pic}^d_{C_k/k}$ is an isomorphism onto its image, which proves the "if"-part. For the "only if"-part one just notices that if $y \in C^{(d)}_{2,k}$, then the connected component of y of the fiber of $f_{d,y}$ above $0 = f_{d,y}(y)$ contains a $\mathbb{P}^1$. The tangent directions inside this $\mathbb{P}^1$ at y are all send to 0 by $f_{d,y}$ hence it is not a formal immersion. □

Let C be as in the above proposition, let $x \in C_k$ be a closed point, $k(x)$ be its residue field and $q \in \widehat{O_{C_k,x}}$ be a uniformizer. The completed local ring $\widehat{\mathcal{O}_{C_k,x}}$ is isomorphic to $k(x)[[q]]$, and if we have a global 1-form $\omega \in \Omega^1_{C_k/k}(C_k)$, then we can write its pullback to $\widehat{\mathcal{O}_{C_k,x}}$ as $f\mathrm{d}q$ with f in $\widehat{\mathcal{O}_{C_k,x}}$, hence we can write:

$$\omega_{\mathcal{O}_{C_k,x}} = \sum_{n=1}^{\infty} a_n q^{n-1}\mathrm{d}q, \quad a_n \in k(x). \tag{3}$$

The right hand side of the above formula is called the q-expansion of ω.

The map $f_{1,x} : C_{k(x)} \to J_{k(x)}$ induces an isomorphism $f_{1,x}^* : H^0(J_{k(x)}, \Omega^1) \to H^0(C_{k(x)}, \Omega^1)$ and evaluation in zero gives an isomorphism $H^0(J_{k(x)}, \Omega^1) \to \mathrm{Cot}_0 J_{k(x)}$. If $\omega' \in \mathrm{Cot}_0 J_{k(x)}$ corresponds to $\omega \in H^0(C_{k(x)}, \Omega^1)$ under these isomorphisms then we also say that $\sum_{n=1}^{\infty} a_n q^{n-1}\mathrm{d}q$ is the q-expansion of ω'.

The following complete local rings are equal

$$\widehat{\mathcal{O}_{C^{(d)}_{k(x)},dx}} = k(x)[[q_1, \ldots, q_d]]^{S_d} = k(x)[[\sigma_1, \ldots, \sigma_d]] \tag{4}$$

where q_i is the pullback of q along the i'th projection map $\pi_i : C^d_{k(x)} \to C_{k(x)}$ and $\sigma_1 := q_1 + \cdots + q_d$ up to $\sigma_d := q_1 q_2 \cdots q_d$ are the elementary symmetric polynomials in q_1 up to q_d. Let $\overline{\mathrm{d}\sigma_i}$ denote the image of $\mathrm{d}\sigma_i$ in $\mathrm{Cot}_{dx}\, C^{(d)}_{k(x)}$, then $\overline{\mathrm{d}\sigma_1}$ up to $\overline{\mathrm{d}\sigma_d}$ form a basis of $\mathrm{Cot}_{dx}\, C^{(d)}_{k(x)}$. The following Lemma is due to Kamienny and can be found implicitly for example in the proof of Proposition 3.1 of Kamienny [1992a].

Lemma 3.6. *Let d be an integer, C, J and $f_{d,dx} : C^{(d)}_{k(x)} \to J_{k(x)}$ be as in the setup of Lemma 3.1 for $x \in C_k$ a closed point. Let q be a uniformizer at x, q_i, σ_i as above and $\omega \in \mathrm{Cot}_0 J_{k(x)}$ an element with q-expansion $\sum_{n=1}^{\infty} a_n q^{n-1}\mathrm{d}q$. Then*

$$\sum_{n=1}^{d} (-1)^{n-1} a_n \overline{\mathrm{d}\sigma_n} = f^*_{d,dx}\omega \in \mathrm{Cot}_{dx}\, C^{(d)}_{k(x)}$$

Proof. Let $p : C^d_{k(x)} \to C^{(d)}_{k(x)}$ denote the quotient map then $f_{d,dx} \circ p = \sum_{i=1}^{d} f_{1,x} \circ \pi_i$ where $\pi_i : C^d_{k(x)} \to C_{k(x)}$ denotes the i'th projection map. In particular,

$$(f_{d,dx} \circ p)^*(\omega) = \sum_{n=1}^{\infty} a_n \left(\sum_{i=1}^{d} q_i^{n-1}\mathrm{d}q_i \right).$$

For a ring B consider the map of $B[[\sigma_1, \ldots, \sigma_d]]$-modules

$$D_B : \bigoplus_{j=1}^{d} B[[\sigma_1, \ldots, \sigma_d]]\mathrm{d}\sigma_j \to \bigoplus_{i=1}^{d} B[[q_1, \ldots, q_d]\mathrm{d}q_i$$

given by $\mathrm{d}\sigma_j \mapsto \sum_{i=1}^{d} \frac{\partial \sigma_j}{\partial q_i}\mathrm{d}q_i$. If we define $s_j := \sum_{i=1}^{d} q_i^j$ for all integers j and $\sigma_j = 0$ for all $j > d$, then Newton's identities give

$$s_n + \sum_{j=1}^{n-1} (-1)^j \sigma_j s_{n-j} = (-1)^{n-1} n \sigma_n.$$

Applying d to this expression shows that

$$(-1)^{n-1}\mathrm{d}\sigma_n - \sum_{i=1}^{d} q_i^{n-1}\mathrm{d}q_i = \frac{1}{n}\mathrm{d}\left(\sum_{j=1}^{n-1}(-1)^j\sigma_j s_{n-j}\right)$$

for $B = \mathbb{Q}$. The right hand side is actually contained in $\bigoplus_{j=1}^{n-1} I\mathrm{d}\sigma_j$ where $I \subset \mathbb{Z}[[\sigma_1,\ldots,\sigma_d]]$ is the ideal generated σ_1 up to σ_d. The proposition follows by base changing $D_{\mathbb{Z}}$ to D_k and quotient out by $\bigoplus_{j=1}^{d} I\mathrm{d}\sigma_j$. □

In the proposition below and its proof we identify $C^{(d)}(k)$ with the set of k rational effective divisors of degree d on C.

Proposition 3.7. *Let $y \in C^{(d)}(k)$ be a point and write $y = \sum_{j=1}^{m} n_j y_j$ with $y_j \in C^{(d)}(\bar{k})$ distinct and $m, n_1, \ldots, n_m \in \mathbb{N}_{>0}$. Let q_j be a uniformizer at y_j, e be a positive integer and $\omega_1, \cdots, \omega_e \in t^*(\mathrm{Cot}_0 A_{\bar{k}}) \subseteq \mathrm{Cot}_0 J_{\bar{k}}$. For $1 \le i \le e$ and $1 \le j \le m$ let $a(\omega_i, q_j, n_j) := (a_1(\omega_i), \ldots, a_{n_j}(\omega_i))$ be the row vector of the first n_j coefficients of ω_i's q_j-expansion.*

Then $t \circ f_{d,y} : C_k^{(d)} \to A_k$ is a formal immersion at y if the matrix

$$A := \begin{bmatrix} a(\omega_1, q_1, n_1) & a(\omega_1, q_2, n_2) & \cdots & a(\omega_1, q_1, n_m) \\ a(\omega_2, q_1, n_1) & a(\omega_2, q_2, n_2) & \cdots & a(\omega_2, q_1, n_m) \\ \vdots & \vdots & \ddots & \vdots \\ a(\omega_e, q_1, n_1) & a(\omega_e, q_2, n_2) & \cdots & a(\omega_e, q_1, n_m) \end{bmatrix} \tag{5}$$

has rank d. If $\omega_1, \cdots, \omega_e$ generate $t^*(\mathrm{Cot}_0 A_{\bar{k}})$, then the previous statement even becomes an equivalence.

Proof. The natural map $\prod_{j=1}^{m} C_{\bar{k}}^{(n_j)} \to C_{\bar{k}}^{(d)}$ is étale at $(n_1y_1, n_2y_2, \ldots, n_my_m)$, hence we get an isomorphism of cotangent spaces

$$\mathrm{Cot}_y C_{\bar{k}}^{(d)} \cong \bigoplus_{j=1}^{m} \mathrm{Cot}_{d_jy_j} C_{\bar{k}}^{(d_j)}.$$

For j from 1 up to m and $1 \le i \le n_j$ let $\sigma_{j,i}$ be the symmetric functions associated to q_j as in (4). The elements $(-1)^{i-1}\overline{\mathrm{d}\sigma_{j,i}}$ with $1 \le j \le m$ and $1 \le i \le n_j$ form a basis of $\mathrm{Cot}_y C_{\bar{k}}^{(d)}$ under this isomorphism. The corollary follows since if $1 \le h \le e$ is an integer then the h'th row of A is just $f_{d,y}^*(\omega_h)$ with respect to this basis. □

If one takes $R = \mathbb{Z}_\ell$ with $\ell > \max_i(n_i)$ then the matrix A in Theorem 1 of Siksek [2009] is obtained by dividing the columns of the matrix A above by certain column dependent integers $\le \max_i(n_i)$. Actually, there is a huge overlap between Theorem 1 of Siksek [2009] and the result one gets when combining Lemma 3.1 and Proposition 3.7. Our version has the advantage that one doesn't have the conditions $\ell > \max_i(n_i)$. The reason is that the formal immersion criterion, and more generally the formal group over R approach, do not introduce denominators in the matrix A, while the p-adic logarithm in Siksek's Chabauty approach does introduce them, since it is only defined over K and not over R. Theorem 1 of Siksek [2009] has the advantage that one has more freedom in the choice of the one-forms ω_i. For example, our version is useless if J is simple and has rank $r > 1$, while Siksek's version is still applicable in cases where $r + d \leq g$ where g is the genus of C, although this problem can be mitigated by replacing the map $t : J \to A$ by a map of formal groups in Lemma 3.1. The reason for not using the results of Siksek [2009] is that we really want to take $\ell = 2$, since in general the number of points on $C^{(d)}(\mathbb{F}_\ell)$ is the smallest for $\ell = 2$, so that we need to check the formal immersion condition (2) of Lemma 3.1 for fewer points.

The entire strategy in this section depends on the existence of a map $t : J \to A$ of Abelian varieties whose image contains only finitely many rational points as in Proposition 3.4. The main goal of the following section is to explicitly describe a quotient of J that has only finitely many rational points in the case that C is a modular curve.

4. The winding quotient

In this section we will let N be an integer and $H \subseteq (\mathbb{Z}/N\mathbb{Z})^*$ a subgroup. The curve X_H over $\mathbb{Z}[1/N]$ is defined to be the quotient curve $X_1(N)/H$ where $(\mathbb{Z}/N\mathbb{Z})^*$ acts as the diamond operators. Taking $H = 1$ gives $X_1(N)$ and $H = (\mathbb{Z}/N\mathbb{Z})^*$ gives $X_0(N)$.

Integration gives a map

$$H_1(X_H(\mathbb{C}), cusps, \mathbb{Z}) \to \mathrm{Hom}_{\mathbb{C}}(H^0(X_H(\mathbb{C}), \Omega^1), \mathbb{C}) \cong H_1(X_H(\mathbb{C}), \mathbb{R}).$$

By a theorem of Manin and Drinfeld the image of this map is contained in $H_1(X_H(\mathbb{C}), \mathbb{Q})$. Let $\{0, \infty\} \in H_1(X_H(\mathbb{C}), cusps; \mathbb{Z})$ be the element coming from a path from 0 to $i\infty$ in the complex upper half plane.

Definition 4.1. The element $\mathbf{e} := \omega \mapsto \int_{\{0,\infty\}} \omega \in H_1(X_H(\mathbb{C}), \mathbb{Q})$ is called the winding element and the corresponding ideal $\mathcal{A}_{\mathbf{e}} := \mathrm{Ann}(\mathbf{e}) \subseteq \mathbb{T}$, consisting of the elements annihilating $\mathbf{e}$, is called the winding ideal. The quotient $J_H^{\mathbf{e}} := J_H/\mathcal{A}_{\mathbf{e}} J_H$ is called the winding quotient.

One can also define $X_{\mu,H}$ to be the quotient of $X_\mu(N)$ by H. The winding element and the winding quotient can be defined in the same way, and the latter will be denoted by $J_{\mu,H}^{\mathbf{e}}$. The isomorphism

$$W_N \colon X_\mu(N) \to X_1(N) \tag{6}$$

sending $(E, f : \mu_N \to E[N])$ to $(E/\mathrm{im}(f), f^\vee : \mathbb{Z}/N\mathbb{Z} \to E[N]/\mathrm{im}(f))$ is defined over $\mathbb{Z}[1/N]$. It interchanges the cusps 0 and ∞ and commutes with taking the quotient by H. This isomorphism sends the winding ideal of $X_{\mu,H}$ to the winding ideal of X_H and hence we get an isomorphism $J_{\mu,H}^{\mathbf{e}} \cong J_H^{\mathbf{e}}$.

The essential property of the winding quotient is that its group of rational points is finite.

Theorem 4.2. *The rank of* $J_H^{\mathbf{e}}(\mathbb{Q})$ *and* $J_{\mu,H}^{\mathbf{e}}(\mathbb{Q})$ *are* 0 .

Merel was in [Merel, 1996, §1] the first one to introduce the winding quotient for $J_0(p)$ with p prime, where he also proves that its rank is finite using a result from Kolyvagin and Logachëv [1989]. This result states that an abelian variety A over $\mathbb{Q}$ that is a quotient of $J_0(N)_{\mathbb{Q}}$ has Mordel-Weil rank 0 if its analytic rank is zero. Parent in [Parent, 1999, §3.8] generalized Merels statement it to composite numbers N. The result of Kolyvagin and Logachev was generalized by Kato [Kato, 2004, Cor. 14.3] to abelian varieties that are a quotient of $J_1(N)_{\mathbb{Q}}$. In both Parent [2000] and Parent [2003] it is mentioned that the theorem follows from using Kato's generalization. Here is a short sketch how to deduce the finiteness of the winding quotient form the work of Kato, where we closely follow the arguments of [Parent, 1999, §3.8].

Proof. Because $J_H^{\mathbf{e}}$ is a quotient of $J_1^{\mathbf{e}}(N)$ and $J_H^{\mathbf{e}} \cong J_{\mu,H}^{\mathbf{e}}$, it suffices to show the theorem for $J_1^{\mathbf{e}}(N)$.

The Hecke algebra $\mathbb{T}_{\mathbb{Q}}$ viewed as subalgebra of the endomorphism ring of $S_2(\Gamma_1(N))_{\mathbb{Q}}$ can be written as

$$\mathbb{T}_{\mathbb{Q}} := R_{f_1} \times R_{f_2} \ldots \times R_{f_k}$$

where the f_i range over all Galois orbits of newforms for Γ_1 of level M_i dividing N and R_{f_i} is the restriction of $\mathbb{T}_{\mathbb{Q}}$ to the subspace $\mathcal{E}_{f_i}$ of $S_2(\Gamma_1(N))_{\mathbb{Q}}$ consisting of all elements that can be written as $\mathbb{Q}$-linear combinations of the Galois conjugates of $B_d(g)$ with $g \in f_i$ and $d \mid N/M_i$ and $B_d \colon X_1(N) \to X_1(M_)$ the degeneracy maps [Parent, 1999, Thm. 3.5]. Now let M be an integer that divides N and d an integer dividing N/M. The degeneracy map $B_d \colon X_1(N) \to X_1(M)$ gives rise to $B_d^* \colon J_1(M)_{\mathbb{Q}} \to J_1(N)_{\mathbb{Q}}$ and we can define

$$J_1(N)_{\mathbb{Q}}^{new} := J_1(N)_{\mathbb{Q}} / \sum_{M|N, M \neq N, d|N/M} \operatorname{im} B_d^*.$$

And we can use the maps $B_{d,*} \colon J_1(N)_{\mathbb{Q}} \to J_1(M)_{\mathbb{Q}}$ to define a map of abelian varieties

$$\Phi \colon J_1(N)_{\mathbb{Q}} \to \bigoplus_{M|N} \bigoplus_{d|N/M} J_1(M)_{\mathbb{Q}}^{new}.$$

Now the identification

$$S_2(\Gamma_1(N))_{\mathbb{C}} \cong H^0(X_1(N)_{\mathbb{C}}, \Omega^1) \cong H^0(J_1(N)_{\mathbb{C}}, \Omega^1) \cong \operatorname{Cot}_0(J_1(N)_{\mathbb{C}}$$

together with the isomorphism $\oplus_{M|N} \oplus_{d|N/M} S^2(\Gamma_1(M))_{\mathbb{C}}^{new} \to S^2(\Gamma_1(M))_{\mathbb{C}}^{new}$ shows that $\Phi_{\mathbb{C}}$ is an isogeny, so Φ is one also. We also have an isogeny $J_1(M)^{new} \to \oplus J_f$ where f runs over the Galois orbits of newforms in $S_2(\Gamma_1(M))$ and J_f is the abelian variety attached to such a Galois orbit. Combining these isogenies with Φ we get an isogeny

$$J_1(N)_{\mathbb{Q}} \to \bigoplus_i \bigoplus_{d|N/M_i} J_{f_i,\mathbb{Q}}.$$

where the f_i range over all Galois orbits of newforms for Γ_1 of level M_i dividing N. Define R^{f_i} as $\oplus_{j\neq i} R_{f_j}$, with this definition the product $\oplus_{d|N/M_i} J_{f_i,\mathbb{Q}}$ will be isogenous to $J_1(N)_\mathbb{Q}/R^{f_i}J_1(N)_\mathbb{Q}$.

Now Parent shows that if the integration pairing $\langle \mathbf{e}, f_i\rangle$ is non-zero, then $\mathcal{A}_{\mathbf{e},\mathbb{Q}} \cap R_{f_i} = 0$ and conversely that if $\langle \mathbf{e}, f_i\rangle = 0$, then $\mathcal{A}_{\mathbf{e},\mathbb{Q}} \cap R_{f_i} = R_{f_i}$. Now since $L(f_i, 1) = 2\pi\langle \mathbf{e}, f_i\rangle$ we can write

$$\mathcal{A}_{\mathbf{e},\mathbb{Q}} = \bigoplus_{i:L(f_i,1)=0} R_{f_i}.$$

Combining this with the previous discussion we get an isogeny

$$J_1^{\mathbf{e}}(N) \to \bigoplus_{i:L(f_i,1)\neq 0} J_1^{\mathbf{e}}(N)/R^{f_i}J_1^{\mathbf{e}}(N) \to \bigoplus_{i:L(f_i,1)\neq 0} \oplus_{d|N/M_i} J_{f_i,\mathbb{Q}}$$

where the latter product has rank 0 by Kato's theorem. □

5. The conditions of 3.3 for $X_\mu(N) \to X_{\mu,H}$.

Let p be a prime. In order to determine $X_1(p)^{(d)}(\mathbb{Q})$, or equivalently $X_\mu(p)^{(d)}(\mathbb{Q})$ by using the isomorphism W_p defined in (6), we will apply Theorem 3.3 to the quotient map $f : X_\mu(p) \to X_{\mu,H}$ where $H \subseteq (\mathbb{Z}/p\mathbb{Z})^*$ is some subgroup such that we manage to verify all conditions. Much of the strategy also works if one drops the assumption that p is a prime.

5.1. **Condition 1: Using the winding quotient.** Let N be an integer, $\ell \nmid N$ a prime and $H \subseteq (\mathbb{Z}/N\mathbb{Z})^*$ a subgroup. Then we can use Theorem 4.2 to construct a $t : J_{\mu,H} \to A$ for some Abelian variety A such that (1) of 3.3 holds, i.e. such that $t(J^1_{\mu,H}(\mathbb{Z}_{(\ell)})) = 0$ where $J^1_{\mu,H}(\mathbb{Z}_{(\ell)})$ is the kernel of reduction. The combination of this Theorem and this Proposition gives.

Proposition 5.1. *Let $\ell > 2$ be a prime coprime to N then condition* (1) *of 3.3 is satisfied with $R = \mathbb{Z}_{(\ell)}$ for the quotient map $t : J_{\mu,H} \to J^{\mathbf{e}}_{\mu,H}$.*

This proposition will not be used in this text, but it is stated since it allows for comparison with other approaches of determining or bounding $S(d)$.

The proposition above is used for $J_0(p)$ with p prime and an ℓ that depends on p in the argument of Merel [1996], and is used for $J_0(p^n)$ for $\ell = 3$ or 5 in the argument of Parent [1999]. It was used by Oesterlé with $\ell = 3$ to prove his exponential bound $(3^{d/2} + 1)^2$, although it is only implicitly used in the Appendix since the part of Oesterlé's argument that uses it is replaced by a citation to Parent [1999]. The need for $\ell > 2$ is also the reason for the occurrence of 3 and not 2 as the base for the exponent in Oesterlé's bound.

The set $X_\mu(N)^{(d)}(\mathbb{F}_\ell)$ has fewer elements for elements for smaller ℓ so one would like to use $\ell = 2$ if $\ell \nmid N$ in view of applying Lemma 3.1. However, there are two difficulties that arise when doing so. The first one is that it is not necessarily true that the $J_{\mu,H}(\mathbb{Q})_{tors}$ injects into $J_{\mu,H}(\mathbb{F}_2)$. The second difficulty arises when determining which elements in $\mathrm{Cot}_0\,(J_{\mu,H})_{\mathbb{F}_l}$ come from $\mathrm{Cot}_0\,(J^{\mathbf{e}}_{\mu,H})_{\mathbb{F}_l}$ as needed for Proposition 3.7. This is because the exact sequence that relates $\mathrm{Cot}_0\,(J_{\mu,H})_{\mathbb{F}_l}$ to $\mathrm{Cot}_0\,(J^{\mathbf{e}}_{\mu,H})_{\mathbb{F}_l}$ for $\ell > 2$ is not necessarily exact for $\ell = 2$. In Parent [2000] there is already a way of dealing with these difficulties when using $X_\mu(N)$. His solution is to take $t_1 : J_\mu(N) \to J_\mu(N)$ to be a Hecke operator that factors via $J^{\mathbf{e}}_\mu$ and $t_2 : J_\mu(N) \to J_\mu(N)$ such that it kills all the two torsion in $J^1_\mu(N)(\mathbb{Z}_{(2)})$ and apply Proposition 3.4.

The operator t_2 as needed for Proposition 3.4 can be obtained using the following proposition.

Proposition 5.2. *Let $q \nmid N$ be a prime, then $(T_q - \langle q\rangle - q)(Q) = 0$*[2] *for all $Q \in J_{\mu,H}(\mathbb{Q})_{tors}$ of order coprime to q.*

Proof. Let $Q \in J_{\mu,H}(\mathbb{Q})$ be torsion of order coprime to q, then $(T_q - \langle q\rangle - q)(Q)$ is also a point of order coprime to q. Now let $Q_{\mathbb{F}_q} \in J_H(p)_{\mathbb{F}_q}(\mathbb{F}_q)$ be its specialisation and let Frob_q be the Frobenius on $J_H(p)_{\mathbb{F}_q}$ and Ver_q its dual (verschiebung). Then we have the Eichler-Shimura relation $T_{q,\mathbb{F}_q} = \langle q\rangle\,\mathrm{Frob}_q + \mathrm{Ver}_q$ see [Diamond and Im, 1995, p. 87] and $\mathrm{Ver}_q \circ \mathrm{Frob}_q = q$ in $\mathrm{End}_{\mathbb{F}_q}(J_H(p)_{\mathbb{F}_q})$. So

$$T_{q,\mathbb{F}_q}(Q_{\mathbb{F}_q}) = \langle q\rangle\,\mathrm{Frob}_q(Q_{\mathbb{F}_q}) + \mathrm{Ver}_q(Q_{\mathbb{F}_q}) = \langle q\rangle Q_{\mathbb{F}_q} + qQ_{\mathbb{F}_q}$$

giving $(T_{q,\mathbb{F}_q} - \langle q\rangle - q)(Q_{\mathbb{F}_q}) = 0$. Since specializing a point on a group scheme can only change its order by a power of the characteristic of the residue field we see that the order of $(T_q - \langle q\rangle - q)(Q)$ must be a power of q, and coprime to q at the same time hence $(T_q - \langle q\rangle - q)(Q) = 0$. □

What we need now is to find a way to find a Hecke operator t_1 as in Proposition 3.4. Now suppose if $t_1 \in \mathbb{T}$ is such that $t_1\mathcal{A}_e = 0$ then t_1 is a Hecke operator such that $t_1\colon J_{\mu,H} \to J_{\mu,H}$ factors via $J^{\mathbf{e}}_{\mu,H}$. Lemma 1.9 of Parent [1999] already gives a way of finding such Hecke operators for J_μ as soon as we have found an element t that generates the Hecke algebra $\mathbb{T}_1(N)_\mathbb{Q}$. If N is a prime then the Hecke algebra $\mathbb{T}_1(N)_\mathbb{Q}$ is of prime level and weight 2 so it is a product of number fields. In particular we know that such a t exists. By just trying “random” elements we should probably find such a t reasonably fast. However if N is composite this is not necessarily true. And even in the prime case testing whether t is a generator is a computationally expensive task if t is represented by a huge matrix, so we don’t want to try many different t’s. Therefore we generalize his Lemma slightly so that we don’t need t to be a generator.

[2] This is slightly different from [Parent, 2000, prop. 1.8], in that proposition it should also read $a_q := T_q - \langle q\rangle - q$. The mistake in that paper comes from Parent using the Eichler-Shimura relation for the $X_1(N)$ while in his article he is working with $X_\mu(N)$, although he denotes our $X_\mu(N)$ by $X_1(N)$. For more details on the Eichler-Shimura relations on $X_\mu(N)$ and $X_1(N)$ see [Diamond and Im, 1995, p. 87]

Proposition 5.3. *Let $t \in \mathbb{T}_{\Gamma_H}$ be an element and let $P(X) = \prod_{i=1}^n P_i(X)^{e_i}$ its factorized characteristic polynomial when viewing t as an element of* $\operatorname{End} S_2(\Gamma_H)_\mathbb{Q}$. *Define*

$$I := \{i \in \{1, \ldots, n\} \mid (P/P_i)(t)\mathbf{e} = 0 \text{ or } e_i > 1\}$$

then $t_1(t) := \prod_{i\in I} P_i^{e_i}(t)$ is such that $t_1\mathcal{A}_e = 0$.

Proof. We have already seen that the Hecke algebra $\mathbb{T}_{\Gamma_H,\mathbb{Q}}$ viewed as sub algebra of the endomorphism ring of $S_2(\Gamma_H)_\mathbb{Q}$ can be written as

$$\mathbb{T}_{\Gamma_H,\mathbb{Q}} := R_{f_1} \times R_{f_2} \ldots \times R_{f_k}$$

where the f_i range over all Galois orbits of newforms for Γ_H of level M_i dividing N and the R_{f_i} are the restriction of $\mathbb{T}_{\Gamma_H,\mathbb{Q}}$ to certain subspaces $\mathcal{E}_{f_i}$ of $S_2(\Gamma_{\Gamma_H,\mathbb{Q}})_\mathbb{Q}$. And we have also seen that $\mathcal{A}_{e,\mathbb{Q}} = \oplus_{i:L(f_i,1)=0} R_{f_i}$. Now define $\mathcal{E}_e := \oplus_{i:L(f_i,1)=0}\mathcal{E}_{f_i}$ and $\mathcal{E}_e^\perp := \oplus_{i:L(f_i,1)\neq 0}\mathcal{E}_{f_i}$ then $S_2(\Gamma_H)_\mathbb{Q} = \mathcal{E}_e \oplus \mathcal{E}_e^\perp$ and $\mathcal{A}_{e,\mathbb{Q}} := \left\{t' \in \mathbb{T}_\mathbb{Q} \mid t'|_{\mathcal{E}_e^\perp} = 0\right\}$ so in particular $t_1\mathcal{A}_{e,\mathbb{Q}} = 0$ if $t_1|_{\mathcal{E}_e} = 0$. So it suffices to show that $t_1|_{\mathcal{E}_{f_i}} = 0$ for all i such that $L(f_i, 1) = 0$. Now all $\mathcal{E}_i$ are contained in some generalized eigenspace corresponding to the factor $P_{j_i}^{e_{j_i}}$ for some j_i depending on i. Now for the i such that $e_{j_i} > 1$ we have $P_{j_i}^{e_{j_i}}(t)|\mathcal{E}_{f_i} = 0$ so $t_1|\mathcal{E}_{f_i} = 0$. For the other i we have $e_{j_i} = 1$ and in particular $\mathcal{E}_{f_i} = \ker P_{j_i}(t)$ so that we have $P/P_{j_i}(t) \in R_i$, now $L(f_i, 1) = 0$ implies $P/P_{j_i}(t)e = 0$ hence $j_i \in I$ and hence $t_1|_{\mathcal{E}_{f_i}} = t_1|_{\ker P_{j_i}(t)} = 0$ □

If N is composite then one can get away with a smaller set than I in the previous proposition, because then not all the terms with $e_i > 1$ are needed. On can see which ones are not needed by studying the action of t on the space of new forms or $\Gamma_1(M)\Gamma_H$ for all $M \mid N$. But this is not necessary for our application.

5.2. **Condition 2: Kamienny's criterion.** Let N be an integer and $H \subseteq (\mathbb{Z}/N\mathbb{Z})^*$ a subgroup, denote by $S_\infty \subseteq X_{\mu,H}(\mathbb{Q})$ the set of cusps that map to the cusp ∞ under the map $X_{\mu,H} \to X_0(N)$. On has that there are exactly $\phi(N)/\#\{\pm H\}$ elements in S_∞, where ϕ is Euler's totient function. Actually $(\mathbb{Z}/N\mathbb{Z})^*/\{\pm H\}$ acts transitively and freely on them. Define

$$S_\infty^{(d)} := \pi(S_\infty^d) \subseteq X_{\mu,H}^{(d)}(\mathbb{Q}), \tag{7}$$

where $\pi : X_{\mu,H}^d \to X_{\mu,H}^{(d)}$ is the quotient map. Then we want to be able to check whether condition (2) of Theorem 3.3 holds for $S = S_\infty^{(d)}$. In order to do this we make the following definition.

Definition 5.4. Let d be an integer, $n_0 \geq n_1 \geq \ldots \geq n_i \geq 1$ a sequence of integers that sum to d and $c_0, \ldots, c_i$ pairwise distinct cusps in $X_{\mu,H}$ that lie above $\infty \in X_0(N)$, then we call $n_0c_0 + \ldots + n_ic_i$ an **ordered sum ∞ cusps (of degree d)**.

It is clear that every element of $S_\infty^{(d)}$ can be written as an ordered sum of cusps in a unique way.

Remark. If $X_{\mu,H} = X_0(p)$ there is only one ordered sum of ∞ cusps of degree d, namely $d\infty$. So in this case condition (2) is the easiest to verify.

The proposition we will use to verify (2) of Theorem 3.3 is the following variant of Kamienny's Criterion which is a slight generalization of the variant [Parent, 2000, Prop. 2.8].

Proposition 5.5 (Kamienny's Criterion). *Let $\ell \nmid N$ be a prime, $c = n_1c_1 + \ldots + n_mc_m$ be an ordered sum of ∞ cusps of $X_{\mu,H}$ of degree d. Let $\langle d_1\rangle, \ldots, \langle d_m\rangle \in (\mathbb{Z}/N\mathbb{Z})^*/\{\pm H\}$ be the diamond operators such that $\infty = \langle d_j\rangle c_j$, where this time ∞ is the cusp of $X_{\mu,H}$ corresponding to $\infty \in \mathbb{P}^1(\mathbb{Q})$. Let $f_{d,c}\colon X^{(d)}_{\mu,H} \to J_{\mu,H}$ as in Eq. (2), let $t \in \mathbb{T}_{\Gamma_H}$ and view t as a map $J_{\mu,H} \to J_{\mu,H}$ then $t \circ f_{d,c}$: is a formal immersion at $c_{\mathbb{F}_\ell}$ if and only if the d Hecke operators*

$$(T_i\langle d_j\rangle t)_{\substack{j\in 1,\ldots,m \\ i\in 1,\ldots,n_i}} \tag{8}$$

are $\mathbb{F}_\ell$ linearly independent in $\mathbb{T}_{\Gamma_H} \otimes \mathbb{F}_\ell$.

Specializing to the case $X_{\mu,H} = X_0(N)$ where $S^{(d)}_\infty = \{d\infty\}$ the condition in Eq. (8) above becomes: The map $t \circ f_{d,d\infty}$ is a formal immersion at $d\infty_{\mathbb{F}_\ell}$ if and only if the d Hecke operators

$$T_1t, T_2t, \ldots, T_dt \tag{9}$$

are $\mathbb{F}_\ell$ linearly independent in $\mathbb{T}_{\Gamma_0(N)} \otimes \mathbb{F}_\ell$.

Proof of Proposition 5.5. We have $k(t \circ f_{d,c}(c_{\mathbb{F}_\ell})) = k(0_{\mathbb{F}_\ell}) = \mathbb{F}_l = k(c_{\mathbb{F}_\ell})$ so we only need to check that the linear independence criterion is equivalent to

$$(t \circ f_{d,c})^*\colon\ \mathrm{Cot}_{0_{\mathbb{F}_\ell}} J_{\mu,H} \to \mathrm{Cot}_{c_{\mathbb{F}_\ell}} X^{(d)}_{\mu,H}$$

being surjective.

Let $E_q/\mathbb{Z}[1/N][[q]]$ be the Tate curve. It has a canonical $\mu_{N,\mathbb{Z}[1/N][[q]]}$ embedding α coming from the unifomization map. The pair (E_q, α) gives a formal coordinate at the cusp ∞ of $X_\mu(N)_{\mathbb{Z}[1/N]}$ and since $X_\mu(N) \to X_{\mu,H}$ is unramified at ∞ it also gives a formal coordinate on $X_{\mu,H}$ at ∞. An element $\omega \in H^0(X_{\mu,H,\mathbb{Z}[1/N]}, \Omega^1)$ with q-expansion $\sum_{i=1}^\infty a_iq^{i-1}\mathrm{d}q$ is sent to the cusp form $f_\omega := \sum_{i=1}^\infty a_iq^i$ under the isomorphism $H^0(X_{\mu,H,\mathbb{Z}[1/N]}, \Omega^1) \cong S_2(\Gamma_H, \mathbb{Z}[1/N])$. Let $q_j = \langle d_j\rangle^*q$, then q_j is a formal coordinate at c_j. And the q_j expansion of ω at c_j is $\langle d_j\rangle f\mathrm{d}q_j/q_j$. This shows that the $a(\omega, q_j, n_j)$ defined as in Proposition 3.7 is given by

$$a(\omega, q_j, n_j) = a_1(\langle d_j\rangle f_\omega), a_2(\langle d_j\rangle f_\omega), \ldots, a_{n_j}(\langle d_j\rangle f_\omega).$$

The q expansion of $t^*\omega$ is tf_ω, now let $\omega_1, \cdots, \omega_g$ be generators of $H^0(X_{\mu,H,\mathbb{Z}[1/N]}, \Omega^1)$, then $t^*\omega_1, \cdots, t^*\omega_g$ generate $t^*H^0(X_{\mu,H,\mathbb{Z}[1/N]}, \Omega^1)$. In particular, using Proposition 3.7 we see that $t^*f_{d,c}$ is a formal immersion at $c_{\mathbb{F}_\ell}$ if and only if the matrix

$$A := \begin{bmatrix} a(t^*\omega_1, q_1, n_1) & a(t^*\omega_1, q_2, n_2) & \cdots & a(t^*\omega_1, q_1, n_m) \\ a(t^*\omega_2, q_1, n_1) & a(t^*\omega_2, q_2, n_2) & \cdots & a(t^*\omega_2, q_1, n_m) \\ \vdots & \vdots & \ddots & \vdots \\ a(t^*\omega_g, q_1, n_1) & a(t^*\omega_g, q_2, n_2) & \cdots & a(t^*\omega_g, q_1, n_m) \end{bmatrix} \quad (10)$$

has rank d over $\mathbb{F}_\ell$.

Now by formula (5.13) of Diamond and Shurman [2005] we have for an integer $1 \leq n \leq n_j$ that $a(t^*\omega_i, q_j, n_j)_n = a_n(\langle d_j \rangle t f_{\omega_i}) = a_1(T_n \langle d_j \rangle t f_{\omega_i})$. Using the isomorphism $\mathbb{T}_{\Gamma_H}/\ell\mathbb{T}_{\Gamma_H} \to Hom(S_2(\Gamma_H, \mathbb{F}_\ell), \mathbb{F}_\ell)$ [Diamond and Im, 1995, Prop. 12.4.13] [3] given by $T \mapsto (f \mapsto a_1(Tf))$ we see that for all $1 \leq i \leq g$ and $n \leq n_j$ we can replace the column of A that contains the elements $a(t^*\omega_i, q_j, n_j)_n$ by $T_n \langle d_j \rangle t$. $\square$

5.2.1. *Making the testing of Kamienny's criterion for $X_{\mu,H}$ faster.* As we have already seen Kamienny's criterion for $X_\mu(N)$ requires the testing of a lot of linear independence relations while Kamienny's criterion for $X_0(N)$ requires testing only 1 linear independence relation. To be more specific about what we mean by a lot, suppose that d is the degree and $p = N$ the the prime for which we want to check the Kamienny's criterion of $X_\mu(p)$ and we only consider the ordered sums of ∞ cusps $n_1c_1, \ldots, n_ic_i$ where the multiplicities $n_1, \ldots, n_i$ are all equal to 1 (hence $i = d$) then there are already $\binom{(p-3)/2}{d-1}$ different linear independencies we need to verify. So when doing actual computations using a computer we rather use $X_0(p)$ instead of $X_\mu(p)$ whenever possible. While doing the explicit computations, it turned out that the $X_0(p)$ version of the criterion sometimes fails for primes which are too big to make it practical to just try the $X_\mu(p)$ criterion for all possible ordered cusp sums. For example, we were unable to find t_1 and t_2 such that the $X_0(p)$ version of the criterion was satisfied for $d = 7$ and $p = 193$. In this case the $X_\mu(p)$ version would require verifying more than 869 million linear independencies and the matrices involved are 1457 by 1457. But luckily we can do something smarter.

We again restrict our attention to the ordered sums of ∞ cusps $n_1c_1 + \cdots + n_ic_i$ where the multiplicities $n_1, \ldots, n_i$ are all equal to 1 and hence $d = i$. Checking Kamienny's criterion for all these sums of cusps comes down to checking whether

$$\langle d_1 \rangle t, \ldots, \langle d_i \rangle t$$

are linearly independent for each set of pairwise distinct diamond operators $\langle d_1 \rangle, \ldots, \langle d_i \rangle$ where the first one is the identity. However, equivalently we can also check that all linear dependencies over $\mathbb{F}_l$ between the Hecke operators $\langle 1 \rangle t, \ldots, \langle (p-1)/2 \rangle t$ involve at least $d + 1$ nonzero coefficients. It turned out that the dimension of this space of linear dependencies was often zero or of very low dimension, so it takes no time at all to use a brute force approach and just calculate the number of nonzero

[3] There they show it only for $\Gamma_H = \Gamma_0(N)$ or $\Gamma_H = \Gamma_1(N)$, however the statement for Γ_H follows from the statement for $\Gamma_1(N)$, because $S_2(\Gamma_H, \mathbb{Z}) = S_2(\Gamma_1(N), \mathbb{Z})^{\Gamma_H/\Gamma_1(N)}$ and $\mathbb{T}_{\Gamma_H} = \mathbb{T}_{\Gamma_1(N)}|_{S_2(\Gamma_H, \mathbb{Z})}$.

coefficients of all linear dependencies. The following lemma generalizes this example to the case where the $n_1, \ldots, n_i$ are not necessarily equal to 1. This trick makes it more feasible to check the $X_\mu(N)$ version of the criterion on the computer.

Lemma 5.6. *Let $\ell \nmid N$ be a prime, d be an integer and $t \in \mathbb{T}_{\Gamma_H}$ and let $D \subset \mathbb{Z}$ be a set of representatives of $(\mathbb{Z}/N\mathbb{Z})^* / \{\pm H\}$ such that $1 \in D$. Define for all integers r with $\lfloor \frac{d}{2} \rfloor \leq r \leq d$ the following set*

$$D_r := \{(1, i) \mid d - r < i \leq r\} \cup \{(k, i) \mid 1 \leq i \leq d - r, k \in D\}.$$

Suppose that for all r with $\lfloor \frac{d}{2} \rfloor \leq r \leq d$ there is no $\mathbb{F}_\ell$ linear dependence among d of the elements $(T_i \langle k \rangle t)_{(k,i) \in D_r}$ in $\mathbb{T}_{\Gamma_H} / \ell \mathbb{T}_{\Gamma_H}$. Then $t \circ f_{d,c} \colon X^{(d)}_{\mu,H} \to J_{\mu,H}$ is a formal immersion at $c_{\mathbb{F}_\ell}$ for all ordered sums of ∞ cusps $c := n_1 c_1 + \cdots + n_m c_m$ of degree d.

Proof. Suppose for contradiction that there is an ordered sum of ∞ cusps $c := n_1 c_1 + \cdots + n_m c_m$ of degree d such that $t \circ f_{d,c}$ is not a formal immersion at $c_{\mathbb{F}_\ell}$. Write $c_1 = \langle d_j \rangle c_j$ with $d_j \in D$ for $1 \leq j \leq m$ then by 5.5 we see that the d vectors

$$(T_i \langle k \rangle t)_{((k,i) \in S)}, \quad S := \{(d_j, i) \mid 1 \leq j \leq m, \, 1 \leq i \leq n_j\}$$

are $\mathbb{F}_l$ linearly dependent in $\mathbb{T}_{\Gamma_H} \otimes \mathbb{F}_l$. We know that

$$\min(n_1, d - n_1) \geq n_2 \geq n_3 \geq \ldots \geq n_m.$$

So if $n_1 \geq \lfloor \frac{d}{2} \rfloor$ then $S \subseteq D_{n_1}$ and if $n_1 \leq \lfloor \frac{d}{2} \rfloor$ then $S \subseteq D_{d-n_1}$ so both cases lead to a contradiction. □

5.2.2. *Testing the criterion.* Using a computer program written in Sage we first tested the criterion for $X_0(p)$. The program and the output generated by it will be available at `http://www.math.leidenuniv.nl/nl/theses/`, the location where this thesis is published. The results of testing the criterion are summarised in the following propositions.

Proposition 5.7. *If $p = 131, 139, 149, 151, 167, 173, 179, 181, 191$ or p is a prime with $193 < p \leq 2281$ then there are $t_1, t_2 \in \mathbb{T}_{\Gamma_0(p)}$ as in Proposition 3.4 with $C = X_0(p)$ and $J = A = J_0(p)$ such that $T_1 t_1 t_2 T, \ldots, T_7 t_1 t_2$ are $\mathbb{F}_2$ linearly independent in $\mathbb{T}_{\Gamma_0(p)} \otimes \mathbb{F}_2$.*

Proof. The computer tested the criterion for all $17 \leq p \leq 2281$ using different choices of t_1 and t_2. The t_1 that were tried are $t_1 = t_1(t)$ as in Proposition 5.3, using $t = T_2, \ldots, T_{60}$, and the t_2 that were tried are $t_2 = T_q - q - 1$ for all primes $2 < q < 20$ with $q \neq p$. For all primes mentioned above the computer found at least one pair t_1, t_2 such that the linear independence holds. The total time used was about 2 hours[4] when checking the criterion for about 8 primes in parallel so it could be used to check the criterion for bigger d and p. □

[4]This is not a very precise timing and meant for indicative purposes only.

Testing the fast version of the criterion for $X_\mu(p)$ gives the following proposition:

Proposition 5.8. *For all pairs (p,d) with p a prime $p \leq 193$ and $3 \leq d \leq 7$ not satisfying any of the following conditions:*

- $d = 3$ *and* $p \in \mathrm{Primes}(17)$
- *($d = 4$ or $d = 5$) and* $p \in \mathrm{Primes}(19) \cup \{29\}$
- *($d = 6$ or $d = 7$) and* $p \in \mathrm{Primes}(37)$

there are $t_1, t_2 \in \mathbb{T}_{\Gamma_1(p)}$ as in Proposition 3.4 with $C = X_\mu(p)$ and $J = A = J_\mu(p)$ such that for $t = t_1t_2$ the D_r as in lemma 5.6 do not contain a subset of size d which is linearly dependent over $\mathbb{F}_2$.

Proof. This was again verified using the computer. This time the t_1, t_2 that were tried are $t_1 = t_1(t)$ for $t = T_2, \ldots, T_{20}$ and $t_2 = T_q - q - \langle q \rangle$ for the primes $2 < q < 20$ only trying new choices of t_1 and t_2 if no successful pair combination of t_1 and t_2 had been found yet. The most time was spent on the case $p = 193$ which took about 14 hours.[4]And that while only one combination of t_1 and t_2 was tried since $t_1 = t_1(T_2)$ and $t_2 = T_3 - 3 - \langle 3 \rangle$ already gave the desired result. □

5.3. **Condition 3: Study of $X_1(p)^{(d)}(\mathbb{F}_2)$.** For a prime $p > 7$ we know from Mazur [1977] that $Y_1(p)(\mathbb{Q}) = \emptyset$ and hence that $X_1(p)(\mathbb{Q})$ consists of $(p-1)/2$ cusps that map to the cusp 0 on $X_0(p)$. Let $S_0 \subseteq X_1(p)(\mathbb{Q})$ be the set of these $(p-1)/2$ cusps mapping 0 on $X_0(p)$, and define

$$S_0^{(d)} := \pi(S_0^d) \subseteq X_1(p)^{(d)}(\mathbb{Q}) \tag{11}$$

where $\pi : X_1(p)^d = X_1(p)^{(d)}$ is the quotient map, then $S_0^{(d)} = W_p^{(d)}(S_\infty^{(d)})$. We would like to verify condition (3) of Theorem 3.2 with $S = S_\infty^{(d)}$ and $C = X_\mu(p)$ when taking $R = \mathbb{Z}_{(\ell)}$, or condition (3) of *Theorem* 3.3 with $C = X_\mu(p)$, $D = X_0(p)$, $S = S_\infty^{(d)}$ and $T = \{d\infty\}$. However since the moduli interpretation of $X_1(p)$ is easier than that of $X_\mu(p)$, we instead apply $W_p^{(d)}$ so that we verify it for $S = S_0^{(d)}$ and $C = X_1(p)$ instead. One situation in which condition (3) of *Theorem* 3.3 is trivially satisfied is if $S = S_0^{(d)}$, $X_1(p)^{(d)}(\mathbb{F}_\ell) = \mathrm{red}_{\mathbb{F}_\ell}(S)$ and $T = f^{(d)}(S)$. For this it is useful to know $X_1(p)^{(d)}(\mathbb{F}_\ell)$. Let $y \in X_1(p)^{(d)}(\mathbb{F}_\ell)$, then y can be written as $\sum_{i=1}^m e_i y_i^{(f_i)}$ with $m, e_i, f_i \in \mathbb{N}_{\geq 0}$ and $y_i \in X_1(p)(\mathbb{F}_{\ell^{f_i}})$ such that each of the y_i does not come from a subfield of $\mathbb{F}_{\ell^{f_i}}$ and such that all the $y_i^{(f_j)}$ are distinct.

Theorem 5.9. *[Waterhouse, 1969, Thm 4.1] Let p, ℓ be distinct primes and d be an integer then*

$$Y_1(p)(\mathbb{F}_{\ell^d}) = \emptyset$$

if and only if the following 5 statements are true

1. *p does not divide any integer n such that both $|n - \ell^d - 1| < 2\ell^{d/2}$ and* $\gcd(n-1, \ell) = 1$.
2. *If d is even then $p \nmid \ell^d + 1 \pm 2\ell^{d/2}$.*

(3) *If d is even and $l \not\equiv 1 \mod 3$ then $p \nmid \ell^d + 1 \pm \ell^{d/2}$.*
(4) *If d is odd and $\ell = 2$ or 3 then $p \nmid \ell^d + 1 \pm \ell^{(d+1)/2}$.*
(5) *If d is odd or $l \not\equiv 1 \mod 4$ then $p \nmid \ell^d + 1$.*

and if (1) is false then all points in $Y_1(p)(\mathbb{F}_{\ell^d})$ are supersingular.

The theorem as stated above only follows from [Waterhouse, 1969, Thm 4.1] for $p > 4$ since for those primes the moduli problem for $Y_1(p)$ is representable over $\mathbb{Z}[1/p]$, but one easily verifies that $Y_1(p)(\mathbb{F}_{\ell^d}) \neq \emptyset$ and that statement 1 is false for $p = 2$ or 3.

If we again assume that $p > 4$ then $X_1(p)(\overline{\mathbb{Q}})$ has aside from the $(p-1)/2$ cusps defined over $\mathbb{Q}$, also $(p-1)/2$ cusps defined over the real subfield of $\mathbb{Q}(\zeta_p)$. The reduction of these $(p-1)/2$ non-rational cusps mod ℓ are definable over $\mathbb{F}_{\ell^d}$ if and only if $p \mid \ell^d - 1$ or $p \mid \ell^d + 1$. In particular the above theorem implies that $X_1(p)(\mathbb{F}_{\ell^{d'}}) = X_1(p)(\mathbb{F}_\ell) = \mathrm{red}_k X_1(p)(\mathbb{Q})$ holds for all $d' \leq d$ if $p > (\ell^{d/2}+1)^2$. Specializing to the case $\ell = 2$ and $3 \leq d \leq 7$ one can with a small computation for the primes $p < (\ell^{d/2}+1)^2$ show the following:

Proposition 5.10. *Let $3 \leq d \leq 7$ be an integer and p be a prime such that*

$$\begin{aligned} p \geq 11 \text{ and } p \neq 13, &\quad \text{if } d = 3,\\ p \geq 19, &\quad \text{if } d = 4,\\ p \geq 23 \text{ and } p \neq 31, 41, &\quad \text{if } d = 5,\\ p \geq 23 \text{ and } p \neq 29, 31, 37, 41, 73, &\quad \text{if } d = 6, \text{ and}\\ p = 47, 53 \text{ or } (p \geq 79 \text{ and } p \neq 113, 127), &\quad \text{if } d = 7. \end{aligned}$$

then

$$X_1(p)(\mathbb{F}_{2^{d'}}) = X_1(p)(\mathbb{F}_2) = \mathrm{red}_{\mathbb{F}_2}(X_1(p)(\mathbb{Q})) = \mathrm{red}_{\mathbb{F}_2}(S_0)$$

for all $d' \leq d$.

Corollary 5.11. *If one takes p, d as in the above proposition and one lets $S_0^{(d)}$ as in Eq.* (11)*, then*

$$X_1(p)^{(d)}(\mathbb{F}_2) = \mathrm{red}_{\mathbb{F}_2}(S_0^{(d)})$$

and hence condition (3) *of Theorem 3.2 holds for $C = X_1(p), S = S_0^{(d)}$ and $R = \mathbb{Z}_{(2)}$. Additionally condition* (3) *of Theorem 3.3 holds for $C = X_1(p), D = X_0(p), S = S_0^{(d)}, T = \{d0\}$ and $R = \mathbb{Z}_{(2)}$. By applying the Atkin-Lehner operator $W_p : X_1(p) \to X_\mu(p)$ condition* (3) *of Theorems 3.2 and 3.3 hold for $C = X_\mu(p), D = X_0(p), S = S_\infty^{(d)}, T = \{d\infty\}$ and $R = \mathbb{Z}_{(2)}$, where $S_\infty^{(d)}$ is as in Eq.* (7)*.*

6. Proof of Theorem 1.1

Proposition 6.1. *Let $d \leq 7$ be an integer. If $S(d) \subseteq$* Primes(2281)*, then $S(d) \subseteq$* Primes(193)*.*

Proof. It suffices to show that if $d \leq 7$ and $193 < p \leq 2281$ is a prime, then $p \notin S(d)$. This is done by applyingTheorem 3.3 with $C = X_\mu(p), D = X_0(p), S = S_\infty^{(d)}, T = \{d\infty\}$ and $R = \mathbb{Z}_{(2)}$. By Propositions 3.4, 5.5 and 5.7 we see that there exists a t such that conditions (1) and (2) are satisfied. By Corollary 5.11 we see that condition (3) is satisfied so we can indeed apply Theorem 3.3. It follows that $S_\infty^{(d)} = X_\mu(p)^{(d)}(\mathbb{Q})$, showing that the only points in $X_\mu(p)$ defined over a number field of degree $\leq d$ are cusps and hence $p \notin S(d)$. □

Proposition 6.2. *If* $S(d) \subseteq \text{Primes}(193)$ *for all* $d \leq 7$ *then*

$$\begin{aligned} S(3) &= \text{Primes}\,(17), \\ S(4) &= \text{Primes}\,(17) \cup \{29\}\,, \\ S(5) &= \text{Primes}\,(19) \cup \{29, 31, 41\}\,, \\ S(6) &= \text{Primes}\,(19) \cup \{29, 31, 37, 41, 73\} \qquad \textit{and} \\ S(7) &\subseteq \text{Primes}\,(43) \cup \{59, 61, 67, 71, 73, 113, 127\}\,. \end{aligned}$$

Proof. This is proven almost the same as Proposition 6.1, with the difference that this time one has to use Theorem 3.2 instead of Theorem 3.3, with C and S still $X_\mu(p)$ and $S_\infty^{(d)}$. Also Propositions 5.5 and 5.7 have to be replaced by Lemma 5.6 and Proposition 5.8. □

So in order to prove Theorem 1.1 it remains to deal with the primes $17, 29, 31, 41$ and 73.

6.1. **Proof of Theorem 1.1 for** $p = 17, 29, 31$ **or** 41**.** We quote [Conrad et al., 2003, Prop. 6.2.1.] in an equivalent formulation using that $J_1(p) \cong J_\mu(p)$ and adding some more information from Section 6.2 in loc.cit.

Proposition 6.3. *The primes p such that $J_\mu(p)$ has rank zero are the primes $p \leq 31$ and* 41*,* 47*,* 59*, and* 71*.*

For all of these, except possibly $p = 29$, the Mordell-Weil group is generated by differences of rational cusps, and for all except $p = 17$, 29, 31 and 41, the order of $J_1(p)(\mathbb{Q})$ is odd.

We can add to this the following new result.

Theorem 6.4. *The group $J_1(29)(\mathbb{Q})$ is generated by differences of rational cusps.*

Proof. Instead of proving this statement for $J_1(29)$ we will prove it for $J_\mu(29)$. This suffices because $X_1(N)$ and $X_\mu(N)$ are isomorphic over $\mathbb{Q}$ by an isomorphism that sends cusps to cusps. This allows us to use the description for the action of Galois on the cusps of $X_\mu(N)$ described in Stevens [1982]. It is already known that $J_\mu(29)(\mathbb{Q})[p^\infty]$ is generated by differences of rational cusps for all $p \neq 2$ prime (see the discussion after Conjecuture 6.2.2 of Conrad et al. [2003]). So it suffices to prove that $J_\mu(29)(\mathbb{Q})[2^\infty]$ is generated by the rational cusps.

Let $q \neq 2, 29$ be a prime then Proposition 5.2 implies that

$$J_\mu(29)(\mathbb{Q})[2^\infty] \subseteq J_\mu(29)(\overline{\mathbb{Q}})[2^\infty, T_q - \langle q \rangle - q].$$

Let $\tau : J_\mu(29)(\overline{\mathbb{Q}}) \to J_\mu(29)(\overline{\mathbb{Q}})$ be complex conjugation, then also

$$J_\mu(29)(\mathbb{Q})[2^\infty] \subseteq J_\mu(29)(\overline{\mathbb{Q}})[2^\infty, \tau - 1].$$

Using the isomorphism $J_\mu(29)(\overline{\mathbb{Q}})[2^\infty] \cong \varinjlim 2^{-i} H_1(X_\mu(29), \mathbb{Z})/H_1(X_\mu(29), \mathbb{Z})$ it is possible to compute the kernels of $\tau - 1$ and $T_q - \langle q \rangle - q$ seen as maps on $J_\mu(29)(\overline{\mathbb{Q}})[2^\infty]$ purely in terms of modular symbols. Let

$$M := J_\mu(29)(\overline{\mathbb{Q}})[2^\infty, T_5 - \langle 5 \rangle - 5, \tau - 1]$$

then a Sage computation shows that $M \cong (\mathbb{Z}/4\mathbb{Z})^6$. Let $C \subseteq J_\mu(29)(\mathbb{Q}(\zeta_{29}))$ be the subgroup generated by all cusps: using a Sage computation we showed $M = C[2^\infty]$. Using the explicit description of the action of $G := \mathrm{Gal}(\mathbb{Q}(\zeta_{29})/\mathbb{Q})$ on the cusps in Stevens [1982] we verified that $C[2^\infty]^G = J_\mu(29)(\mathbb{Q})[2^\infty]$ is indeed generated by the differences of rational cusps. □

This shows that for all primes p such that $J_1(p)(\mathbb{Q})$ is finite, the latter group is generated by differences of rational cusps. Now if $J_1(p)(\mathbb{Q})$ is finite and $J_1(p)(\mathbb{Q})[2] \hookrightarrow J_1(p)(\mathbb{F}_2)$ then condition (1) of Theorem 3.2 is satisfied for $t = \mathrm{Id}_{J_1(p)}$. For the primes $p = 3$, 5, 7, 11, 13, 19, 23, 47, 59 and 71, $J_1(p)(\mathbb{Q})[2] \hookrightarrow J_1(p)(\mathbb{F}_2)$ is trivially satisfied, since the group has odd order. Ironically the primes of Proposition 6.3 missing from this sequence are exactly the primes we are interested in.

Proposition 6.5. *For $p = 17, 29, 31$ or 41 one has $J_1(p)(\mathbb{Q})[2] \hookrightarrow J_1(p)(\mathbb{F}_2)$, and hence condition* (1) *of Theorem 3.2 is satisfied for $t = \mathrm{Id}_{J_1(p)}$.*

Proof. We only have to consider $p = 17$, 29, 31 and 41. We know that $J_1(p)(\mathbb{Q})$ is generated by differences of rational cusps, see Proposition 6.3 and Theorem 6.4. It is also known what the order of this group is, see [Conrad et al., 2003, § 6.2.3 and Table 1]. We now use Magma Bosma et al. [1997] to compute a model of $X_1(p)$ over $\mathbb{F}_2$ and check that the subgroup of its Picard group generated by differences of its $\mathbb{F}_2$-points (which are the images of the cusps under reduction mod 2) has the correct order. In fact, it suffices to check that the 2-primary part of the group has the correct order. For $p = 17$, we do this directly. For the other three primes, we use an intermediate curve X_H such that the predicted order of the 2-primary part of $J_H(\mathbb{Q})$ equals that of $J_1(p)(\mathbb{Q})$, since the computation using $X_1(p)$ directly would be too involved. We check that the subgroup of $J_H(\mathbb{F}_2)$ generated by differences of the images of cusps has 2-primary part of the correct size. For $p = 29$, we use X_H corresponding to $d = 7$ in the notation of Conrad et al. [2003], for $p = 31$, we use the curve corresponding to $d = 3$, and for $p = 41$ we use the curve corresponding to $d = 4$. In each case, the computation gives the desired result. (It is also possible and not taking too much time to do the computation directly on $X_1(p)$ over $\mathbb{F}_2$ for $p = 29$ and $p = 31$.) □

Lemma 6.6. *Condition* (3) *of Theorem 3.2 is satisfied for* $C = X_1(29)$ *and* $C = X_1(31)$*,* $S = S_0^{(d)}$*,* $R = \mathbb{Z}_{(2)}$ *and* $d \leq 7$*. Here* $S_0^{(d)}$ *is as in Eq.* (11)*.*

Proof. For $d < 5$, this is covered by Proposition 5.10. For $d = 5, 6, 7$, we check it by a Magma calculation. In this calculation we computed the images in $\mathrm{Pic}_{C_{\mathbb{F}_2}/\mathbb{F}_2}(\mathbb{F}_2)$ of all points $s \in C^{(d)}(\mathbb{F}_2)$ not coming from a point in $S_0^{(d)}$. We verified that these images are not in the subgroup of $\mathrm{Pic}_{C_{\mathbb{F}_2}/\mathbb{F}_2}(\mathbb{F}_2)$ generated by the points coming from $\mathbb{Q}$-rational cusps, and we know that the $\mathbb{Q}$ rational cups generate $\mathrm{Pic}_{C_{\mathbb{Q}}/\mathbb{Q}}(\mathbb{Q})$ for these two curves by Proposition 6.3 and Theorem 6.4. □

The above proof involves computing $\mathrm{Pic}_{C_{\mathbb{F}_2}/\mathbb{F}_2}(\mathbb{F}_2)$ in Magma. For $C = X_1(41)$ this would probably take too long to be practical. Therefore we deal with $C = X_1(41)$ in a slightly different way:

Lemma 6.7. *Condition* (3) *of Theorem 3.2 is satisfied for* $C = X_1(41)$*,* $S = S_0^{(d)}$*,* $R = \mathbb{Z}_{(2)}$ *and* $d \leq 7$*.*

Proof. There is no elliptic curve E over $\mathbb{F}_{2^e}$ with $41 \mid \#E(\mathbb{F}_{2^e})$ if $e = 1, 2, 3, 4, 6$ or 7. There is exactly one elliptic curve E over $\mathbb{F}_{2^5}$ with $\#E(\mathbb{F}_{2^5}) = 41$; this is the curve $y^2 + y = x^3 + x + 1$ already defined over $\mathbb{F}_2$. Its automorphism group over $\mathbb{F}_{2^5}$ is cyclic of order 4; we therefore obtain only $10 = (41 - 1)/4$ distinct $\mathbb{F}_{2^5}$-points on $X_1(41)$ that are not cusps. Let X_H be the intermediate curve corresponding to $d = 4$ in Conrad et al. [2003]. Then $X_1(41) \to X_H$ is an étale cover of degree 5, and the ten $\mathbb{F}_{2^5}$-points on $X_1(41)$ map to two $\mathbb{F}_2$-points on X_H. In fact, $X_H(\mathbb{F}_2)$ consists of six points; four of them are cusps, and the other two are the ones just mentioned. It can be checked that these two points do not map into the subgroup of $\mathrm{Pic}_{X_{H,\mathbb{F}_2}/\mathbb{F}_2}(\mathbb{F}_2)$ generated by the four cusps, which implies condition (3). □

Proposition 6.8. *The following exclusions hold:*

$$\begin{aligned} 17 &\notin S(3), \\ 29 &\notin S(4), \\ 29, 31, 41 &\notin S(5), \\ 29, 31, 41 &\notin S(6) \quad \textit{and} \\ 29, 31, 41 &\notin S(7). \end{aligned}$$

The proof of $17 \notin S(3)$ is similar to that in Parent [2003] although we manage to avoid the careful analysis of the formal group of $J_1(p)_{\mathbb{Z}_2}$ since we have proven that $J_1(p)(\mathbb{Q})[2] \hookrightarrow J_1(p)(\mathbb{F}_2)[2]$ in Proposition 6.5.

Proof. This is again done by applying Theorem 3.2 over $R = \mathbb{Z}_{(2)}$, this time with $C = X_1(p)$ and $S = S_0^{(d)}$ for the p, d for which we want to show $p \notin S(d)$. We check that Theorem 3.2 can indeed be applied by verifying that its conditions (1),(2) and (3) are satisfied using $t = \mathrm{Id} : J_1(p) \to J_1(p)$.

(1) This follows from Proposition 6.5.
(2) For $(p,d) = (17,3)$ this is in [Parent, 2000, §4.3].
For $p = 29$ resp. 31 it is known that the $\mathbb{F}_2$ gonality of $X_1(p)$ is 11 resp. 12 [Derickx and van Hoeij, 2014, Tbl. 1, Rmk. 1]. So condition (2) is satisfied by Proposition 3.5.
For $p = 41$ this follows from Proposition 5.8 together with Lemma 5.6 using the isomorphism $W_p : X_\mu(p) \to X_1(p)$.
(3) For $p = 17$ this is Corollary 5.11, for $p = 29, 31, 41$ it follows from Lemmas 6.6 and 6.7.

□

This leaves us with only one case which we also found the hardest to prove.

6.2. **Proof of Theorem 1.1 for $p = 73$.** First we start by analysing the points in $X_1(73)(\mathbb{F}_2^d)$ for $d \leq 6$. The first thing to notice is that for $d \leq 6$ the only points in $X_1(73)(\mathbb{F}_{2^d}) \setminus Y_1(73)(\mathbb{F}_{2^d})$ are the points mapping to the cusp 0 on $X_0(73)$, because $2^d \not\equiv \pm 1 \mod 73$ for $d \leq 6$. Using the isomorphism $W_p : X_1(p) \to X_\mu(p)$ and applying Lemma 5.6 and Proposition 5.8 shows that the conditions of Lemma 3.1 are satisfied for all cuspidal points of $X_1(73)^{(6)}(\mathbb{F}_2)$. As a result we only need to study the residue classes in $X_1(73)^{(6)}(\mathbb{F}_2)$ that do not consist entirely of cusps. After a detailed study of these residue classes the proof will be finished by Proposition 6.9.

We continue by analysing the points of $X_1(73)^{(6)}(\mathbb{F}_2)$ that do not consist completely of cusps. For this we first describe the Tate normal form see Knapp [1992] of a point $(E,P) \in Y_1(N)(K)$ for K a field and $N \geq 4$ an integer coprime to the characteristic to K. For every pair (E,P) where E is an elliptic curve over K and P a point of order exactly N there are unique $b, c \in K$ such that $(E,P) \cong (E_{b,c},(0,0))$ where $E_{b,c}$ is the elliptic curve given by the Weierstrass equation

$$y^2 + (1-c)xy - by = x^3 - bx^2. \tag{12}$$

By Theorem 5.9 one sees that there are no points in $Y_1(73)(\mathbb{F}_{2^d})$ for $d \leq 5$ and that all points in $Y_1(73)(\mathbb{F}_{2^6})$ are supersingular. To explicitly find the Tate normal form of all points in $Y_1(73)(\mathbb{F}_{2^6})$ note that $E_{b,c}$ has discriminant $\Delta_{b,c} := b^3(c^4+c^3+c^2+b+c)$ and j-invariant $\frac{(c+1)^{12}}{\Delta_{b,c}}$ in characteristic 2. The curve $E_{b,c}$ is supersingular if and only if $j = 0$, which is equivalent to $c = 1$. By computing the 73 division polynomial for $E_{b,1}$ one sees that the solutions of

$$(b^6+b+1)(b^6+b^3+1)(b^6+b^5+b^2+b+1)(b^6+b^5+b^4+b+1) \tag{13}$$

are exactly the values of $b \in \mathbb{F}_{2^6}$ such that $(0,0)$ is of order 73. This calculation shows that $X_1(73)^{(6)}(\mathbb{F}_2)$ has exactly 4 points that do not consist entirely of cusps, namely the points corresponding to the 4 factors of (13). Explicitly calculating the action of $(\mathbb{Z}/73\mathbb{Z})^*/\{\pm 1\}$ on these 4 points one can show that the diamond operator $\langle 10 \rangle$ of order 4 acts transitively on them. Let $H \subseteq (\mathbb{Z}/N\mathbb{Z})^*/\{\pm 1\}$ be the subgroup generated by 10, then 4 points in $Y_1(73)^{(6)}(\mathbb{F}_2)$ map to a single point on $Y_H^{(6)}(\mathbb{F}_2)$ by the discussion above.

If E is an elliptic curve with $73 = 2^6+1+8$ points over $\mathbb{F}_{2^6}$ then the characteristic polynomial of frobenius is

$$x^2 - 8x - 2^6 = (x - 8\zeta_3)(x + 8\zeta_3 + 8).$$

Let E_{ζ_3} be an elliptic curve $\mathbb{Q}(\zeta_3)$ that has complex multiplication by $\mathbb{Q}(\zeta_3)$, then E_{ζ_3} has two isogenies of degree 73 over $\mathbb{Q}(\zeta_3)$ namely $8\zeta_3 - 1$ and $-8\zeta_3 - 9$. The map $X_1(73) \to X_0(73)$ is of degree $36 = (73-1)/2$, and since the automorphism ζ_3 of order 3 preserves the kernels of the isogenies $8\zeta_3-1$ and $-8\zeta_3-9$ we see that the ramification index of $\pi : X_1(73) \to X_0(73)$ at the points corresponding to the isogenies $8\zeta_3 - 1$ and $-8\zeta_3 - 9$ is 3. Showing that $S := \pi^{-1}(\{(E, 8\zeta_3 - 1), (E, -8\zeta_3 - 9)\}) \subseteq X_1(73)(\overline{\mathbb{Q}})$ is a set of size 24. The action of Galois on S is transitive because there are no CM elliptic curves with a 73 torsion point over number fields of degree < 24 [Clark et al., 2013, Table 1]. If one fixes a prime ℓ above 2 in $\overline{\mathbb{Q}}$, then reduction modulo ℓ gives a bijection between S and $Y_1(73)(\mathbb{F}_{2^6})$. The existence of this bijection can be shown either by explicit computation in Sage or by pure thought by showing that for $(E,P) \in Y_1(73)^{(6)}(\mathbb{F}_2)$ the canonical lift (or Deuring lift) (E_0, ϕ_0) of $(E, \mathrm{Frob}_{\mathbb{F}_{2^6}}/8)$ to $\bar{\mathbb{Q}}$ is either (E, ζ_3) or its Galois conjugate $(E, -\zeta_3 - 1)$.

The above discussion shows that if one takes $x_1, \ldots, x_6 \in X_H(\bar{\mathbb{Q}})$ to be the 6 points corresponding to the 6 orbits of $\langle 10 \rangle$ acting on S, that then

$$x^{(6)} : x_1 + \cdots + x_6 \in X_H^{(6)}(\mathbb{Q}) \tag{14}$$

is a point that reduces to the unique point in the image of $Y_1(73)^{(6)}(\mathbb{F}_2) \to Y_H^{(6)}(\mathbb{F}_2)$.

Since $x^{(6)}$ corresponds to a CM curve and CM curves over number fields of degree < 24 have no 73 torsion as mentioned before, and we know that a point in $y \in X_H^{(6)}(\mathbb{Q})$ coming from $X_1(73)^{(6)}(\mathbb{Q})$ has to specialize to $x_{\mathbb{F}_2}^{(6)}$ we can prove that $73 \notin S(6)$ by showing:

Proposition 6.9. *Let $H \subseteq (\mathbb{Z}/73\mathbb{Z})^*/\{\pm 1\}$ the subgroup generated by 10. Then the point $x^{(6)}$ defined above is the unique point in $X_H^{(6)}(\mathbb{Q})$ reducing to $x_{\mathbb{F}_2}^{(6)}$ modulo 2.*

Proof. We do this by proving instead that $W_p^{(6)}(x^{(d)}) \in X_{\mu,H}^{(6)}(\mathbb{Q})$ is the unique point reducing to $W_p^{(6)}(x_{\mathbb{F}_2}^{(d)})$. This allows us to work with a model where the cusp at infinity is rational. We are going to prove that the matrix A of Proposition 3.7 at $W_p^{(6)}(x_{\mathbb{F}_2}^{(d)})$ has rank 6 using an explicit model of $X_{\mu,H,\mathbb{F}_2}$, we know that its genus is 43. Using Sage to compute an explicit basis of $H^0(X_{\mu,H,\mathbb{F}_2}, \Omega^1) = S_2(\Gamma_H, \mathbb{F}_2)$ shows that q^{47} is the largest leading term among all modular forms. So giving the coefficients of a modular form up to and including q^{47} is enough to determine it uniquely. The subspace $H^0(X_{\mu,H,\mathbb{F}_2}, \Omega^1(-41\infty)) \subseteq H^0(X_{\mu,H,\mathbb{F}_2}, \Omega^1)$ is 3 dimensional and has as basis

$$\begin{aligned}\omega_1 &:= (q^{42}+q^{47}+q^{49}+O(q^{50}))\tfrac{\mathrm{d}q}{q}\\ \omega_2 &:= (q^{43}+q^{49}+O(q^{50}))\tfrac{\mathrm{d}q}{q}\\ \omega_3 &:= (q^{47}+q^{48}+O(q^{50}))\tfrac{\mathrm{d}q}{q}.\end{aligned}$$

Let $\mathcal{L} \subseteq \Omega^1_{X_{\mu,H,\mathbb{F}_2}}$ be the line bundle generated by $\omega_1, \omega_2, \omega_3$ then $\mathcal{L}$ has degree at most $2 \cdot 43 - 2 - 41 = 43$. Viewing $\omega_1, \omega_2, \omega_3$ as sections of $\mathcal{L}$ gives us a map $\phi : X_{\mu,H,\mathbb{F}_2} \to \mathbb{P}^2_{\mathbb{F}_2}$ given by $\phi(P) = (\omega_1(P) : \omega_2(P) : \omega_3(P))$. Its image is given by a homogeneous polynomial of degree at most 43. Indeed, using the computer to compare the q-expansions of products of ω_1, ω_2 and ω_3 we found a homogeneous polynomial $f_H \in \mathbb{F}_2[x_0, x_1, x_2]$ of degree 41 describing the image of ϕ, since this is only 2 smaller then expected we know that $\Omega^1_{X_{\mu,H,\mathbb{F}_2}}/\mathcal{L}$ is an effective divisor $\mathbb{F}_2$ of degree 2, in particular there are no points with residue field $\mathbb{F}_{2^6}$ in its support, meaning that at least one of $\omega_1, \omega_2, \omega_3$ is a generator of $\Omega^1_{X_{\mu,H,\mathbb{F}_2}}$ at the points we are interested in. The polynomial f_H takes about two pages in LaTeX so we did not include it here, but we could use Sage to compute with it. Let C_H be the curve with equation f_H, using Sage we computed its geometric genus. Its genus turned out to be 43, so we know it has to be birational to $X_{\mu,H,\mathbb{F}_2}$.

The next step is to find the points in $X_{\mu,H}(\mathbb{F}_{2^6})$ that are supersingular, for this we use the Hasse invariant A_2, it is a modular form of weight 1 over $\mathbb{F}_2$ whose zero's are exactly the supersingular curves and its q expansions is $1 \in \mathbb{F}_2[q]$. Using Magma we listed all points with residue field $\mathbb{F}_{2^6}$ on the desingularisation of $\operatorname{im}\phi \subset \mathbb{P}^2_{\mathbb{F}_2}$ none of these points had 0 as their 3-th coordinate. So we know that $g := A_2^2/\omega_3$ is a function on $X_{\mu,H,\mathbb{F}_2}$ which has a zero at all the supersingular points in $X_{\mu,H}(\mathbb{F}_{2^6})$, comparing q-expansions we found two homogeneous polynomials $g^{num}, g^{den} \in \mathbb{F}_2[x_0, x_1, x_2]$ of degree 40 such that

$$A_2^2 g^{den}(\omega_1, \omega_2, \omega_3) = \omega_3 g^{num}(\omega_1, \omega_2, \omega_3),$$

so that $g = g^{num}/g^{den}$. Choose a $c \in \mathbb{F}_{2^6}$ such that $c^6 + c^5 + 1 = 0$. By looking at the zero's of g we found that, up to relabeling, the points

$$x_i := (0 : c^{2^{i-1}} : 1) \in (\operatorname{im}\phi)(\mathbb{F}_{2^6}) \subset \mathbb{P}^2(\mathbb{F}_{2^6}), \quad 1 \le i \le 6$$

correspond to the points $x_1, \ldots, x_6$ of Eq. (14). Define $T = (T_3 - \langle 3 \rangle - 3)t_1(T_5)$ where t_1 is as in Proposition 5.3. Then T is as in Proposition 3.4. The matrix of T when seen as acting on $S_2(\Gamma_H, \mathbb{F}_2)$ was seen to be of rank 39 showing that the dimension of $T^*(\mathrm{Cot}_0\, J_{\mu,H,\mathbb{F}_2})$ is 39, providing good hope that we can find ω_i such that the matrix A of Proposition 3.7 has rank 6.

Define

$$\begin{aligned}
\omega_1' &:= (q^{40} + q^{41} + q^{46} + O(q^{48}))\tfrac{\mathrm{d}q}{q}\\
\omega_2' &:= (q^{37} + q^{43} + O(q^{48}))\tfrac{\mathrm{d}q}{q}\\
\omega_3' &:= (q^{36} + q^{38} + q^{39} + q^{41} + q^{46} + q^{47} + O(q^{48}))\tfrac{\mathrm{d}q}{q}\\
\omega_4' &:= (q^{34} + q^{39} + q^{43} + q^{44} + q^{45} + O(q^{48}))\tfrac{\mathrm{d}q}{q}\\
\omega_5' &:= (q^{33} + q^{39} + q^{45} + O(q^{48}))\tfrac{\mathrm{d}q}{q}\\
\omega_6' &:= (q^{32} + q^{41} + q^{44} + q^{46} + q^{47} + O(q^{48}))\tfrac{\mathrm{d}q}{q}
\end{aligned}$$

Let q_j be a uniformizer at x_j such that and write $\omega_3 = f_j\mathrm{d}q_j$ and $\omega_i' = f_{i,j}\mathrm{d}q_j$. Then the coefficient $a(\omega_i', q_j, 1)$ of the matrix A is just $f_{i,j}(0)$. If we view $g_i := \omega_i'/\omega_3$ as a function on $X_{\mu,H,\mathbb{F}_2}$ then because as we saw earlier that $f_j(0) \neq 0$ we see that g_i does not have a pole at x_j and $g_i(x_j) = f_{i,j}(0)/f_j(0)$. The rank does not change if we scale the q_j'th row by $f_j(0)$ so the rank of the matrix A is the same as that of $(g_i(x_j))_{i,j=1}^6$. Comparing q-expansions like we did to write $g = g^{num}/g^{den}$ we again managed to find the function g_i explicitly on our model, allowing us to compute

$$(g_i(x_j))_{i,j=1}^6 := \begin{pmatrix} c^{46} & c^{29} & c^{58} & c^{53} & c^{43} & c^{23}\\ c^{14} & c^{28} & c^{56} & c^{49} & c^{35} & c^{7}\\ c^{8} & c^{16} & c^{32} & c & c^{2} & c^{4}\\ c^{35} & c^{7} & c^{14} & c^{28} & c^{56} & c^{49}\\ c & c^{2} & c^{4} & c^{8} & c^{16} & c^{32}\\ c^{5} & c^{10} & c^{20} & c^{40} & c^{17} & c^{34} \end{pmatrix}.$$

The fact that each column is the square of the previous column is explained by

$$g_i(x_j)^2 = \mathrm{Frob}_2(g_i(x_j)) = g_i(\mathrm{Frob}_2(x_j)) = g_i(x_j).$$

The determinant of the above matrix is 1 showing that the map

$$T \circ f_{6,x^{(6)}_{\mathbb{F}_2}} : X^{(6)}_{\mu,H,\mathbb{F}_2} \to J_{\mu,H,\mathbb{F}_2}$$

is a formal immersion at $x^{(6)}_{\mathbb{F}_2}$. So we can apply Lemma 3.1 to get the proposition. ☐

References

Siegfried Bosch, Werner Lütkebohmert, and Michel Raynaud. *Néron models*, volume 21 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1990. ISBN 3-540-50587-3. doi: 10.1007/978-3-642-51438-8. URL `http://dx.doi.org/10.1007/978-3-642-51438-8`.

Wieb Bosma, John Cannon, and Catherine Playoust. The Magma algebra system. I. The user language. *J. Symbolic Comput.*, 24(3-4):235–265, 1997. ISSN 0747-7171. doi: 10.1006/jsco.1996.0125. URL `http://dx.doi.org/10.1006/jsco.1996.0125`. Computational algebra and number theory (London, 1993).

Pete L. Clark, Brian Cook, and James Stankewicz. Torsion points on elliptic curves with complex multiplication (with an appendix by Alex Rice). *Int. J. Number Theory*, 9(2):447–479, 2013. ISSN 1793-0421. doi: 10.1142/S1793042112501436. URL http://dx.doi.org/10.1142/S1793042112501436.

Brian Conrad, Bas Edixhoven, and William Stein. $J_1(p)$ has connected fibers. *Doc. Math.*, 8:331–408, 2003. ISSN 1431-0635.

Maarten Derickx and Mark van Hoeij. Gonality of the modular curve $X_1(N)$. *J. Algebra*, 417:52–71, 2014. ISSN 0021-8693. doi: 10.1016/j.jalgebra.2014.06.026. URL http://dx.doi.org/10.1016/j.jalgebra.2014.06.026.

Fred Diamond and John Im. Modular forms and modular curves. In *Seminar on Fermat's Last Theorem (Toronto, ON, 1993–1994)*, volume 17 of *CMS Conf. Proc.*, pages 39–133. Amer. Math. Soc., Providence, RI, 1995.

Fred Diamond and Jerry Shurman. *A first course in modular forms*, volume 228 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2005. ISBN 0-387-23229-X.

Noam D. Elkies. Elliptic and modular curves over finite fields and related computational issues. In *Computational perspectives on number theory (Chicago, IL, 1995)*, volume 7 of *AMS/IP Stud. Adv. Math.*, pages 21–76. Amer. Math. Soc., Providence, RI, 1998.

Daeyeol Jeon, Chang Heon Kim, and Yoonjin Lee. Families of elliptic curves over cubic number fields with prescribed torsion subgroups. *Math. Comp.*, 80(273): 579–591, 2011a. ISSN 0025-5718. doi: 10.1090/S0025-5718-10-02369-0. URL http://dx.doi.org/10.1090/S0025-5718-10-02369-0.

Daeyeol Jeon, Chang Heon Kim, and Yoonjin Lee. Families of elliptic curves over quartic number fields with prescribed torsion subgroups. *Math. Comp.*, 80(276): 2395–2410, 2011b. ISSN 0025-5718. doi: 10.1090/S0025-5718-2011-02493-2. URL http://dx.doi.org/10.1090/S0025-5718-2011-02493-2.

S. Kamienny. Torsion points on elliptic curves over fields of higher degree. *Internat. Math. Res. Notices*, (6):129–133, 1992a. ISSN 1073-7928. doi: 10.1155/S107379289200014X. URL http://dx.doi.org/10.1155/S107379289200014X.

S. Kamienny. Torsion points on elliptic curves and q-coefficients of modular forms. *Invent. Math.*, 109(2):221–229, 1992b. ISSN 0020-9910. doi: 10.1007/BF01232025. URL http://dx.doi.org/10.1007/BF01232025.

Kazuya Kato. p-adic Hodge theory and values of zeta functions of modular forms. *Astérisque*, (295):ix, 117–290, 2004. ISSN 0303-1179. Cohomologies p-adiques et applications arithmétiques. III.

Anthony W. Knapp. *Elliptic curves*, volume 40 of *Mathematical Notes*. Princeton University Press, Princeton, NJ, 1992. ISBN 0-691-08559-5.

V. A. Kolyvagin and D. Yu. Logachëv. Finiteness of the Shafarevich-Tate group and the group of rational points for some modular abelian varieties. *Algebra i Analiz*, 1(5):171–196, 1989. ISSN 0234-0852.

B. Mazur. Modular curves and the Eisenstein ideal. *Inst. Hautes Études Sci. Publ. Math.*, (47):33–186 (1978), 1977. ISSN 0073-8301. URL `http://www.numdam.org/item?id=PMIHES_1977__47__33_0`.

B. Mazur. Rational isogenies of prime degree (with an appendix by D. Goldfeld). *Invent. Math.*, 44(2):129–162, 1978. ISSN 0020-9910. doi: 10.1007/BF01390348. URL `http://dx.doi.org/10.1007/BF01390348`.

Loïc Merel. Bornes pour la torsion des courbes elliptiques sur les corps de nombres. *Invent. Math.*, 124(1-3):437–449, 1996. ISSN 0020-9910. doi: 10.1007/s002220050059. URL `http://dx.doi.org/10.1007/s002220050059`.

Pierre Parent. Bornes effectives pour la torsion des courbes elliptiques sur les corps de nombres. *J. Reine Angew. Math.*, 506:85–116, 1999. ISSN 0075-4102. doi: 10.1515/crll.1999.009. URL `http://dx.doi.org/10.1515/crll.1999.009`.

Pierre Parent. Torsion des courbes elliptiques sur les corps cubiques. *Ann. Inst. Fourier (Grenoble)*, 50(3):723–749, 2000. ISSN 0373-0956. URL `http://www.numdam.org/item?id=AIF_2000__50_3_723_0`.

Pierre Parent. No 17-torsion on elliptic curves over cubic number fields. *J. Théor. Nombres Bordeaux*, 15(3):831–838, 2003. ISSN 1246-7405. URL `http://jtnb.cedram.org/item?id=JTNB_2003__15_3_831_0`.

The Developers of Sage. *SageMath, the Sage Mathematics Software System (Version 6.4)*, 2014. `http://www.sagemath.org`.

Samir Siksek. Chabauty for symmetric powers of curves. *Algebra Number Theory*, 3(2):209–236, 2009. ISSN 1937-0652. doi: 10.2140/ant.2009.3.209. URL `http://dx.doi.org/10.2140/ant.2009.3.209`.

Glenn Stevens. *Arithmetic on modular curves*, volume 20 of *Progress in Mathematics*. Birkhäuser Boston, Inc., Boston, MA, 1982. ISBN 3-7643-3088-0.

William C. Waterhouse. Abelian varieties over finite fields. *Ann. Sci. École Norm. Sup. (4)*, 2:521–560, 1969. ISSN 0012-9593.

Mathematisch Instituut, Universiteit Leiden, P.O. Box 9512, 2300 RA Leiden, The Netherlands
E-mail address: `maarten@mderickx.nl`

University of Southern California, 3620 South Vermont Ave., KAP 108, Los Angeles, California 90089-2532, USA
E-mail address: `kamienny@usc.edu`

University of Washington, Department of Mathematics, Box 354350, Seattle, WA 98195-4350, USA.
E-mail address: `wstein@gmail.com`

Mathematisches Institut, Universität Bayreuth, 95440 Bayreuth, Germany.
E-mail address: `Michael.Stoll@uni-bayreuth.de`

APPENDIX A: OESTERLÉ'S BOUND

A.1. **Introduction.** The goal of this appendix is to publish a proof of the following well known theorem.

Theorem A.1 (Oesterlé, 1994, unpublished)**.** *Let $K/\mathbb{Q}$ be a number field of degree d, E/K an elliptic curve and $P \in E(K)$ a point of prime order p then*

$$p \leq (3^{d/2} + 1)^2.$$

Loïc Merel already proved this theorem in 1994 with a bound of d^{3d^2}, published in [Merel, 1996]. Shortly after Merel, Joseph Oesterlé proved the theorem above for $(d, p) \neq (3, 43)$ and in fact Oesterlé's improvement is already announced in Merel's article. The case case $(d, p) = (3, 43)$) was later dealt with in [Parent, 2000]. This appendix closely follows Oesterlé's notes which he made available to the first author, although this appendix contains some minor simplifications using literature which didn't exist in 1994. The better bound of Oesterlé is an essential starting point in order to make the explicit computations in the article to which this Appendix is attached possible. Conversely, because of Theorem 1.1 of the main text and the results that $p \leq 7$ if $d = 1$ of [Mazur, 1977] and $p \leq 13$ if $d = 2$ of [Kamienny, 1992b] it suffices to prove the following weaker theorem:

Theorem A.2. *Let $K/\mathbb{Q}$ be a number field of degree d, E/K an elliptic curve and $P \in E(K)$ a point of prime order p then:*

(1) $p \leq (3^{d/2} + 1)^2$ *if* $d \geq 6$.
(2) $p < 410$ *if* $d = 3, 4$ *or* 5.

Actually, in his notes Oesterlé first establishes Theorem A.2, and then later goes on to prove Theorem A.1 for $(d, p) \neq (3, 43)$ using a comparable but slightly different strategy. The section of his notes where Oesterlé proves Theorem A.1 for $d = 3, 4, 5$, $p < 410$ and $(d, p) \neq (3, 43)$ contains no surprising new techniques. This section is omitted since it is covered by the computations in the main text.

Several of Oesterlé's ideas can already be found in the literature, since Pierre Parent generalized several of his ideas to points on elliptic curves whose order is a prime power in [Parent, 1999]. In fact, Theorem A.2 for $d > 25$ is an easy corollary of the following theorem, as will be shown in section A.2.

Theorem A.3. *[Parent, 1999, Thm. 1.6] Let E be an elliptic curve over a number field K of degree d over $\mathbb{Q}$ possessing a K-rational point P of prime power order p^n. Let l be prime different from 2^* and p. Suppose that for every prime ideal ℓ of O_K*

*Parent only mentions the condition $l \neq p$ in his Theorem and not $l \neq 2$. However he mentions it at the beginning of §1.3 and this condition is necessary for his proof of this Theorem to work.

dividing l one has that E has split multiplicative reduction and that P has order p^n in the component group of the Néron model of E, then

$$p^n < 65(2d)^6 \text{ if } p > 2 \quad \text{and} \quad p^n < 129(3d)^6 \text{ if } p = 2.$$

But not all ideas of Oesterlé were generalized by Parent. The main ingredients that are not yet in the literature are the intersection formulas in sections A.5.3 and A.5.4.

Note that from the work of Parent it is also possible to deduce a version of A.2 with the weaker bound $p < 65(2d)^6$ for $d \leq 25$. However the results of the main text would have been very difficult to obtain starting from this weaker bound (although maybe not impossible), since it would require significantly more computer computations as the following table indicates.

d	4	5	6	7	25
$\lfloor (3^{d/2}+1)^2 \rfloor$	100	275	784	2281	847×10^9
$65(2d)^6$	17,039,360	65,000,000	194,088,960	489,419,840	$1,015 \times 10^9$

Acknowledgements. I would like to thank Tessa Schild for her proofreading, Loïc Merel for his help explaining Oesterlé's notes and Bas Edixhoven for his useful discussions and his detailed reading of this text. But most of all I would like to thank Joseph Oesterlé for his help in understanding his proof and giving me his notes. I also want to thank him for allowing me to use it to write this appendix. The two sections A.5.3 and A.5.4 and section A.6 are a translation into English of Oesterlé's notes, where I added some details making claims easier to verify for the reader and replaced certain arguments by references. The rest of this article is a summary of needed background theory and results already present in the literature, much of which was also already in Oesterlé's notes.

A.2. **Proof of Theorem A.2 for $d > 25$.** To be able to use [Parent, 1999, Thm. 1.6] we first have to check whether its conditions are satisfied. This means we first need to prove the following proposition which is similar to Proposition 1.4 of [Parent, 1999].

Proposition A.4. *Let $K/\mathbb{Q}$ be a number field of degree d, E/K an elliptic curve with Néron model $\tilde{E}$ and $P \in E(K)$ a point of prime order p. If $p > (3^{d/2}+1)^2$ then $\tilde{E}$ has split multiplicative reduction at all primes ℓ of $\mathcal{O}_K$ dividing 3 and $\tilde{P}_{\mathcal{O}_K/\ell}$ does not lie in the identity component of $\tilde{E}_{\mathcal{O}_K/\ell}$.*

Remark. The map $X_0(p) \to X_0(1)$ is unramified at the cusp ∞ and ramified of order p at the cusp 0 see [Mazur, 1977, p. 64], so one sees that because $\tilde{P}_{\mathcal{O}_K/\ell}$ lies in a component that is not the identity implies that the pair $(\tilde{E}_{\mathcal{O}_K/\ell}, \langle \tilde{P}_{\mathcal{O}_K/\ell} \rangle)$ has to be the cusp 0 of $X_0(p)$ [Deligne and Rapoport, 1975, VII, §2]. This however is inconsistent with the modular interpretation of the cusps on page 159 of [Mazur, 1977]. The description of the cusps in [Deligne and Rapoport, 1975, VII, §2] shows that moduli interpretation of the unramified cusp of $X_0(p)$ should be a Néron 1-gon and that of the ramified cusp a Néron p-gon. Luckily this mistake does not affect the main results of [Mazur, 1977] since one can apply the Atkin-Lehner operator W_p to swap the cusps 0 and ∞. This mistake also propagated to works that cite Mazur his article, among for example [Kamienny, 1992a,b, Kamienny and Mazur, 1995], the first author has notified Kamienny and Mazur of this mistake and an erratum is being written.

Proof. Let ℓ be a prime ideal of $\mathcal{O}_K$ dividing 3 and k be its residue field. We want to rule out all types of reduction except split multiplicative where $\tilde{P}_k$ does not lie in the identity component.

The first thing to notice is that $p > (3^{1/2}+1)^2 > 3 = \operatorname{char} k$. This means that the map $\tilde{E}[p](\mathcal{O}_K) \to \tilde{E}[p](k)$ is injective and in particular that $\tilde{P}_k \in \tilde{E}[p](k)$ has order p.

- $\tilde{E}$ does not have good reduction at ℓ, because if it has good reduction, then $\tilde{E}_k$ is an elliptic curve and hence the Hasse bound gives
$$\#\tilde{E}(k) \leq (\sqrt{\#k}+1)^2 \leq (3^{d/2}+1)^2$$
which clearly contradicts that $\tilde{E}(k)$ has a point of order $p > (3^{d/2}+1)^2$.
- $\tilde{E}$ does not have additive reduction at ℓ. This is because additive reduction means that we have an exact sequence:
$$\mathbb{G}_a(k) \to \tilde{E}(k) \to \phi(k)$$
where ϕ is the component group of $\tilde{E}_k$. This means that either $\tilde{P}_k$ lies in the image of $\mathbb{G}_a(k)$, in which case $p = 3$ or $p \mid \#\phi(k) \leq 4$, with both possibilities leading to a contradiction with $p > (3^{1/2}+1)^2 > 7$.
- $\tilde{E}$ does not have non-split multiplicative reduction at ℓ. This is because this would mean that we have an exact sequence
$$\tilde{\mathbb{G}}_{m,k}(k) \to \tilde{E}(k) \to \phi(k)$$
$\tilde{\mathbb{G}}_{m,k}$ is the quadratic twist of the multiplicative group over k. In this case either
$$p | \#\tilde{\mathbb{G}}_m(k) = \#k + 1 < (3^{d/2}+1)^2 \text{ or } p | \#\phi(k) \leq 2,$$
with both possibilities leading to a contradiction with $p > (3^{d/2}+1)^2$.
- If $\tilde{E}$ has split multiplicative reduction, then $\tilde{P}_k$ cannot lie in the identity component of $\tilde{E}_k$. This is because the identity component is isomorphic to $\mathbb{G}_m$ and $\#\mathbb{G}_m(k) = \#k - 1 < (3^{d/2}+1)^2 < p$.

□

Now Theorem A.2 easily follows from Theorem A.3 using the following inequality:

$$\text{If } d \geq 26 \text{ then } (3^{d/2}+1)^2 > 65(2d)^6. \tag{A.1}$$

Indeed, suppose that K is a number field of degree $d \geq 26$ over $\mathbb{Q}$, E/K an elliptic curve and $P \in E(K)$ of prime order p. Then Proposition A.4 says that either $p < (3^{d/2}+1)^2$, in which case we are done, or the hypotheses of Theorem A.3 are satisfied. In the latter case one gets

$$(3^{d/2}+1)^2 > 65(2d)^6 > p$$

and Theorem A.2 follows.

A.3. **The Winding Quotient.** This section only contains a short summary about the winding quotient $J_0^{\mathbf{e}}(\mathbb{Q})$. For more details and the fact that $J_0^{\mathbf{e}}(\mathbb{Q})$ is finite, see either [Merel, 1996, §1] or [Parent, 1999, §3.8] or §4 of the main text. Note that the finiteness of $J_0^{\mathbf{e}}(\mathbb{Q})$ is proved by using the analytic rank 0 implies algebraic rank 0 case of the BSD conjecture as proven in [Kolyvagin and Logachëv, 1989] completed by [Bump et al., 1990] or [Murty and Murty, 1991].

If $a, b \in \mathbb{Q} \cup \{\infty\}$, then we define $\{a, b\} \in H_1(X_0(p)(\mathbb{C}), cusps, \mathbb{Z})$ to be the element given by a path from a to b in $\mathbb{H} \cup \mathbb{Q} \cup \{\infty\}$. The element $\{a, b\}$ is called a modular symbol. If $k \in \mathbb{Z}_{(p)}$ is a fraction whose denominator is not divisible by p, then define

$$\lambda(k) := \{0, 1/k\}\,. \tag{A.2}$$

The element $\lambda(k)$ only depends on $k \mod p$, hence one can also see λ as a map

$$\lambda : \mathbb{Z}/p\mathbb{Z} \to H_1(X_0(p)(\mathbb{C}), cusps, \mathbb{Z}).$$

The $\lambda(k)$ where k ranges over $\mathbb{Z}/p\mathbb{Z}$ are known to generate $H_1(X_0(p)(\mathbb{C}), cusps, \mathbb{Z})$ and if $k \not\equiv 0 \mod p$ then $\lambda(k) \in H_1(X_0(p)(\mathbb{C}), \mathbb{Z})$, and hence the element $\lambda(0) = \{0, \infty\}$ generates the rank 1 $\mathbb{Z}$-module $H_1(X_0(p)(\mathbb{C}), cusps, \mathbb{Z})/H_1(X_0(p)(\mathbb{C}), \mathbb{Z})$.

We have an isomorphism $H_1(X_0(p)(\mathbb{C}), \mathbb{R}) \cong H^1(X_0(p)(\mathbb{C}), \Omega^1)^\vee$, of real vector spaces, given by integration. So the map

$$\begin{aligned} \mathbf{e} : H^1(X_0(p)(\mathbb{C}), \Omega^1) &\to \mathbb{C} \\ \omega &\mapsto -\int_{\{0,\infty\}} \omega \end{aligned} \tag{A.3}$$

defines an element $\mathbf{e} \in H_1(X_0(p)(\mathbb{C}), \mathbb{R})$ under this isomorphism, which is called the winding element. Actually $(p-1)\mathbf{e} \in H_1(X_0(p)(\mathbb{C}), \mathbb{Z})$ showing that $\mathbf{e} \in H_1(X_0(p)(\mathbb{C}), \mathbb{Q})$. Let $\mathbb{T}$ be the sub algebra of $\operatorname{End} H_1(X_0(p)(\mathbb{C}), \mathbb{Z})$ generated by the Hecke operators and the Atkin-Lehner involution, then $\mathbb{T}$ also acts faithfully on $J_0(p)$, the Jacobian of $X_0(p)$ over $\mathbb{Z}[1/p]$. Let $\mathcal{J}_{\mathbf{e}} \subseteq \mathbb{T}$ be the annihilator of $\mathbf{e}$, then

$$J_0^{\mathbf{e}} := J_0(p)/\mathcal{J}_{\mathbf{e}} J_0(p)$$

is called the winding quotient.

Let $X_0(p)^{(d)}$ be the d-th symmetric power of the modular curve $X_0(p)$, then one has a natural map $X_0(p)^{(d)} \to J_0(p)$ by sending a divisor D of degree d to the linear equivalence class of $D - d\infty$. Composing with the quotient map $J_0(p) \to J_0^{\mathbf{e}}$ gives us the map

$$f_d : X_0(p)^{(d)} \to J_0^{\mathbf{e}}. \tag{A.4}$$

Now if $x \in X_0(p)(K)$ is a point where K is a number field of degree d and $\sigma_1, \ldots, \sigma_d : K \to \bar{\mathbb{Q}}$ are the different embeddings, then define

$$x^{(d)} := \sigma_1(x) + \cdots + \sigma_d(x) \in X_0(p)^{(d)}(\mathbb{Q}).$$

We will also write $x^{(d)}$ for $\sum_{i=1}^{d} x$ if $x \in X_0(p)(\mathbb{Q})$.

A.4. **Kamienny's Criterion.** The discussion that follows is based on section 4.12 of [Parent, 1999], who himself says that he is following Oesterlé's unpublished exposition. The main reason for following Parent, is because this allows certain proofs to be skipped and instead just cite Parent. This section is called Kamienny's criterion because the main ideas originate from [Kamienny, 1992a, §3] , although many of Kamienny's arguments have been sharpened to get the needed statement of this section. The following proposition is a slight variation of [Parent, 1999][Thm. 4.15], although his Theorem is much shorter. The reason the statement of Theorem 4.15 of Parent is so much shorter is because Parent did not include his running hypotheses in his Theorem.

Proposition A.5. *Let d be an integer and p be a prime such that $p > (3^{d/2} + 1)^2$. If there exists a number field $K/\mathbb{Q}$ of degree d, an elliptic curve E/K and a point $P \in E(K)$ of prime order p, then the map $f_d : X_0(p)^{(d)} \to J_0^e$ of equation A.4 above is not a formal immersion at $\infty_{\mathbb{F}_3}^{(d)}$.*

Proof. Let $K/\mathbb{Q}$ be a number field of degree d, E/K an elliptic curve and $0 \neq P \in E(K)[p]$. Consider j resp. $j' \in X_0(p)(K)$ to be the points corresponding to $(E, \langle P \rangle)$ resp. $(E/\langle P \rangle, E[p]/\langle P \rangle)$. By proposition A.4 one sees that $j_{\mathbb{F}_3}^{(d)} = 0_{\mathbb{F}_3}^{(d)}$ and hence $j'^{(d)}_{\mathbb{F}_3} = \infty_{\mathbb{F}_3}^{(d)}$. Now because $J_0^e(\mathbb{Q})$ is torsion and $f_d(j'^{(d)})_{\mathbb{F}_3} = f_d(\infty^{(d)})_{\mathbb{F}_3} = 0$ we get $f_d(j'^{(d)}) = f_d(\infty^{(d)}) = 0$. But $j'^{(d)} \neq \infty^{(d)}$, hence we can apply [Parent, 1999][Lemma 4.13] to get the proposition. □

The above proposition reduces the proof of Theorem A.2 to checking whether f_d is a formal immersion.

Theorem A.6. *[Parent, 1999][Thm 4.18] Let $l > 2$ be a prime, then the following two statements are equivalent:*

1. *f_d is a formal immersion at $\infty_{\mathbb{F}_l}^{(d)}$.*
2. *$T_1\mathbf{e}, \ldots, T_d\mathbf{e}$ are linearly independent in $\mathbb{T}\mathbf{e}/l\mathbb{T}\mathbf{e}$.*

A.5. **Intersection numbers of modular symbols.** Since we can view $X_0(p)(\mathbb{C})$ as a smooth oriented real manifold we get an intersection pairing on homology. The intersection pairing $\bullet : H_1(X_0(p)(\mathbb{C}), \mathbb{Z}) \times H_1(X_0(p)(\mathbb{C}), \mathbb{Z}) \to \mathbb{Z}$ also gives a pairing $\bullet : H_1(X_0(p)(\mathbb{C}), \mathbb{F}_l) \times H_1(X_0(p)(\mathbb{C}), \mathbb{F}_l) \to \mathbb{F}_l$. It would be convenient to be able to use these pairings to check the linear independence of $T_1\mathbf{e}, \ldots, T_d\mathbf{e}$ in $\mathbb{T}\mathbf{e}/l\mathbb{T}\mathbf{e}$. However while $\mathbb{T}\mathbf{e} \subset H_1(X_0(p)(\mathbb{C}), \mathbb{Q})$, it is not true that $\mathbb{T}\mathbf{e} \subset H_1(X_0(p)(\mathbb{C}), \mathbb{Z})$, so checking the linear independence cannot be checked directly with the intersection pairing. The solution, which will be worked out in more detail later, is to chose a Hecke operator I in such a way that $I\mathbf{e} \subseteq H_1(X_0(p)(\mathbb{C}), \mathbb{Z})$ and use this to write down a linear map

$$I : \mathbb{T}\mathbf{e} \to H_1(X_0(p)(\mathbb{C}), \mathbb{F}_l)$$

after which we can use the intersection pairing to check linear independence.

A.5.1. *Action of the Hecke operators on homology.* For $r > 0$ an integer and define $\sigma_1(r) := \sum_{d|r, d>0} d$. Using this definition one can compute $(T_r - \sigma_1(r))\,\mathbf{e}$ as follows.

Lemma A.7. *[Merel, 1996, Lemma 2] If p is a prime and $r < p$ a positive integer, then the following equality holds in $H_1(X_0(p)(\mathbb{C}), \mathbb{Q})$*

$$(T_r - \sigma_1(r))\,\mathbf{e} = -\sum_{\substack{a > b \geq 0 \\ d > c > 0 \\ ad - bc = r}} \lambda(c/d).$$

Where one should note that our element $\lambda(k)$ is denoted by $\xi(k)$ in [Merel, 1996].

Remark. Note that since $p > r = ad - bc \geq ad - (a-1)(d-1) \geq d > c > 0$, we see that none of the c and d in the sum are divisible by p. This means that the right hand side actually is an element of $H_1(X_0(p)(\mathbb{C}), \mathbb{Z})$. Since $H_1(X_0(p)(\mathbb{C}), \mathbb{Z})$ is torsion free, the equality actually holds in $H_1(X_0(p)(\mathbb{C}), \mathbb{Z})$, and in particular $(T_r - \sigma_1(r))\,\mathbf{e} \in H_1(X_0(p)(\mathbb{C}), \mathbb{Z})$. This is also something that could have been seen directly by noting that the boundary of $(T_r - \sigma_1(r))\,\{0, \infty\}$ is zero.

A.5.2. *The intersection number $\lambda(k) \bullet \lambda(k')$.* For p a prime and $1 \leq k < p$ an integer let k^* be the integer such that $1 \leq k^* < p$ and $kk^* \equiv -1 \mod p$ and let C_k denote the oriented straight line segment in $\mathbb{C}$ from $e^{2\pi i k/p}$ to $e^{2\pi i k^*/p}$. Recall that if $k \in \mathbb{Z}/p\mathbb{Z}^*$ then $\lambda(k)$ was defined as $\{0, 1/k\} \in H_1(X_0(p)(\mathbb{C}), \mathbb{Z})$. The intersection number of $\lambda(k)$ and $\lambda(k')$ can be computed as follows.

Lemma A.8. *[Merel, 1996, Lemma 4.] Let k, k' be two integers such that $1 \leq k < p$ and $1 \leq k' < p$. If $k' \neq k$ and $k' \neq k^*$ then $\lambda(k) \bullet \lambda(k')$ equals the intersection number $C_{k'} \bullet C_k$ and $\lambda(k) \bullet \lambda(k') = 0$ otherwise.*

Where in [Merel, 1996] the element k^* is denoted by k_*. The fact that $\lambda(k) \bullet \lambda(k') = 0$ if $k' = k$ or $k' = k^*$ is not mentioned by Merel. But this follows easily from the fact that $\bullet$ is an alternating bilinear form and $\lambda(k) = -\lambda(k^*)$.

The reason that the order of intersection is swapped is because Merel first proves $\lambda(k)\bullet\lambda(k') = C'_k \bullet C'_{k'}$ where C'_k denotes the oriented straight line segment in $\mathbb{C}$ from $e^{-2\pi ik/p}$ to $e^{-2\pi ik^*/p}$, and then concludes by $C'_k \bullet C'_{k'} = C_{k'} \bullet C_k$ because both complex conjugation and reversing the order of intersection changes sign. The lemma above is independent of the choice of orientation on $\mathbb{C}$ as long as one takes the orientation on $X_0(p)(\mathbb{C})$ to be the one compatible with the map $\mathbb{H} \to X_0(p)(\mathbb{C})$. From now on we will take the orientation on $\mathbb{C}$ such that $[-1,1] \bullet [-i,i] = 1$ where $[a,b]$ denotes the oriented straight line segment from a to b.

Definition A.9. Let $H : \mathbb{R} \to \mathbb{R}$ be the function given by

$$H(x) = \begin{cases} 1 \text{ if } x > 0 \\ \frac{1}{2} \text{ if } x = 0 \\ 0 \text{ if } x < 0 \end{cases}$$

With this definition the above lemma translates to

$$\lambda(k) \bullet \lambda(k') = -H(k'-k) + H(k'-k^*) + H(k'^*-k) - H(k'^*-k^*).$$

This equality can be verified by first checking that the both sides only depend on the cyclic ordering, with possible equalities, of k, k^*, k', k'^* in $\mathbb{Z}/p\mathbb{Z}$. And then verifying it holds for the possible cyclic orderings.

A.5.3. *The intersection number $I_r\mathbf{e} \bullet \lambda(k)$.* Let p be a prime and let $1 \le r < p$ be an integer. Define

$$I_r := T_r - \sigma_1(r),$$

then $I_r\mathbf{e} \in H_1(X_0(p)(\mathbb{C}), \mathbb{Z})$.

Proposition A.10. *Let p be a prime number and let r, k be integers such that $1 \le k < p$ and $1 \le r < p$, then one has*

$$I_r\mathbf{e} \bullet \lambda(k) = \sum_{s|r} \left(\left\lfloor \frac{sk}{p} \right\rfloor - \left\lfloor \frac{sk^*}{p} \right\rfloor \right) + v_r(k) - v_r(k^*)$$

where for $i \in \mathbb{Z}$ one defines $v_r(i)$ to be the following quantity

$$v_r(i) = \#\left\{a',b',c',d' \in \mathbb{N}_{\ge 1} \mid a'd' + b'c' = r, d'i \equiv c' \mod p\right\}$$

Proof. Define the map $x \mapsto k_x$ from $\mathbb{P}^1(\mathbb{Q})$ to the set $\{1,\dots,p\}$ by sending a simple fraction x where p does not divide the denominator to the unique element congruent to it modulo p, and one defines $k_x = p$ for $x = \infty$ and the fractions where p divides the denominator. Combining Lemmas A.7 and A.8 one gets

$$I_r\mathbf{e}\bullet\lambda(k) = \sum_{\substack{a > b \ge 0 \\ d > c > 0 \\ ad - bc = r}} \left(H(k - k_{c/d}) - H(k - k_{-d/c}) - H(k^* - k_{c/d}) + H(k^* - k_{-d/c})\right)$$

The equality stays true if we also include the terms with $c = 0$ in the sum, since those terms are all 0. Now let B_r be the set of all matrices $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ of determinant r with $a > b \geq 0$,$d > c \geq 0$ and let B'_r (resp. B''_r) be the set of matrices in B_r with $b \neq 0$ (resp. $c \neq 0$). Now we have a bijection between B'_r and B''_r by sending the matrix $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ to $\left[\begin{smallmatrix} b & -a+mb \\ d & -c+md \end{smallmatrix}\right]$ where m is the unique integer such that $0 \leq -a+mb < b$ (its inverse is obtained by sending $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ to $\left[\begin{smallmatrix} -b+na & a \\ -d+na & c \end{smallmatrix}\right]$ where n is the unique integer such that $0 \leq -d+na < c$). This shows

$$I_r\mathbf{e} \bullet \lambda(k) = S_1 - S_2 + S_3, \text{ where}$$

$$\begin{aligned}
S_1 &= \sum_{B_r \setminus B'_r} \big(H(k - k_{c/d}) - H(k^* - k_{c/d}\big) \\
S_2 &= \sum_{B_r \setminus B''_r} \big(H(k - k_{-d/c}) - H(k^* - k_{-d/c})\big) \\
S_3 &= \sum_{B'_r} \big(H(k - k_{c/d}) - H(k - k_{(c-md)/d}) - H(k^* - k_{c/d}) + H(k^* - k_{(c-md)/d})\big)
\end{aligned}$$

Let's start by calculating S_2. The matrices in $B_r \setminus B'_r$ are the matrices of the form $\left[\begin{smallmatrix} a & 0 \\ c & d \end{smallmatrix}\right]$ with $ad = r$ and $0 \leq c < d$. For $s \mid r$ let $S_1(s)$ be the contribution to S_1 of coming from the matrices such that $d = s$. The contribution to $S_1(s)$ of the matrix with $c = 0$ is 0. For $1 \leq c < d$ the number $k_{c/d}$ is equal to $\frac{up+c}{d}$, where u is the element $1 \leq u < d$ congruent to $-c/p \mod d$, and $k_{c/d}$ is the smallest integer $\geq \frac{up}{d}$. The map which associates u to c is a permutation of $\{1, \dots, d-1\}$. So the number of $c \in \{1, \dots, d-1\}$ such that $k_{c/d} \leq k$ is equal to the number of $u \in \{1, \dots, d-1\}$ such that $\frac{up}{d} \leq k$. An analogues argument with k replaced by k^* gives that

$$S_1 = \sum_{s \mid r} \left(\left\lfloor \frac{sk}{p} \right\rfloor - \left\lfloor \frac{sk^*}{p} \right\rfloor \right) - \frac{1}{2}S'_1 + \frac{1}{2}S''_1,$$

where S'_1 (resp. S''_1) is the number of pairs of integers (c, d) such that $d \mid r$, $1 \leq c < d$ and $k_{c/d} = k$ (resp. $k_{c/d} = k^*$).

The matrices in $B_r \setminus B''_r$ all have $c = 0$, hence $k_{-d/c} = p$ and $H(k - k_{-d/c}) = H(k^* - k_{-d/c}) = 0$ implying

$$S_2 = 0.$$

What remains is to determine S_3. Let $x = c/d$ be a rational number occurring in S_3, then one has that $p > r \geq d > 0$ hence $p \nmid d$. In particular if $k_x \neq 1$, then $k_{x-1} = k_x - 1$ and hence $H(k - k_x) - H(k - k_{x-1})$ equals $-\frac{1}{2}$ if $k = k_x$ or $k = k_{x-1}$ and equals 0 otherwise. If $k_x = 1$ then $k_{x-1} = p$ and $H(k - k_x) - H(k - k_{x-1})$ equals $1/2$ if $k = 1$ and 1 if $1 < k < p$. In particular, whether $k_x = 1$ or $k_x \neq 1$, the following always holds

$$H(k - k_x) - H(k - k_{x-1}) - H(k^* - k_x) + H(k^* - k_{x-1}) =$$

$$\frac{1}{2}\left(\#(\{k^*\}\cap\{k_x,k_{x-1}\}) - \#(\{k\}\cap\{k_x,k_{x-1}\})\right).$$

By induction on m, one sees that for all $m \geq 1$,

$$H(k-k_x) - H(k-k_{x-m}) - H(k^*-k_x) + H(k^*-k_{x-m})$$

equals the number of integers $i \in \{0,\dots,m\}$ such that $k^* = k_{x-i}$ minus the number of integers such that $k = k_{x-i}$, taking into account that one counts $i=0$ and $i=m$ only for half an integer.

Now to evaluate S_3, let us first define U (resp. U', resp. U'') as the set of pairs $(\left[\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right], i)$ with $\left[\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right] \in B_r'$ and $1 \leq i < m$ (resp. $i = 0$, resp. $i = m$) where m is the unique integer such that $0 \leq -a + mb < b$. Let $u(k)$ (resp. $u'(k)$, resp. $u''(k)$) be the number of these pairs such that $k = k_{(c-id)/d}$. This means that

$$S_3 = u(k^*) + \frac{1}{2}u'(k^*) + \frac{1}{2}u''(k^*) - u(k) - \frac{1}{2}u'(k) - \frac{1}{2}u''(k).$$

The map $(\left[\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right], i) \mapsto \left[\begin{smallmatrix} b & -a+ib \\ d & -c+id\end{smallmatrix}\right]$ is a bijection between U and the set of matrices of the form $\left[\begin{smallmatrix} a' & -b' \\ c' & d'\end{smallmatrix}\right]$ with a', b', c', d' integers ≥ 1 with $a'd' + b'c' = r$ (its inverse is given by sending $\left[\begin{smallmatrix} a' & -b' \\ c' & d'\end{smallmatrix}\right]$ to $\left(\left[\begin{smallmatrix} b'+ja' & a' \\ -d'+jc' & c'\end{smallmatrix}\right], j\right)$ where j is the unique integer such that $0 \leq -d' + jc' < c'$). Under this bijection, $k = k_{(c-id)/d}$ if and only if $k \equiv -d'/c' \mod p$ or equivalently if $k^* \equiv c'/d' \mod p$. This shows that $u(k) = v_r(k^*)$ and $u(k^*) = v_r(k)$.

The integer $u'(k)$ equals the number of quadruples of integers (a,b,c,d) such that $a > b > 0$, $d > c \geq 0$, $ad - bc = r$ and $k \equiv c/d \mod p$. The bijection between B_r' and B_r'', one can show that $u''(k)$ is equal to the number of quadruples (a,b,c,d) such that $a > b \geq 0$, $d > c > 0$, $ad - bc = r$ and $k \equiv -d/c \mod p$. From this it follows that $u''(k) = u'(k^*) + S_1''$ and $u''(k^*) = u'(k) + S_1'$ and hence

$$S_3 = v_r(k) - v_r(k^*) + \frac{1}{2}S_1' - \frac{1}{2}S_1''.$$

Putting the formulas for S_1, S_2 and S_3 together finally finishes the proof. □

If one defines $v_r'(i)$ by the following

$$v_r'(i) := \#\left\{a', b', c', d' \in \mathbb{N}_{\geq 1} \mid \gcd(c', d') = 1,\ a'd' + b'c' = r,\ d'i \equiv c' \mod p\right\},$$

then for $r < p$ one has $v_r(k) = \sum_{s|r} v_s'(k)$. If one also defines the Hecke operators I_r' for $1 \leq r < p$ to be such that

$$I_r = \sum_{s|r} I_s', \tag{A.5}$$

then an equivalent form of the above proposition is obtained by using he Möbius inversion formula to remove the sum over the divisors of r.

Proposition A.11. *Let p be a prime number and let $1 \leq k, k^* < p$ be integers such that and $kk^* \equiv -1 \mod p$. If r is an integer such that $1 \leq r < p$ then*

$$I_r'\mathbf{e} \bullet \lambda(k) = \left\lfloor \frac{rk}{p} \right\rfloor - \left\lfloor \frac{rk^*}{p} \right\rfloor + v_r'(k) - v_r'(k^*)$$

where for $i \in \mathbb{Z}$ one defines $v_r'(i)$ to be the following quantity

$$v_r'(i) = \#\left\{a', b', c', d' \in \mathbb{N}_{\geq 1} \mid \gcd(c', d') = 1,\ a'd' + b'c' = r,\ d'i \equiv c' \mod p\right\}$$

A.5.4. *The intersection number $I_r'\mathbf{e} \bullet \left\{0, \frac{a}{c}\right\}$.*

Proposition A.12. *Let p be a prime and r, c, d be integers such that $1 \leq r$, $1 \leq d < c < \frac{p}{r}$ and c and d are coprime. Define a, b to be the integers such that $ad - bc = 1$, $0 \leq a < c$ and $0 \leq b < d$. Define $1 \leq k < p$ and $1 \leq k^* < p$ to be the integers that are equal to c/d and $-d/c$ modulo p and finally let u, u^* be such that $dk = up + c$ and $ck^* = u^*p - d$. Then $0 \leq u < d$, $0 \leq u^* < c$ and*

$$I_r'\mathbf{e} \bullet \lambda(k) = \left\lfloor \frac{ru}{d} \right\rfloor - \left\lfloor \frac{rb}{d} \right\rfloor + \left\lfloor \frac{ra}{c} \right\rfloor - \left\lfloor \frac{ru^*}{c} \right\rfloor$$

Proof. Because $\frac{rk}{p} = \frac{ru}{d} + \frac{rc}{pd}$ and $0 \leq \frac{rc}{pd} < \frac{1}{d}$ one has

$$\left\lfloor \frac{rk}{p} \right\rfloor = \left\lfloor \frac{ru}{d} \right\rfloor.$$

And because $\frac{rk^*}{p} = \frac{ru^*}{c} - \frac{rd}{pc}$ and $0 < \frac{rd}{pc} < \frac{1}{c}$ one has

$$\left\lfloor \frac{rk^*}{p} \right\rfloor = \left\lfloor \frac{ru^* - 1}{c} \right\rfloor.$$

Now let a', b', c', d' be a quadruple as in the definition of $v_r'(k)$, because $d'k \equiv c' \mod p$ one has $c'd \equiv cd' \mod p$. Because $1 \leq cd' < cr < p$ and $1 \leq c'd < rd < p$, one even has $c'd = cd'$ and because $\gcd(c', d') = \gcd(c, d) = 1$, it follows that $c = c'$ and $d = d'$. Since $rad - rbc = r = a'd + b'c$ there exists an integer t such that $tc = ra - a'$ and $td = rb + b'$. The fact that $a', b' \geq 1$ translate into $\left\lfloor \frac{rb}{d} \right\rfloor < t \leq \left\lfloor \frac{ra-1}{c} \right\rfloor$ and since $\frac{rb}{d} < \frac{ra}{c}$ one has $\left\lfloor \frac{rb}{d} \right\rfloor \leq \left\lfloor \frac{ra-1}{c} \right\rfloor$.

This shows that under the assumptions on r, k and p, that $v_r'(k)$ is equal to the number of integers t satisfying $\left\lfloor \frac{rb}{d} \right\rfloor < t \leq \left\lfloor \frac{ra-1}{c} \right\rfloor$, or in formulas:

$$v_r'(k) = \left\lfloor \frac{ra - 1}{c} \right\rfloor - \left\lfloor \frac{rb}{d} \right\rfloor.$$

Now let a', b', c', d' be a quadruple as in the definition of $v_r'(k^*)$. Since $d'k^* \equiv c' \mod p$, we get $cc' + dd' \equiv 0 \mod p$. Now $a'd' + b'c' = r$ implies $c' + d' \leq r$ and hence $1 \leq cc' + dd' < c(c' + d') \leq cr < p$ which is incompatible with $cc' + dd' \equiv 0 \mod p$ so,

$$v_r'(k^*) = 0.$$

Putting the above equalities together one gets

$$I'_r\mathbf{e}\bullet\lambda(k) = \left\lfloor \frac{rk}{p} \right\rfloor - \left\lfloor \frac{rk^*}{p} \right\rfloor + v'_r(k) - v'_r(k^*) = \left\lfloor \frac{ru}{d} \right\rfloor - \left\lfloor \frac{rb}{d} \right\rfloor + \left\lfloor \frac{ra-1}{c} \right\rfloor - \left\lfloor \frac{ru^*-1}{c} \right\rfloor .$$

What remains to be shown is

$$\left\lfloor \frac{ra}{c} \right\rfloor - \left\lfloor \frac{ra-1}{c} \right\rfloor = \left\lfloor \frac{ru^*}{c} \right\rfloor - \left\lfloor \frac{ru^*-1}{c} \right\rfloor$$

But this is indeed the case. Since c is coprime with both u^* and a, one sees that the left and right hand side are 1 if c divides r and 0 otherwise. □

Taking $1 < k < p/r$ an integer and $d = 1$ and $c = k$ in the above proposition gives $a = 1$ which proves:

Corollary A.13. *Let p be prime and $k \geq 2$, $r \geq 1$ be integers such that $kr < p$, and let $1 \leq u^* < k$ be the inverse of p modulo k then*

$$I'_r\mathbf{e}\bullet\lambda(k) = \left\lfloor \frac{r}{k} \right\rfloor - \left\lfloor \frac{ru^*}{k} \right\rfloor .$$

Proposition A.14. *Let $c \geq 2, r \geq 1$ be integers such that $cr < p$ and $1 \leq a < c$ an integer coprime to c. Let $1 \leq u^* < c$ be such that $apu^* \equiv 1 \mod c$ then*

$$I'_r\mathbf{e}\bullet\left\{0, \frac{a}{c}\right\} = \left\lfloor \frac{ra}{c} \right\rfloor - \left\lfloor \frac{ru^*}{c} \right\rfloor .$$

Proof. We do this by induction on c. If $c = 2$ then $a = 1$ and it follows from the above corollary.

For larger c, let b, d such that $ad - bc = 1$ with $1 \leq d < c$. Because $a < c$ it follows that $b < d$. The case $d = 1$ implies $b = 0$ and hence $a = 1$ which is dealt with by the above corollary, so we can assume $d \geq 2$.

Let $1 \leq k < p$ be such that $k \equiv c/d \mod p$, then

$$\begin{bmatrix} a-bk & b \\ c-dk & d \end{bmatrix} \left\{0, \frac{1}{k}\right\} = \left\{\frac{b}{d}, \frac{a}{c}\right\}.$$

Since $k \equiv c/d \mod p$ the above matrix is in $\Gamma_0(p)$ and hence $\lambda(k) = \left\{\frac{b}{d}, \frac{a}{c}\right\}$. Since $ad \equiv 1 \mod c$ we see that the u^* of this proposition agrees with that of Proposition A.12. If we take u to be such that $pu = dk - c$, and using $bc \equiv -1 \mod d$ we get that $1 \leq u < d$ and $bpu \equiv 1 \mod d$. So using the induction hypothesis we have $I'_r\mathbf{e}\bullet\left\{0, \frac{b}{d}\right\} = \left\lfloor \frac{rb}{d} \right\rfloor - \left\lfloor \frac{ru}{d} \right\rfloor$. Writing $\left\{0, \frac{a}{c}\right\} = \left\{0, \frac{b}{d}\right\} + \left\{\frac{b}{d}, \frac{a}{c}\right\} = \left\{0, \frac{b}{d}\right\} + \lambda(k)$ finally gives

$$I'_r\mathbf{e}\bullet\left\{0, \frac{a}{c}\right\} = \left\lfloor \frac{rb}{d} \right\rfloor - \left\lfloor \frac{ru}{d} \right\rfloor + \left\lfloor \frac{ru}{d} \right\rfloor - \left\lfloor \frac{rb}{d} \right\rfloor + \left\lfloor \frac{ra}{c} \right\rfloor - \left\lfloor \frac{ru^*}{c} \right\rfloor = \left\lfloor \frac{ra}{c} \right\rfloor - \left\lfloor \frac{ru^*}{c} \right\rfloor$$

□

A.6. **Putting it all together.** With all these intersection formulas now at our disposal it is time to return to the question of when the morphism

$$f_d : X_0(p) \to J_0^e$$

of (A.4) is a formal immersion at $\infty_{\mathbb{F}_l}^{(d)}$ using Theorem A.6.

Let T_r' be the Hecke operators such that $T_r = \sum_{s|r} T_s'$ then one easily sees that for $r < p$ one has $\sum_{s|r} I_s' = T_r - \sigma_1(r) = \sum_{s|r}(T_s' - s)$ and hence $I_s' = T_s' - s$. Define $L_r := T_{2r}' - 2T_r'$ then $L_r = I_{2r}' - 2I_r'$. Using $T_{2r} = T_2T_r$ if r is odd and $T_{2r} = T_2T_r - 2T_{r/2}$ if r is even, one can deduce that for $1 \leq r < p$:

$$\sum_{s|r} I_2T_s' = (T_2 - 3)T_r = \sum_{s|r} L_s - \sum_{s|r,\, s \text{ even}} L_{s/2},$$

from which it follows that

$$I_2T_r' = \begin{cases} L_r & \text{if } r \text{ is odd} \\ L_r - L_{r/2} & \text{if } r \text{ is even} \end{cases}$$

Since $I_2\mathbf{e} \in H_1(X_0(p)(\mathbb{C}), \mathbb{Z})$ we have that I_2 induces a linear map $I_2 : \mathbb{T}\mathbf{e}/l\mathbb{T}\mathbf{e} \to H_1(X_0(p)(\mathbb{C}), \mathbb{F}_l)$, and we get the following addition to A.6.

Theorem A.15. *If $l > 2, p$ are distinct is primes and $d > 0$ an integer with $2d < p$ then $f_d : X_0(p)^{(d)} \to J_0^e$ is a formal immersion at $\infty_{\mathbb{F}_l}^{(d)}$ if either*

(1) *$L_1\mathbf{e}, L_2\mathbf{e} \ldots, L_d\mathbf{e}$ are linearly independent in $H_1(X_0(p)(\mathbb{C}), \mathbb{F}_l)$,*
(2) *$I_2'\mathbf{e}, I_3'\mathbf{e}, \ldots, I_{2d}'\mathbf{e}$ are linearly independent in $H_1(X_0(p)(\mathbb{C}), \mathbb{F}_l)$, or*
(3) *$I_2\mathbf{e}, I_3\mathbf{e}, \ldots, I_{2d}\mathbf{e}$ are linearly independent in $H_1(X_0(p)(\mathbb{C}), \mathbb{F}_l)$.*

In the above theorem the statements 2 and 3 are equivalent and they both imply the first. In Oesterlé's notes there is a part where he proved that the linear independence condition 2 of the above theorem always holds if $d > 2$ and $p/\log^4 p \geq (2d)^6$, giving a proof of Theorem A.2 for $d > 36$. We skip this part of the argument since a variation of this argument is already in [Parent, 1999, §5]. For the smaller d Oesterlé verified the linear independence 1 using the following proposition.

Proposition A.16. *Let $d \geq 1$ be an integer, $M \geq 3$ an odd integer and $l \geq 3$ a prime. Let $\varepsilon : (\mathbb{Z}/M\mathbb{Z})^* \to 0, 1$ be the map such that $\varepsilon(n) = 0$ if n is represented by an integer between 0 and $M/2$ and 1 otherwise. Let $u \in (\mathbb{Z}/M\mathbb{Z})^*$ and define the matrix $R_{d,u}$ to be the matrix with rows indexed by $\{1, \ldots, d\}$ and columns indexed by $(\mathbb{Z}/M\mathbb{Z})^*$ and whose (r, a) entry is $\varepsilon(ra) - \varepsilon(ru/a)$.*

If the matrix $R_{d,u}$ has rank d modulo l, then $L_1\mathbf{e}, \ldots, L_d\mathbf{e}$ are linearly independent in $H_1(X_0(p)(\mathbb{C}), \mathbb{F}_l)$ for all primes p such that $p > 2dM$, and $pu \equiv 1 \mod M$.

Proof. The congruence $pu \equiv 1 \mod M$ implies that $ap(u/a) \equiv 1 \mod M$ and hence $u^* \equiv u/a \mod M$ where u^* is as in Proposition A.14 with $c = M$. Now because $L_r = I_{2r}' - 2I_r'$ and $\varepsilon(n) = \left\lfloor \frac{2n}{M} \right\rfloor - 2\left\lfloor \frac{n}{M} \right\rfloor$, it follows from A.14 that for all primes p such that $p > 2dM$ and $pu \equiv 1 \mod M$ that $L_r\mathbf{e} \bullet \{0, a/M\} = \varepsilon(ra) - \varepsilon(ru/a)$. Hence the linear independence holds if $R_{d,u}$ has rank d modulo l. $\square$

A.6.1. *Proof of Theorem A.2 for* $3 \leq d \leq 25$. The following table lists for all integers $3 \leq d \leq 26$ an integer M_d such that reduction of the matrix $R_{d,u}$ modulo 3 of the above proposition has rank d for all $u \in \mathbb{Z}/M\mathbb{Z}^*$.

d	3	4	5	6	7	8	9	10	11	12	13	14
M_d	29	37	41	43	47	47	53	53	53	61	73	73

d	15	16	17	18	19	20	21	22	23	24	25	26
M_d	79	79	89	89	89	101	101	109	109	109	127	127

These values of M_d have been found using a computer and the code can be found at `https://sage.math.leidenuniv.nl/home/pub/51`. Since the M_d in the table satisfy $2dM_d < (3^{d/2}+1)^2$ if $d > 6$ and $2dM_d \leq 410$ for $d = 3, 4, 5$ it follows from Proposition A.16 that $L_1\mathbf{e}, \ldots, L_d\mathbf{e}$ are linearly independent in $H_1(X_0(p)(\mathbb{C}), \mathbb{F}_3)$ for all $p > \max((3^{d/2}+1)^2, 410)$. Hence from Theorem A.15 it follows that $f_d : X_0(p)^{(d)} \to J_0^\varepsilon$ is a formal immersion at $\infty_{\mathbb{F}_3}^{(d)}$ for all $p > \max((3^{d/2}+1)^2, 410)$, so that Theorem A.2 follows from Proposition A.5.

References

Daniel Bump, Solomon Friedberg, and Jeffrey Hoffstein. Nonvanishing theorems for L-functions of modular forms and their derivatives. *Invent. Math.*, 102(3): 543–618, 1990. ISSN 0020-9910. doi: 10.1007/BF01233440. URL `http://dx.doi.org/10.1007/BF01233440`.

P. Deligne and M. Rapoport. Correction to: "Les schémas de modules de courbes elliptiques" (*modular functions of one variable, ii* (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972), pp. 143–316, Lecture Notes in Math., Vol. 349, Springer, Berlin, 1973). In *Modular functions of one variable, IV (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972)*, pages p. 149. Lecture Notes in Math., Vol. 476. Springer, Berlin, 1975.

S. Kamienny. Torsion points on elliptic curves over fields of higher degree. *Internat. Math. Res. Notices*, (6):129–133, 1992a. ISSN 1073-7928. doi: 10.1155/S107379289200014X. URL `http://dx.doi.org/10.1155/S107379289200014X`.

S. Kamienny. Torsion points on elliptic curves and q-coefficients of modular forms. *Invent. Math.*, 109(2):221–229, 1992b. ISSN 0020-9910. doi: 10.1007/BF01232025. URL `http://dx.doi.org/10.1007/BF01232025`.

S. Kamienny and B. Mazur. Rational torsion of prime order in elliptic curves over number fields. *Astérisque*, (228):3, 81–100, 1995. ISSN 0303-1179. With an appendix by A. Granville, Columbia University Number Theory Seminar (New York, 1992).

V. A. Kolyvagin and D. Yu. Logachëv. Finiteness of the Shafarevich-Tate group and the group of rational points for some modular abelian varieties. *Algebra i Analiz*, 1(5):171–196, 1989. ISSN 0234-0852.

B. Mazur. Modular curves and the Eisenstein ideal. *Inst. Hautes Études Sci. Publ. Math.*, (47):33–186 (1978), 1977. ISSN 0073-8301. URL `http://www.numdam.org/item?id=PMIHES_1977__47__33_0`.

Loïc Merel. Bornes pour la torsion des courbes elliptiques sur les corps de nombres. *Invent. Math.*, 124(1-3):437–449, 1996. ISSN 0020-9910. doi: 10.1007/s002220050059. URL `http://dx.doi.org/10.1007/s002220050059`.

M. Ram Murty and V. Kumar Murty. Mean values of derivatives of modular L-series. *Ann. of Math. (2)*, 133(3):447–475, 1991. ISSN 0003-486X. doi: 10.2307/2944316. URL `http://dx.doi.org/10.2307/2944316`.

Pierre Parent. Bornes effectives pour la torsion des courbes elliptiques sur les corps de nombres. *J. Reine Angew. Math.*, 506:85–116, 1999. ISSN 0075-4102. doi: 10.1515/crll.1999.009. URL `http://dx.doi.org/10.1515/crll.1999.009`.

Pierre Parent. Torsion des courbes elliptiques sur les corps cubiques. *Ann. Inst. Fourier (Grenoble)*, 50(3):723–749, 2000. ISSN 0373-0956. URL `http://www.numdam.org/item?id=AIF_2000__50_3_723_0`.