Algebra & Number Theory

Volume 17
2023
No. 2

Torsion points on elliptic curves over number fields of small degree

Maarten Derickx, Sheldon Kamienny, William Stein and Michael Stoll

msp

ALGEBRA AND NUMBER THEORY 17:2 (2023)
https://doi.org/10.2140/ant.2023.17.267

Torsion points on elliptic curves over number fields of small degree

Maarten Derickx, Sheldon Kamienny, William Stein and Michael Stoll

Dedicated to the memory of Bas Edixhoven

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 $d = 4, 5, 6$, and 7.

1. Introduction

Let K be an algebraic number field and let E be an elliptic curve over K. Then the group $E(K)$ of K-rational points on E is a finitely generated abelian group; in particular, its torsion subgroup $E(K)_{\mathrm{tors}}$ is a finite abelian group, and one can ask which finite abelian groups can occur as the torsion subgroup of $E(K)$ for some elliptic curve over some number field K of degree d.

For $K = \mathbb{Q}$ (equivalently, $d = 1$), Mazur [1977; 1978] famously proved that the known finite list of possibilities for the torsion subgroup is complete. This was later extended by Merel [1996], who showed that for any given degree d, there are only finitely many possibilities for $E(K)_{\mathrm{tors}}$ when $[K : \mathbb{Q}] = d$.

One key step in these finiteness results is to show that there are only finitely many prime numbers p that can divide the order of $E(K)_{\mathrm{tors}}$, i.e., can occur as the order of an element of $E(K)$ for K of degree d. We therefore make the following definition (following [Kamienny and Mazur 1995]).

Definition 1.1. Let $d \geq 1$ be an integer. Then we define $S(d)$ to be the set of all prime numbers p such that there exists a number field K of degree d, an elliptic curve E over K and a point $P \in E(K)$ such that P has order p.

We write $\mathrm{Primes}(x)$ for the set of all prime numbers p such that $p \leq x$.

Mazur showed that

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

Kamienny [1992a] determined

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

Merel [1996, Propositions 2 and 3] showed that

$$S(d) \subseteq \mathrm{Primes}(2^{d+1} d!^{5d/2})$$

MSC2010: primary 11G05; secondary 14G05, 14G25, 14H52.
Keywords: elliptic curve, torsion point, torsion subgroup, number fields of small degree.

for $d \geq 4$. Parent [1999] gave the better bound (for all d)

$$S(d) \subseteq \text{Primes}(65(3^d - 1)(2d)^6).$$

However, Oesterlé had improved this already (as mentioned in Parent's paper) to

$$S(d) \subseteq \text{Primes}((3^{d/2} + 1)^2) \tag{1-1}$$

(except for not ruling out that $43 \in S(3)$) in his unpublished notes [Oesterlé 1994]. It should be noted that Parent actually shows that his bound is valid for prime *power* order p^n of a torsion point when $p \geq 5$ (and he has similar bounds for powers of 2 and 3); this is the main point of his work. Parent [2000; 2003], extending the techniques used by Mazur and Kamienny and relying on Oesterlé's work, proved that

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

The main result of this paper is the following theorem, which extends these results to $d = 4$, 5, 6, and 7.

Theorem 1.2.

$$S(4) = \text{Primes}(17), \quad S(5) = \text{Primes}(19), \quad S(6) = \text{Primes}(19) \cup \{37\}, \quad S(7) = \text{Primes}(23).$$

We also give a simplified proof of Parent's result on $S(3)$. Since we rely on Oesterlé's bound (1-1), a proof of which has not been published so far, we include a proof here that is based on Oesterlé's notes, which he kindly made available to us.

It is much easier to determine the set $S'(d)$ of primes p such that there are *infinitely many* elliptic curves E over number fields K of degree d with distinct j-invariants that have a K-point of order p. This is mostly a question about the gonality of the modular curve $X_1(p)$. The following is known.

Proposition 1.3.

$$S'(1) = \text{Primes}(7), \quad S'(2) = \text{Primes}(13), \quad S'(3) = \text{Primes}(13), \quad S'(4) = \text{Primes}(17),$$
$$S'(5) = \text{Primes}(19), \quad S'(6) = \text{Primes}(19), \quad S'(7) = \text{Primes}(23), \quad S'(8) = \text{Primes}(23).$$

For $d = 1, 2, 3, 4$, this is shown in [Mazur 1977; Kamienny 1992a; Jeon et al. 2011a; 2011b] respectively; for $5 \leq d \leq 8$, this follows from [Derickx and van Hoeij 2014, Theorem 3]. Since clearly $S'(d) \subseteq S(d)$, these results, together with the fact that a quadratic twist $E_{6,37}$ over the sextic number field $K = \mathbb{Q}\big(\sqrt{5}, \cos\big(\frac{2\pi}{7}\big)\big)$ of the elliptic curve

$$1225.b2\colon\ y^2 + xy + y = x^3 + x^2 - 8x + 6$$

has a point of order 37 over K [Elkies 1998, Equation 108], reduce the proof of Theorem 1.2 to showing the inclusions "$\subseteq$".

We give the following more precise result in the case $d = 6$.

Proposition 1.4. *Let K be a number field of degree 6 and let E/K be an elliptic curve such that there is a point $P \in E(K)$ of exact order 37. Then $j(E) = j(E_{6,37}) = -9317$.*

We prove Proposition 1.4 at the end of Section 8.

The gonality of $X_1(p)$ grows like p^2 [Abramovich 1996]; this implies that $S'(d) \subset \mathrm{Primes}(O(\sqrt{d}))$. On the other hand, denoting by $S_{\mathrm{CM}}(d)$ the set of primes that can occur as orders of points on elliptic curves over a number field of degree d that have complex multiplication, the results of [Clark et al. 2013] show that $S_{\mathrm{CM}}(s) \subset \mathrm{Primes}(O(d))$ and that $3d+1 \in S_{\mathrm{CM}}(d)$ when $3d+1$ is prime. (Let $p = 3d+1$. There is a pair of quadratic points defined over $\mathbb{Q}(\sqrt{-3})$ with j-invariant zero on $X_0(p)$. The set-theoretic preimage gives a Galois orbit of points of degree $2 \cdot \frac{1}{2}(p-1) \cdot \frac{1}{3} = d$ on $X_1(p)$, since the covering $X_1(p) \to X_0(p)$ ramifies with index 3 above the points with j-invariant zero.) So we will certainly have $S'(d) \subsetneq S(d)$ for infinitely many d. It is perhaps tempting to assume that for large enough d, the only sporadic points of degree d on $X_1(p)$ are CM points, as this seems to be the expectation for rational points on modular curves. This would imply that $S(d) \subseteq \mathrm{Primes}(3d+1)$ for large d. However, consulting the table in [van Hoeij 2012], it appears that there are many sporadic non-CM points (like the degree 6 points on $X_1(37)$ we have mentioned above). Still, the bound $p \leq 3d+1$ is consistent with this information for $d \geq 13$.

The strategy. To show the inclusions "$\subseteq$" in Theorem 1.2, we have to verify that $p \notin S(d)$ for every prime number p that is not in the set on the right-hand side. This is equivalent to the statement that all points of degree dividing d on the modular curve $X_1(p)$ over $\mathbb{Q}$ are cusps. Recall that noncuspidal points on $X_1(N)$, for $N \in \mathbb{Z}_{\geq 2}$, correspond to pairs (E, P), where E is an elliptic curve and $P \in E$ is a point of exact order N. See Section 2 for some background on modular curves.

Now if $x \in X_1(p)(K)$ is a point defined over a number field K of degree d, but not over a smaller field, then the sum of its Galois conjugates gives a $\mathbb{Q}$-rational effective divisor of degree d on $X_1(p)$. If x is defined over a smaller field K', then the degree d' of K' divides d, and we can take d/d' times the sum of the conjugates of x to obtain a $\mathbb{Q}$-rational effective divisor of degree d again. Effective divisors of degree d on a curve X correspond to points on its d-th symmetric power $X^{(d)}$ (which is the quotient of X^d by the natural action of the symmetric group on d letters). This leads to the following criterion. We write $C_1(p)$ for the set of cusps on $X_1(p)$.

Lemma 1.5. *Let $d \in \mathbb{Z}_{\geq 1}$ and let p be a prime number. If the composition*

$$\alpha : C_1(p)(\mathbb{Q})^d \to X_1(p)(\mathbb{Q})^d \to X_1(p)^{(d)}(\mathbb{Q})$$

of natural maps is surjective, then $p \notin S(d)$.

If $p > 2d+1$ and $p \notin S(d')$ for all $d' \leq d$, then the map above is surjective.

Proof. The assumption is equivalent to the statement that every $\mathbb{Q}$-rational effective divisor of degree d on $X_1(p)$ is a sum of rational cusps. However, if there were a number field K of degree d, an elliptic curve E over K and a point $P \in E(K)$ of order p, then (E, P) would give a K-rational noncuspidal point on $X_1(p)$ and hence, by the discussion above, a $\mathbb{Q}$-rational effective divisor of degree d that is not supported on (rational) cusps, contradicting the assumption.

For the converse, assume that the map is not surjective. Then there is a $\mathbb{Q}$-rational effective divisor D of degree d that is not supported on rational cusps. Since the irrational cusps on $X_1(p)$ form one Galois

orbit of size $(p-1)/2 > d$, D is not supported on cusps. This implies that there is a noncuspidal point on $X_1(p)$ of degree $d' \leq d$, and hence $p \in S(d')$. □

We will follow the strategy that has been established in earlier work by Mazur [1978], Kamienny [1992b; 1992a], Merel [1996], Oesterlé [1994] and Parent [1999; 2000; 2003]. We give an overview of the main steps below; for a nice and more detailed account of Merel's proof of the boundedness of $S(d)$ for all d, see [Rebolledo 2009].

In our exposition, we refer to the existing literature for proofs of many results we are using. Fairly detailed proofs of these statements can be found in an earlier version of this paper [Derickx et al. 2017] or in the doctoral thesis [Derickx 2016].

The task is to show that $p \notin S(d)$ for $3 \leq d \leq 7$ and all primes p not contained in the set on the right-hand side of the equality in Theorem 1.2. We use the criterion of Lemma 1.5, in the equivalent form given below. Before we formulate it, we make some definitions.

Definition 1.6. Let ℓ be a prime. We write $\mathbb{Z}_{(\ell)}$ for the localization of $\mathbb{Z}$ at the prime ideal $(\ell) = \ell\mathbb{Z}$.

Let X be a scheme over $\mathbb{Z}_{(\ell)}$. We denote the natural map $X(\mathbb{Z}_{(\ell)}) \to X(\mathbb{F}_\ell)$ by red_ℓ. Let $\bar{x} \in X(\mathbb{F}_\ell)$. Then $\mathrm{red}_\ell^{-1}(\bar{x})$ is the *residue class of* $\bar{x}$. When X is a model over $\mathbb{Z}_{(\ell)}$ of a projective variety over $\mathbb{Q}$, then $X(\mathbb{Z}_{(\ell)}) = X(\mathbb{Q})$, so that we can think of the residue class of $\bar{x}$ as the set of rational points on X reducing mod ℓ to $\bar{x}$.

Recall that $X_1(p)$ has a smooth model over $\mathbb{Z}[1/p]$; this implies the corresponding statement for the d-th symmetric power $X_1(p)^{(d)}$.

Lemma 1.7. *Let $\ell \neq p$ be a prime. Assume that*:

(a) *The residue class of each point $\bar{x} \in X_1(p)^{(d)}(\mathbb{F}_\ell)$ that is a sum of images under* red_ℓ *of rational cusps contains at most one rational point.*

(b) *The residue class of each point $\bar{x} \in X_1(p)^{(d)}(\mathbb{F}_\ell)$ that is not a sum of images under* red_ℓ *of rational cusps contains no rational point.*

Then $p \notin S(d)$.

Proof. Let $x \in X_1(p)^{(d)}(\mathbb{Q})$ and write $\bar{x} = \mathrm{red}_\ell(x) \in X_1(p)^{(d)}(\mathbb{F}_\ell)$. By assumption (b), $\bar{x} = \bar{x}_1 + \cdots + \bar{x}_d$ is a sum of images of rational cusps. Let $x_1, \ldots, x_d \in X_1(p)(\mathbb{Q})$ be rational cusps such that $\mathrm{red}_\ell(x_j) = \bar{x}_j$ for $1 \leq j \leq d$. Then $x' = x_1 + \cdots + x_d \in X_1(p)^{(d)}(\mathbb{Q})$ is such that $\mathrm{red}_\ell(x') = \bar{x}$. By assumption (a), x is the only rational point in the residue class of $\bar{x}$, so it follows that $x' = x \in \mathrm{im}(\alpha)$ with α as in Lemma 1.5. So α is surjective, and Lemma 1.5 shows that $p \notin S(d)$. □

Fix a rational cusp $c \in X_1(p)(\mathbb{Q})$. We can then define a morphism $\iota : X_1(p)^{(d)} \to J_1(p)$ by sending $x_1 + \cdots + x_d$ to the class of the divisor $x_1 + \cdots + x_d - d \cdot c$; here $J_1(p)$ denotes the Jacobian variety of $X_1(p)$; see Section 2 below. This map is actually defined over $\mathbb{Z}[1/p]$.

The standard way of verifying assumption (a) is to show that there is a morphism $t : J_1(p) \to A$

of abelian varieties such that

(i) $t \circ \iota$ is injective on each residue class of a point $\bar{x}$ as in assumption (a), and

(ii) $\mathrm{red}_\ell : t(J_1(p)(\mathbb{Q})) \to A(\mathbb{F}_\ell)$ is injective.

By standard properties of red_ℓ on the rational torsion subgroup, the second condition is satisfied when $t(J_1(p)(\mathbb{Q}))$ is finite and either ℓ is odd or $\ell = 2$ and $t(J_1(p)(\mathbb{Q}))$ has odd order. We can achieve this by choosing A as a factor of $J_1(p)$ that has Mordell–Weil rank zero and t to be the projection to A (plus some technicalities when $\ell = 2$). By work of Kolyvagin and Logachëv [1989] and Kato [2004], it is known that the "winding quotient" $J_1^e(p)$ of $J_1(p)$ has Mordell–Weil rank zero. Assuming the Birch and Swinnerton-Dyer conjecture for abelian varieties, $J_1^e(p)$ is in fact the largest such quotient. See Section 2 for the definition of the winding quotient.

The first condition follows if it can be shown that $t \circ \iota$ is a "formal immersion" at the relevant points $\bar{x}$; see Section 4.

We can work with $J_0(p)$ in place of $J_1(p)$. Then there is only one point $\bar{x}$ to consider, which is d times the image of the rational cusp ∞ on $X_0(p)$. This is what Mazur and Kamienny used to determine $S(1)$ and $S(2)$ and is also used in Merel's proof of an explicit bound on $S(d)$ for all d and Oesterlé's improvement of the bound. In all this work, odd primes ℓ are used. To deal with $S(3)$, Parent had to work with $J_1(p)$ (which was made possible by Kato's work showing that the winding quotient has rank zero) and also had to use $\ell = 2$ to exclude some of the primes.

One minor innovation we introduce here is that we work with some intermediate curve X_H between $X_1(p)$ and $X_0(p)$; see again Section 2. This can reduce the necessary work in cases when using $J_0(p)$ is not successful, but the dimension of $J_1(p)$ is too large to make computations feasible.

Assuming Oesterlé's bound (1-1), verification of assumption (a) amounts to exhibiting a suitable t for each prime $p \le (3^{d/2}+1)^2$ such that $p \notin S(d)$ and checking that it satisfies the conditions. This can be done by an explicit computation using modular symbols, which is based on a criterion established by Kamienny for $J_0(p)$ and extended to $J_1(p)$ by Parent. In view of assumption (b) (see below), we work with $\ell = 2$, which necessitates using "Parent's trick" to deal with the technicalities that arise when ℓ is not odd.

For certain small primes, this is not sufficient. For $d \le 7$, these primes p have the property that $J_1(p)(\mathbb{Q})$ is finite, which allows us to work with the full Jacobian and perform some more direct computations. This is another new ingredient compared to earlier work. In the course of our work, we establish an open case of a conjecture of Conrad, Edixhoven and Stein: we show that the group $J_1(29)(\mathbb{Q})$ (which is finite) is generated by differences of rational cusps on $X_1(29)$; see Theorem 3.2.

Combining both approaches, we obtain the following result.

Proposition 1.8. *Let $p \le 2281 = \lfloor (3^{7/2}+1)^2 \rfloor$ be a prime. If*

$$\begin{array}{lllllllllll} d=3 & \text{and} & p \ge 17 & \text{or} & d=4 & \text{and} & p \ge 19 & \text{or} & d=5 & \text{and} & p \ge 23 \\ & & & \text{or} & d=6 & \text{and} & p \ge 23 & \text{or} & d=7 & \text{and} & p \ge 29, \end{array}$$

then assumption (a) *of Lemma 1.7 is satisfied.*

Proposition 1.8 is proved in Section 5.

We now consider assumption (b) of Lemma 1.7. The simplest way for the assumption to be satisfied is when there are no points $\bar{x}$ that are not sums of images of rational cusps. Equivalently,

(i) there is no elliptic curve E over $\mathbb{F}_{\ell^{d'}}$ with $d' \le d$ such that $p \mid \#E(\mathbb{F}_{\ell^{d'}})$, and

(ii) $p \nmid \ell^{d'} \pm 1$ for all $d' \le d$.

The first condition excludes the existence of noncuspidal points, whereas the second excludes the possibility that $X_1(p)(\mathbb{F}_{\ell^{d'}})$ contains cusps that are not images of rational cusps. Recall that the irrational cusps are defined over the maximal real subfield of $\mathbb{Q}(\mu_p)$, which has a place of degree dividing d above ℓ if and only if $\ell^d \equiv \pm 1 \bmod p$.

We note the following simple consequence.

Lemma 1.9. *If $p > (\ell^{d/2}+1)^2$, then assumption* (b) *of Lemma 1.7 is satisfied.*

Proof. If there is an elliptic curve E over $\mathbb{F}_{\ell^{d'}}$ with $d' \le d$ such that $p \mid \#E(\mathbb{F}_{\ell^{d'}})$, then by the Hasse bound,

$$p \le \#E(\mathbb{F}_{\ell^{d'}}) \le (\ell^{d'/2}+1)^2 \le (\ell^{d/2}+1)^2,$$

which is not the case, so condition (i) above is satisfied. Since $p > (\ell^{d/2}+1)^2 > \ell^d + 1$, condition (ii) is also satisfied. □

This explains the form of Oesterlé's bound (1-1), which is related to the fact that he is working with $\ell = 3$.

We also see that it is advantageous to use the smallest possible ℓ, because then the condition of Lemma 1.9 covers more primes p. But even using $\ell = 2$, we need to verify assumption (b) for some primes $p < (2^{d/2}+1)^2$. In some cases, we can still show for such primes that there are no points $\bar{x}$ that are not sums of images of rational cusps, but this is not enough: when

$$(d, p) \in \{(5,31), (5,41), (6,29), (6,31), (6,41), (6,73), (7,29), (7,31), (7,37), (7,41), (7,43), (7,59), (7,61), (7,67), (7,71), (7,73), (7,113), (7,127)\},$$

there actually are such points, and we have to work quite a bit harder to show that they are not images of rational points on $X_1(p)^{(d)}$. This is another novel aspect of our work. We use a number of different approaches (for $p = 37$, see further below).

(1) For $p \in \{29, 31, 41\}$, we can again use direct computations based on the fact that $J_1(p)(\mathbb{Q})$ is finite and known; see Lemma 3.7.

(2) For $p \in \{71, 113, 127\}$ and $d = 7$, we use a new criterion based on gonality estimates and working with Hecke operators as correspondences, which shows directly that $p \notin S(d)$; see Corollary 7.2.

(3) For $(d, p) \in \{(6,73), (7,43)\}$, we use an intermediate curve X_H such that $X_H^{(d)}$ possesses a rational point x_H in the image of the relevant residue class and use a formal immersion argument to show that it is the only rational point in this residue class. This implies that every rational point on $X_1(p)^{(d)}$

in the residue class of $\bar{x}$ must map to x_H, but x_H does not lift to a rational point on $X_1(p)^{(d)}$; see Lemmas 8.4 and 8.5.

(4) For $p \in \{59, 61, 67, 73\}$ and $d = 7$, we use another new criterion that shows that a noncuspidal point $\bar{x} \in X_1(p)^{(d)}(\mathbb{F}_2)$ is not the reduction mod 2 of a rational point by showing that its image in $J_1(p)(\mathbb{F}_2)$ is not in the reduction of the Mordell–Weil group; see Lemma 8.7.

We then obtain the following result.

Proposition 1.10. *For the following pairs of an integer* $3 \le d \le 7$ *and a prime* p, *assumption* (b) *of Lemma 1.7 is satisfied*:

$$\begin{array}{llllll} d = 3 & \text{and} & p = 11 & \text{or} & p \ge 17; \\ d = 4 & \text{and} & p \ge 19; & & \\ d = 5 & \text{and} & p \ge 23; & & \\ d = 6 & \text{and} & p \ge 23 & \text{and} & p \ne 37; \\ d = 7 & \text{and} & p \ge 29 & \text{and} & p \ne 37. \end{array}$$

Proposition 1.10 is proved in Section 8.

We still have to show Proposition 1.4 and that $37 \notin S(7)$. We combine the approaches in (3) and (4) to do this. We first show using (4) that no noncuspidal point in $X_1(37)^{(7)}(\mathbb{F}_2)$ is the reduction mod 2 of a rational point and there is essentially only one such point in $X_1(37)^{(6)}(\mathbb{F}_2)$. We then use the formal immersion argument as in (3) to show that the remaining points in $X_1(37)^{(d)}(\mathbb{F}_2)$ for $d = 6, 7$ lift uniquely to rational points; see Lemmas 8.8 and 8.9.

Theorem 1.2 then follows from this and Propositions 1.8 and 1.10, using Lemma 1.7 and Oesterlé's bound (1-1).

A large part of the work done in this paper relies heavily on computations. We provide Magma [Bosma et al. 1997] code (with explanatory comments) for all these computations in an online supplement. The timings we give in some places in this paper were obtained on the last author's current laptop (as of 2020). All computations together took about one day on this machine. We also provide [SageMath] code at the first author's GitHub site [Derickx 2020] that independently verifies the claims made in Section 5. Some of these computations rely on modular symbols. See for example [Stein 2007] for the necessary background.

The structure of the paper. We begin by recalling some background on modular curves in Section 2. In Section 3, we quote the list of primes p such that $J_1(p)(\mathbb{Q})$ is finite from [Conrad et al. 2003] and prove that for such primes, $J_1(p)(\mathbb{Q})$ is generated by differences of rational cusps (the new case being $p = 29$) and that the reduction map $J_1(p)(\mathbb{Q}) \to J_1(p)(\mathbb{F}_2)$ is injective. We use this to prove assumption (b) for $p = 29$, 31, and 41. In Section 4, we introduce formal immersions and state the computational criterion we use to verify assumption (a). Section 5 reports on these computations, and Section 6 contains the proof of Oesterlé's bound (1-1). In Section 7 we state and prove the criterion used to show that $71, 113, 127 \notin S(7)$. Finally, we complete the verification of assumption (b) in Section 8, which also contains the proof of Proposition 1.4.

What is new in this paper? The main new *result* is Theorem 1.2, which extends the list of known sets $S(d)$ from $d \leq 3$ to $d \leq 7$. Completing the determination of $S(6)$, Proposition 1.4 gives a classification of the sporadic points in $X_1(37)^{(6)}(\mathbb{Q})$. Another new result is Theorem 3.2, which confirms a conjecture made in [Conrad et al. 2003] in a case that was left open in that paper.

We also develop some new *techniques* for proving that $p \notin S(d)$ for suitable $d \geq 1$ and primes p. One point is the use of intermediate curves in various computations instead of just either $X_0(p)$ or $X_1(p)$. Another is the use of explicit computations in the Picard group of $X_1(p)_{\mathbb{F}_2}$ when p is such that $J_1(p)(\mathbb{Q})$ is finite. In addition, we derive two new criteria, one that uses the gonality of $X_1(p)$ and can show directly that $p \notin S(d)$ using global arguments (Proposition 7.1), and a related one that works over $\mathbb{F}_2$ using Hecke correspondences (Lemma 8.6). Finally, we extend the formal immersion approach that is traditionally used to show what we call assumption (a) to also apply to assumption (b). All this is necessary to be able to determine $S(7)$.

Why stop at $d = 7$? Obviously, determining $S(d)$ gets harder and harder as d grows. When $d = 1$, the formal immersion condition for $X_0(p)$ is essentially trivially satisfied, and assumption (b) for $\ell = 3$ is automatically satisfied for $p > \lfloor(\sqrt{3}+1)^2\rfloor = 7$. Once the theoretical framework is in place (which, of course, was the key contribution of Mazur [1977; 1978]), no computation is necessary to obtain the desired result.

For $d = 2$, Kamienny had to come up with a criterion that allows one to verify the formal immersion condition (still for $X_0(p)$). In this case, it can still be shown to hold by a theoretical argument. The trivial bound for assumption (b) when $\ell = 3$ is $p > 16$, which is again sufficient.

For $d = 3$ and larger, one needs to work to verify the formal immersion condition. Merel and Oesterlé managed to find a theoretical argument that does this (for $X_0(p)$ and $\ell = 3$) for p larger than some explicit polynomial in d. Oesterlé then came up with another ingenious way to reduce the remaining cases to a finite and manageable amount of computation, thus proving the bound (1-1). To determine $S(3)$, Parent had to rely on this and to come up with a way of using $X_1(p)$ and $\ell = 2$ to cover the primes between $\lceil(2^{3/2}+1)^2\rceil = 15$ and $\lfloor(3^{3/2}+1)^2\rfloor = 38$ (and $p = 43$, which had escaped Oesterlé's approach).

For $d \geq 4$, there are two main difficulties, which each get worse as d increases.

(1) The gap between the best general bound (1-1) and the smallest prime not in $S(d)$ increases exponentially with d. While we can, for each d and each p in this range, verify the formal immersion condition for $X_0(p)$ or some intermediate curve X_H computationally, the computational effort increases considerably with p. For $d = 7$, this part of the computation took about two hours. For $d = 8$, the upper end of this range is about three times as large as for $d = 7$, which lets us expect that doing this in reasonable time would require a massively parallel computation. For $d \geq 9$, this appears to be infeasible in the absence of a major theoretical advance that leads to a significantly reduced general bound.

(2) The gap between the primes in $S(d)$ and the "easy" range for assumption (b) also increases. Most likely, this increase is also exponential, because we expect that $\max S(d)$ should grow only

polynomially (possibly even linearly). This means that there will be more and more primes p for which we have to show assumption (b) when there are indeed points in $X_1(p)^{(d)}(\mathbb{F}_2)$ that are not sums of images of rational cusps. While we could deal with the "rank-zero primes" $p = 29, 31, 41$ by explicit computations and with the one further such prime $p = 73$ for $d = 6$ by a variant of the formal immersion criterion, this is the point where it gets hard when $d = 7$. To rule out the primes 37, 43, 59, 61, 67, 71, 73, 113, and 127, we needed to develop some new criteria, and some of the computations that are then still necessary run for several hours.

However, it appears that our new criteria can be used to go a bit further. This will be explored in a follow-up paper.

2. Preliminaries on modular curves

A good reference for most of the following is [Diamond and Im 1995].

As usual, we define, for $N \in \mathbb{Z}_{\geq 1}$,

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : N \mid c \right\} \quad \text{and} \quad \Gamma_1(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : (c, d) \equiv (0, 1) \bmod N \right\}.$$

$\mathrm{SL}_2(\mathbb{Z})$ and therefore also Γ_0 and Γ_1 act on the complex upper half-plane $\mathfrak{H}$ and on $\mathfrak{H}^* = \mathfrak{H} \cup \mathbb{P}^1(\mathbb{Q})$ by Möbius transformations. Then the quotient $Y_j(N)(\mathbb{C}) = \Gamma_j(N)\backslash\mathfrak{H}$ (for $j = 0, 1$) is a Riemann surface that can be compactified to $X_j(N)(\mathbb{C}) = \Gamma_j(N)\backslash\mathfrak{H}^*$ by adding the finitely many cusps $C_j(N) = \Gamma_j(N)\backslash\mathbb{P}^1(\mathbb{Q})$. The points in $Y_1(N)(\mathbb{C})$ classify pairs (E, P) consisting of an elliptic curve E over $\mathbb{C}$ and a point $P \in E(\mathbb{C})$ of exact order N; in terms of a representative point $\tau \in \mathfrak{H}$, this is given by $E = \mathbb{C}/\Lambda_\tau$ with $\Lambda_\tau = \mathbb{Z} + \mathbb{Z}\tau$ and $P = 1/N + \Lambda$. Similarly, $Y_0(N)(\mathbb{C})$ classifies pairs (E, C) where again E is an elliptic curve over $\mathbb{C}$ and $C \subseteq E(\mathbb{C})$ is a cyclic subgroup of order N. The compact Riemann surfaces $X_j(N)(\mathbb{C})$ can be identified with the set of complex points on projective algebraic curves $X_j(N)$ defined over $\mathbb{Q}$ (or even over $\mathbb{Z}[1/N]$). The rational structure is defined in terms of the q-expansions of functions on $X_j(N)(\mathbb{C})$: such a function f lifts to a modular function with respect to $\Gamma_j(N)$ on $\mathfrak{H}$ and therefore has a Laurent series expansion in terms of $q = e^{2\pi i\tau}$. The function field $\mathbb{Q}(X_j(N))$ is then defined to consist of those f whose q-expansion has coefficients in $\mathbb{Q}$. Since the rational structure is defined in terms of q, the natural moduli interpretation of a point on $X_1(N)$ over $\mathbb{Q}$ (or any field extension K) is as representing a pair (E, φ), where E is an elliptic curve over $\mathbb{Q}$ (or K) and $\varphi : \mu_N \to E$ is an embedding of the group μ_N of N-th roots of unity into E as group schemes. This is because the image of $1/N$ under $\tau \mapsto e^{2\pi i\tau}$ is not rational, but a generator of μ_N. Since (over any field K of characteristic not dividing N) there is a natural bijection between pairs (E, P) and pairs (E', φ) as above, the points on $X_1(N)$ can still be understood as classifying elliptic curves over K together with a point of order N, but we have to keep in mind that this is not the same as the moduli interpretation over $\mathbb{C}$ given above. (The bijection is obtained as follows. Given a pair (E, P) and $\zeta \in \mu_N$, the set of $Q \in E[N]$ such that $e_N(Q, P) = \zeta$ forms a coset C_ζ of the subgroup $\mathbb{Z}P$ generated by P. We then set $E' = E/\mathbb{Z}P$ and $\varphi : \zeta \mapsto C_\zeta/\mathbb{Z}P$.)

The space of cusp forms for $\Gamma_j(N)$ is canonically isomorphic to the space of regular differentials

on $X_j(N)(\mathbb{C})$. Under this isomorphism, regular differentials on $X_j(N)$ over $\mathbb{Q}$ correspond to cusp forms whose q-expansion has rational coefficients.

There is a natural map $X_1(N) \to X_0(N)$ (induced over $\mathbb{C}$ by the identity on $\mathfrak{H}^*$). This makes $X_1(N)$ into a (possibly ramified) Galois covering of $X_0(N)$, whose Galois group consists of the diamond operators $\langle a\rangle$ for $a \in (\mathbb{Z}/N\mathbb{Z})^\times$, where $\langle -1\rangle$ is the identity, so the Galois group is naturally isomorphic to $(\mathbb{Z}/N\mathbb{Z})^\times/\{\pm 1\}$. In terms of the interpretation of points on $Y_1(N)$ as pairs $(E, \varphi : \mu_N \to E)$, the action of $\langle a\rangle$ corresponds to precomposing φ with the a-th power map. If $H \subseteq (\mathbb{Z}/N\mathbb{Z})^\times/\{\pm 1\}$ is a subgroup, then we have an intermediate curve $X_H = H\backslash X_1(N)$ between $X_1(N)$ and $X_0(N)$.

We write $\infty \in X_j(N)$ for the cusp that over $\mathbb{C}$ is the image of $\infty \in \mathbb{P}^1(\mathbb{Q})$. Note that $\infty \in X_j(N)(\mathbb{Q})$, since it corresponds to $q = 0$. When $N = p$ is prime, $X_0(p)$ has the two cusps ∞ and the cusp represented by $0 \in \mathbb{P}^1(\mathbb{Q})$, which are both rational, whereas $X_1(p)$ has $p-1$ cusps, which split into two orbits under the diamond operators, each consisting of $(p-1)/2$ cusps. One of the orbits contains ∞ and consists of rational cusps, the other orbit consists of cusps defined over the maximal totally real subfield of the cyclotomic field $\mathbb{Q}(\mu_p)$; these cusps are all conjugate under the Galois action, and the Galois action is given by diamond operators (since it commutes with them). An analogous statement is true for the cusps of X_H. See [Stevens 1982, Theorem 1.3.1] for a description of the Galois action on the cusps.

We denote the Jacobian varieties of $X_0(N)$, $X_1(N)$ and X_H by $J_0(N)$, $J_1(N)$ and J_H, respectively. They are defined over $\mathbb{Q}$ and extend to abelian schemes over $\mathbb{Z}[1/N]$.

We denote the Hecke algebra, in its various incarnations, by $\mathbb{T}$. It is generated by the Hecke operators T_n for all $n \geq 1$ or, alternatively, by all T_p for p prime together with diamond operators $\langle a\rangle$ for a generating $(\mathbb{Z}/N\mathbb{Z})^\times/\{\pm 1\}$. The Hecke algebra acts on the integral homology $H_1(X_H(\mathbb{C}), \mathbb{Z})$, on the relative homology $H_1(X_H(\mathbb{C}), \text{cusps}, \mathbb{Z})$, on the associated spaces of modular forms or cusp forms, and as endomorphisms of J_H. The Hecke operators T_n and the diamond operators $\langle a\rangle$ can also be viewed as correspondences on X_H. It will always be clear from the context or explicitly stated which interpretation is considered.

The integral relative homology with respect to the cusps is generated as a $\mathbb{Z}$-module by modular symbols $\{\gamma_1, \gamma_2\}$ with $\gamma_1, \gamma_2 \in \mathbb{P}^1(\mathbb{Q})$. There is an integration pairing

$$H_1(X_H(\mathbb{C}), \text{cusps}, \mathbb{Z}) \times H^0(X_{H,\mathbb{C}}, \Omega^1) \to \mathbb{C}, \quad (\xi, \omega) \mapsto \int_\xi \omega$$

(if $\xi = \{\gamma_1, \gamma_2\}$, then the integral is along any path in $\mathfrak{H}^*$ joining γ_1 to γ_2); it induces a perfect pairing of real vector spaces between $H_1(X_H(\mathbb{C}), \mathbb{R})$ and $H^0(X_{H,\mathbb{C}}, \Omega^1)$, and the image of the composition

$$\pi : H_1(X_H(\mathbb{C}), \text{cusps}, \mathbb{Z}) \to H^0(X_{H,\mathbb{C}}, \Omega^1)^* \to H_1(X_H(\mathbb{C}), \mathbb{R})$$

is in the rational homology $H_1(X_H(\mathbb{C}), \mathbb{Q})$ by the Manin–Drinfeld theorem [Manin 1972; Drinfeld 1973].

Definition 2.1. We set

$$\boldsymbol{e} = \pi(-\{0, \infty\}) \in H_1(X_H(\mathbb{C}), \mathbb{Q});$$

this is called the *winding element*. Its annihilator $\mathrm{Ann}(\boldsymbol{e})$ in $\mathbb{T}$ is the *winding ideal*. It acts via endomorphisms on J_H; the quotient $J_H^{\boldsymbol{e}} := J_H/\mathrm{Ann}(\boldsymbol{e})J_H$ is the *winding quotient*.

The definition of the winding element goes back to Mazur [1977, Lemma II.18.6 and the definition preceding it] in the case of $J_0(N)$. We note that there is some ambiguity regarding the sign of the winding element in the literature. We follow [Merel 1996, Section 1] here (but, for example, [Parent 1999] uses the opposite sign.) The winding quotient has the following essential property.

Theorem 2.2. *For each subgroup $H \subseteq (\mathbb{Z}/N\mathbb{Z})^\times/\{\pm 1\}$, the Mordell–Weil group $J_H^{\boldsymbol{e}}(\mathbb{Q})$ is finite.*

Merel [1996, § 1] was the first one to introduce the winding quotient for $J_0(p)$ with p prime, where he also proves that its Mordell–Weil group is finite using a result from [Kolyvagin and Logachëv 1989], which states that an abelian variety A over $\mathbb{Q}$ that is a quotient of $J_0(N)$ has Mordell–Weil rank 0 when $L(A, 1) \neq 0$. Parent [1999, § 3.8] generalized Merel's statement to composite numbers N. The result of Kolyvagin and Logachëv was generalized by Kato [2004, Corollary 14.3] to quotients of $J_1(N)$. In both [Parent 2000] and [Parent 2003], it is mentioned that the theorem follows from Kato's generalization. This can be seen by adapting the arguments of [Parent 1999, § 3.8] accordingly. The key point in the proof is that $J_H^{\boldsymbol{e}}$ is isogenous to a product of simple abelian varieties A over $\mathbb{Q}$ such that $L(A, 1) \neq 0$. Kato's result then shows that $A(\mathbb{Q})$ is finite.

The following is a variant of [Parent 2000, Proposition 1.8]. We remark that, according to [Diamond and Im 1995, p. 87], the Eichler–Shimura relation on $X_1(N)$ with the modular interpretation used here (and in [Parent 2000]) is different from that valid with the more usual interpretation as parametrizing pairs (E, P) of elliptic curves with a point of order N. We therefore believe that our version is correct, and that (the first part of) Parent's statement needs to be modified accordingly.

Proposition 2.3. *Let $q \nmid N$ be a prime and $P \in J_H(\mathbb{Q})_{\mathrm{tors}}$ such that q is odd or P has odd order. Then $(T_q - \langle q\rangle - q)(P) = 0$.*

Proof. Let n be the order of P. Then $(T_q - \langle q\rangle - q)(P) \in J_H(\mathbb{Q})$ is a point of order dividing n. We write $\bar{P}$ for the reduction mod q of P, Frob_q for the Frobenius on $J_{H,\mathbb{F}_q}$ and Ver_q for its dual (Verschiebung). Then we have the Eichler–Shimura relation

$$T_{q,\mathbb{F}_q} = \langle q\rangle \,\mathrm{Frob}_q + \mathrm{Ver}_q \quad \text{and} \quad \mathrm{Ver}_q \circ \mathrm{Frob}_q = q$$

in $\mathrm{End}_{\mathbb{F}_q}(J_{H,\mathbb{F}_q})$; see [Diamond and Im 1995, p. 87]. So, using that $\mathrm{Frob}_q(\bar{P}) = \bar{P}$,

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

which implies that $(T_{q,\mathbb{F}_q} - \langle q\rangle - q)(\bar{P}) = 0$. Since the reduction map is injective on $J_H(\mathbb{Q})_{\mathrm{tors}}$ when q is odd, and it is injective on odd-order torsion when $q = 2$, the claim follows. □

Remark. For the proof of Theorem 3.2, it actually does not matter whether one uses $T_q - \langle q\rangle - q$ as in Proposition 2.3 or $T_q - \langle q\rangle q - 1$ as in [Parent 2000]. Up to composition with $\langle q\rangle$ or its inverse, the two operators are conjugate to each other under the Atkin–Lehner involution w_p; see [Diamond and Im 1995,

p. 56 and Remark 10.2.2]. If we use the "wrong" operators in the proof of Theorem 3.2, then instead of $J_1(29)(\mathbb{Q})_{\mathrm{tors}} \subseteq C$, we find that $w_{29}(J_1(29)(\mathbb{Q})_{\mathrm{tors}}) \subseteq C$, which also implies $J_1(29)(\mathbb{Q})_{\mathrm{tors}} \subseteq w_{29}(C) = C$ (the cusps are permuted by w_p).

Our second application in Corollary 5.2 uses the operators with q odd to kill torsion. By the remark after Corollary 4.3, it is enough to kill 2-torsion. For q an odd prime, the two operators $T_q - \langle q\rangle - q$ and $T_q - \langle q\rangle q - 1$ differ by a multiple of 2 in the Hecke algebra, so they have the same effect on 2-torsion points. This implies that the conclusion of Corollary 4.3 also holds if we use the "wrong" operator and show that $t \circ \iota$ is a formal immersion. In particular, the conclusions of [Parent 2000] are valid.

We will also need the following statement.

Proposition 2.4 (Derickx). *Let $q \nmid N$ be a prime. We consider $t = T_q - \langle q\rangle - q \in \mathbb{T}$ as a correspondence on $X_1(N)$, inducing an endomorphism of the divisor group of $X_1(N)$ over $\mathbb{C}$. Then the kernel of t is contained in the subgroup of divisors supported in cusps.*

Proof. Let D be a divisor in the kernel of t, so that

$$T_q(D) = \langle q\rangle(D) + qD. \tag{2-1}$$

A noncuspidal point $x \in X_1(N)(\mathbb{C})$ corresponds to an elliptic curve E over $\mathbb{C}$ with additional structure. The point $\langle q\rangle(x)$ corresponds to the same curve E (with modified extra structure), and $T_q(x)$ is a sum of points corresponding to all the elliptic curves that are q-isogenous to E. We define the *q-isogeny graph G* to have as vertices the isomorphism classes of all elliptic curves over $\mathbb{C}$; two vertices are connected by an edge when there is a q-isogeny between the corresponding curves. There is a natural map γ from $Y_1(N)(\mathbb{C})$ to the vertex set of G. Let x be a noncuspidal point in the support of D and let G_x be the connected component of G containing $\gamma(x)$. Let E be the elliptic curve given by x. We distinguish two cases.

First, assume that E does not have CM. Then G_x is an infinite $(q+1)$-regular tree. Consider a vertex v of G_x that has maximal possible distance from $\gamma(x)$ among all vertices of the form $\gamma(y)$ for a point y in the support of D. Let $y_1, \ldots, y_n$ be the points in the support of D such that $\gamma(y_j) = v$, and let w be a neighbor of v whose distance from $\gamma(x)$ is larger than that of v. Each $T_q(y_j)$ contains precisely one point y_j' such that $\gamma(y_j') = w$, and these points are distinct for distinct points y_j. Since w is not of the form $\gamma(z)$ for a point z in the support of D, this shows that $T_q(D)$ has points in its support that do not occur in the support of $\langle q\rangle(D) + qD$ (recall that $\gamma(\langle q\rangle(y)) = \gamma(y)$). This contradicts the relation (2-1), and we conclude that there can be no non-CM point x in the support of D.

Now consider the case that E has CM. Then G_x is no longer a tree in general, but has the structure of a "volcano"; see [Sutherland 2013]. For a CM elliptic curve over $\mathbb{C}$, this volcano has infinite depth. Concretely, this means that it consists of a number of rooted $(q+1)$-regular trees whose roots form a cycle (which may have length 1 or 2). We can now argue as in the first case by choosing v to be a vertex of maximal level (i.e., distance from the root cycle) and w to be a neighbor of v whose level is larger by one. This shows that there can be no CM points in the support of D either.

The only points that we have not excluded from the support of D are the cusps; this proves the claim. □

Remark. In the case that $N = p$ is a prime, we can describe the kernel exactly. The rational cusps are killed by t, whereas the irrational cusps are killed by $t^* = T_q - q\langle q\rangle - 1$; compare [Parent 2000, Section 2.4] (the rational cusps are those mapping to the cusp ∞ on $X_0(p)$ under the modular interpretation we use). Since $t - t^* = (q-1)(\langle q\rangle - 1)$ and the divisor group is torsion-free, t kills a divisor supported on irrational cusps if and only if it is invariant under $\langle q\rangle$.

3. Rank-zero primes

We say that a prime p is a *rank-zero prime* when $J_1(p)(\mathbb{Q})$ is finite.

The following result gives us the list of rank-zero primes. This is [Conrad et al. 2003, Proposition 6.2.1]; we include some more information from Section 6.2 of loc. cit.

Proposition 3.1. *The rank-zero primes p are the primes $p \le 31$ and 41, 47, 59, and 71.*

For all of these, except possibly $p = 29$, the group $J_1(p)(\mathbb{Q})$ 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, which confirms Conjecture 6.2.2 in [Conrad et al. 2003] for the smallest open case, $p = 29$.

Theorem 3.2. *The group $J_1(29)(\mathbb{Q})$ is generated by differences of rational cusps.*

Proof. We prove this by a computation using modular symbols, as follows. The group $J_1(29)(\mathbb{C})_{\mathrm{tors}}$ is canonically isomorphic to $M := H_1(X_1(29)(\mathbb{C}), \mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{Q}/\mathbb{Z}$. By Proposition 2.3, the image of the rational torsion subgroup is annihilated by $T_q - \langle q\rangle - q$ for all odd primes $q \neq 29$, and it is also annihilated by $\tau - 1$, where τ is induced by complex conjugation. We let M' be the subgroup of M annihilated by $\tau - 1$ and $T_q - \langle q\rangle - q$ for $q = 3, 5, 7$. We find that

$$J_1(29)(\mathbb{Q})_{\mathrm{tors}} \subseteq M' \cong \frac{\mathbb{Z}}{2\mathbb{Z}} \times \frac{\mathbb{Z}}{2\mathbb{Z}} \times \frac{\mathbb{Z}}{2\mathbb{Z}} \times \frac{\mathbb{Z}}{2^2\mathbb{Z}} \times \frac{\mathbb{Z}}{2^2\mathbb{Z}} \times \frac{\mathbb{Z}}{2^2\cdot 3\cdot 7\cdot 43\cdot 17837\mathbb{Z}}.$$

We can also compute the cuspidal subgroup C as the image in M of the relative homology $H_1(X_1(29)(\mathbb{C}), \text{cusps}, \mathbb{Z})$ via its embedding into $H_1(X_1(29)(\mathbb{C}), \mathbb{Q})$. We obtain that

$$M' \subseteq C \cong \frac{\mathbb{Z}}{2^2\mathbb{Z}} \times \frac{\mathbb{Z}}{2^2\mathbb{Z}} \times \frac{\mathbb{Z}}{2^2\mathbb{Z}} \times \frac{\mathbb{Z}}{2^2\mathbb{Z}} \times \frac{\mathbb{Z}}{2^2\cdot 3\cdot 43\cdot 17837\mathbb{Z}} \times \frac{\mathbb{Z}}{2^2\cdot 3\cdot 7^2\cdot 43\cdot 17837\mathbb{Z}}.$$

Finally, we have an explicit homomorphism $\mathbb{Z}[\text{cusps}]^0 \to C$, where $\mathbb{Z}[\text{cusps}]^0$ denotes the degree-zero part of the free abelian group with basis the cusps of $X_1(29)$. We know that the absolute Galois group of $\mathbb{Q}$ fixes the 14 cusps mapping to the cusp ∞ of $X_0(29)$, whereas the remaining 14 cusps are permuted cyclically via the action of the diamond operators. This allows us to determine

$$J_1(29)(\mathbb{Q}) = J_1(29)(\mathbb{Q})_{\mathrm{tors}} = C^{\mathrm{Gal}_{\mathbb{Q}}} \cong \frac{\mathbb{Z}}{2^2\mathbb{Z}} \times \frac{\mathbb{Z}}{2^2\mathbb{Z}} \times \frac{\mathbb{Z}}{2^2\cdot 3\cdot 7\cdot 43\cdot 17837\mathbb{Z}},$$

and we can verify that this equals the subgroup generated by differences of rational cusps. □

Remark. In our Magma code for the computations in the proof above, we rely only on linear algebra functionality (over $\mathbb{Q}$ and $\mathbb{Z}$): we construct the relevant spaces "by hand" instead of using the built-in modular symbols functionality.

Together with Proposition 3.1, this immediately implies the following.

Corollary 3.3. *If p is a prime such that $J_1(p)(\mathbb{Q})$ is finite, then the latter group is generated by differences of rational cusps on $X_1(p)$.*

We will need the following result on the reduction mod 2.

Proposition 3.4. *If $p > 2$ is a prime such that $J_1(p)(\mathbb{Q})$ is finite, then the reduction map*

$$\mathrm{red}_2 : J_1(p)(\mathbb{Q}) = J_1(p)(\mathbb{Q})_{\mathrm{tors}} \to J_1(p)(\mathbb{F}_2)$$

is injective.

Proof. Let X be a curve over $\mathbb{Q}$ with good reduction at 2, and let J be its Jacobian variety. Then the kernel of the reduction map $J(\mathbb{Q})_{\mathrm{tors}} \to J(\mathbb{F}_2)$ is contained in the 2-torsion subgroup [Parent 2000, Lemma 1.7]. So the claim follows for all p such that $J_1(p)(\mathbb{Q})$ is finite of odd order. For the remaining primes on our list, namely $p \in \{17, 29, 31, 41\}$, we check by an explicit computation that $J_1(p)(\mathbb{Q})[2] \to J_1(p)(\mathbb{F}_2)$ is injective. This then implies the claim for these primes as well.

We now describe this computation. By Corollary 3.3, we know $J_1(p)(\mathbb{Q})$ is generated by differences of rational cusps. The order of this group is known; see [Conrad et al. 2003, § 6.2.3 and Table 1] and note that the order for $p = 29$ given there has to be divided by 2^6 to get the order of the group generated by differences of rational cusps; compare Theorem 3.2. Sutherland (https://math.mit.edu/~drew/X1_altcurves.html) provides equations for planar models of $X_1(p)$ over $\mathbb{Q}$ for the relevant values of p. We use the reduction modulo 2 of this model to check that the subgroup of its Picard group generated by differences of its degree-1 places over $\mathbb{F}_2$ (which correspond to the rational 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 $p = 17$, 29, and 31, this only takes a few minutes; for $p = 41$ the computation of the Picard group of $X_1(p)$ over $\mathbb{F}_2$ takes about eight hours (and 2.5 gigabytes of memory). □

Remark. If one does not want to wait for several hours for the computation for $p = 41$ to finish, one can alternatively use the intermediate curve X_H corresponding to $d = 4$ in the notation of [Conrad et al. 2003] (then H has index 4). The predicted order of the 2-primary part of $J_H(\mathbb{Q})$ equals that of $J_1(p)(\mathbb{Q})$. We check that the 2-primary part of the subgroup of $J_H(\mathbb{F}_2)$ generated by differences of the images of rational cusps has the correct size.

Remark. For $p = 17$, Proposition 3.4 together with the fact that the $\mathbb{Q}$-gonality of $X_1(17)$ is 4 gives a simple alternative proof of the main result of [Parent 2003] that $17 \notin S(3)$; see Corollary 3.6 below. (Note that $17 > (2^{3/2} + 1)^2$.)

Remark. The statement of Proposition 3.4 is false for $J_0(p)$. For example, $J_0(17)$ is the elliptic curve with Cremona label $17a1$. It has $J_0(17)(\mathbb{Q}) \cong \mathbb{Z}/4\mathbb{Z}$, generated by the difference of the two cusps, but the reduction modulo 2 of a generator has only order 2.

We now show that assumption (a) in Lemma 1.7 is satisfied when $\ell = 2$ and p is a rank-zero prime such that the $\mathbb{Q}$-gonality of $X_1(p)$ is strictly larger than d. The $\mathbb{Q}$-*gonality* of a curve X over $\mathbb{Q}$ is the smallest degree of a nonconstant rational function on X defined over $\mathbb{Q}$.

Recall the embedding $\iota : X_1(p)^{(d)} \to J_1(p)$ given by fixing a basepoint $c \in C_1(p)(\mathbb{Q})$.

Corollary 3.5. *Let $d \geq 1$ be an integer. If $p > 2$ is a rank-zero prime and the $\mathbb{Q}$-gonality of $X_1(p)$ is strictly larger than d, then assumption* (a) *in Lemma 1.7 is satisfied for $\ell = 2$.*

Proof. The map $\iota : X_1(p)^{(d)}(\mathbb{Q}) \to J_1(p)(\mathbb{Q})$ is injective when the $\mathbb{Q}$-gonality of $X_1(p)$ exceeds d, since otherwise there are two distinct $\mathbb{Q}$-rational effective divisors D_1 and D_2 of degree d that are linearly equivalent, which means that there is a rational function f on $X_1(p)$ defined over $\mathbb{Q}$ whose divisor is $D_1 - D_2$, and hence f has degree $\leq d$. This contradicts the condition on the $\mathbb{Q}$-gonality. By Proposition 3.4, the reduction map $\mathrm{red}_2 : J_1(p)(\mathbb{Q}) \to J_1(p)(\mathbb{F}_2)$ is injective as well, and therefore $\mathrm{red}_2 \circ \iota = \iota \circ \mathrm{red}_2$ is injective, which implies that red_2 is injective on $X_1(p)^{(d)}(\mathbb{Q})$:

$$\begin{array}{ccc} X_1(p)^{(d)}(\mathbb{Q}) & \xrightarrow{\iota} & J_1(p)(\mathbb{Q}) \\ \downarrow{\scriptstyle \mathrm{red}_2} & & \downarrow{\scriptstyle \mathrm{red}_2} \\ X_1(p)^{(d)}(\mathbb{F}_2) & \xrightarrow{\iota} & J_1(p)(\mathbb{F}_2) \end{array}$$

□

The following is an excerpt of [Derickx and van Hoeij 2014, Table 1]. We write $\mathrm{gon}_{\mathbb{Q}}(X)$ for the $\mathbb{Q}$-gonality of a curve X:

p	11	13	17	19	23	29	31
$\mathrm{gon}_{\mathbb{Q}}(X_1(p))$	2	2	4	5	7	11	12

Also, it follows from [Derickx and van Hoeij 2014, Theorem 3] that $\mathrm{gon}_{\mathbb{Q}}(X_1(p)) > 8$ for $p \in \{41, 47, 59, 71\}$. We deduce the following.

Corollary 3.6. *For the following values of d and p, assumption* (a) *in Lemma 1.7 is satisfied for $\ell = 2$:*

$$\begin{aligned} d = 3 \quad &\textit{and} \quad p \in \{17, 19, 23, 29, 31, 41, 47, 59, 71\}, \\ d = 4 \quad &\textit{and} \quad p \in \{19, 23, 29, 31, 41, 47, 59, 71\}, \\ d = 5 \quad &\textit{and} \quad p \in \{23, 29, 31, 41, 47, 59, 71\}, \\ d = 6 \quad &\textit{and} \quad p \in \{23, 29, 31, 41, 47, 59, 71\}, \\ d = 7 \quad &\textit{and} \quad p \in \{29, 31, 41, 47, 59, 71\}. \end{aligned}$$

We now consider assumption (b) of Lemma 1.7 for $p = 29, 31, 41$. We do this here rather than in Section 8, since the computations we do to show that the assumption is satisfied are closely related to those we do to establish Proposition 3.4.

Lemma 3.7. *For $p \in \{29, 31, 41\}$, $d \leq 7$ and $\ell = 2$, assumption* (b) *of Lemma 1.7 is satisfied.*

Proof. For $d \leq 4$, we have that $p > (2^{d/2}+1)^2$, and the claim follows from Lemma 1.9. For $(d, p) = (5, 29)$, we observe that there is no elliptic curve over $\mathbb{F}_{2^5}$ with 29 points and that the cusps that are not images of rational cusps are not defined over $\mathbb{F}_{2^5}$, so there are no points $\bar{x}$ as in assumption (b).

In the other cases, Corollary 3.3 tells us that $J_1(p)(\mathbb{Q})$ is generated by the differences of the rational cusps. This implies that the reduction mod 2 of any $\mathbb{Q}$-rational point of $X_1(p)^{(d)}$ must map into the subgroup of $J_1(p)(\mathbb{F}_2)$ that is generated by the differences of the images of the rational cusps. We verify that the points $\bar{x}$ as in assumption (b) do not map into that subgroup, which by the above shows that these points are not in the image of the reduction map. This implies the claim. This computation is done together with the computations we do to prove Proposition 3.4. □

Remark. In a similar way as in Proposition 3.4, we can use the following alternative approach for $p = 41$. There is no elliptic curve E over $\mathbb{F}_{2^e}$ with $41 \mid \#E(\mathbb{F}_{2^e})$ if $e \leq 7$ and $e \neq 5$. 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 between $X_1(41)$ and $X_0(41)$ with H of index 4. 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 generated by the differences of the four cusps, so that we can conclude in the same way as above.

4. Formal immersions

When p is not a rank-zero prime, so that $J_1(p)(\mathbb{Q})$ has positive rank, then the reduction map $J_1(p)(\mathbb{Q}) \to J_1(p)(\mathbb{F}_\ell)$ is no longer injective. This means that we need to find a more sophisticated argument to verify assumption (a) of Lemma 1.7.

As mentioned in the introduction, one key idea here is to use a morphism $t : J_1(p) \to A$ of abelian varieties over $\mathbb{Z}_{(\ell)}$. We obtain the following commutative diagram:

$$\begin{array}{ccccc} X_1(p)^{(d)}(\mathbb{Q}) & \xrightarrow{\iota} & J_1(p)(\mathbb{Q}) & \xrightarrow{t} & A(\mathbb{Q}) \\ \downarrow{\scriptstyle \mathrm{red}_\ell} & & \downarrow{\scriptstyle \mathrm{red}_\ell} & & \downarrow{\scriptstyle \mathrm{red}_\ell} \\ X_1(p)^{(d)}(\mathbb{F}_\ell) & \xrightarrow{\iota_\ell} & J_1(p)(\mathbb{F}_\ell) & \xrightarrow{t_\ell} & A(\mathbb{F}_\ell) \end{array} \tag{4-1}$$

Let $\bar{x} \in X_1(p)^{(d)}(\mathbb{F}_\ell)$ be some point. Assuming that red_ℓ is injective on $t(J_1(p)(\mathbb{Q}))$, it will follow that the residue class of $\bar{x}$ contains at most one rational point, if we can show that the diagonal composition $\mathrm{red}_\ell \circ t \circ \iota = t_\ell \circ \iota_\ell \circ \mathrm{red}_\ell$ is injective on the residue class of $\bar{x}$.

The strategy for doing that is to take A such that $A(\mathbb{Q})$, or at least $t(J_1(p)(\mathbb{Q}))$, is finite; then the reduction map on $A(\mathbb{Q})$ (or the image of t) will be injective when ℓ is odd; when $\ell = 2$, we can ensure that the reduction map is injective on $t(J_1(p)(\mathbb{Q}))$ by making sure that this image has odd order. It then remains to show that $t \circ \iota$ is injective on the residue class of any point $\bar{x} \in X_1(p)^{(d)}(\mathbb{F}_\ell)$. To do this, we show that $t \circ \iota$ is a formal immersion at each of the points $\bar{x}$ as above. We recall the definition below.

First, some notation. We write $\mathcal{O}_X$ for the structure sheaf of a scheme X, $\mathcal{O}_{X,x}$ for its local ring at a point x of X, and $\widehat{\mathcal{O}}_{X,x}$ for the completion of the local ring with respect to its maximal ideal $\mathfrak{m}_{X,x}$.

Definition 4.1. Let $\phi : X \to Y$ be a morphism of noetherian schemes and let $x \in X$ be a point. Then ϕ is a *formal immersion at* x if the induced local homomorphism on complete local rings

$$\hat{\phi}^* : \widehat{\mathcal{O}}_{Y,\phi(x)} \to \widehat{\mathcal{O}}_{X,x}$$

is surjective.

The relevant property of formal immersions for our purposes is the following; this is (a consequence of) [Parent 1999, Lemma 4.13].

Lemma 4.2. *Let $\phi : X \to Y$ be a morphism of noetherian schemes over $\mathbb{Z}_{(\ell)}$ that is a formal immersion at $x \in X(\mathbb{F}_\ell)$. Then ϕ induces an injective map on residue classes*

$$\phi : \mathrm{red}_\ell^{-1}(x) \to \mathrm{red}_\ell^{-1}(\phi(x)).$$

Corollary 4.3. *Let $d \in \mathbb{Z}_{\geq 1}$ and let $\ell \neq p$ be primes. Let $t : J_1(p) \to A$ be a morphism of abelian schemes over $\mathbb{Z}_{(\ell)}$ such that*

(i) *$t(J_1(p)(\mathbb{Q}))$ is finite,*

(ii) *$\ell > 2$ or $\#t(J_1(p)(\mathbb{Q}))$ is odd,*

(iii) *$t \circ \iota$ is a formal immersion at all $\bar{x} \in X_1(p)^{(d)}(\mathbb{F}_\ell)$ that are sums of images of rational cusps on $X_1(p)$.*

Then assumption (a) *of Lemma 1.7 is satisfied.*

Proof. Note that $X_1(p)$ and $J_1(p)$ have good reduction at ℓ, and hence $J_1(p)$ can be considered as an abelian scheme over $\mathbb{Z}_{(\ell)}$.

Let $x, x' \in X_1(p)^{(d)}(\mathbb{Q})$ be in the residue class of a point $\bar{x}$ that is a sum of images of rational cusps and write $y = t(\iota(x))$, $y' = t(\iota(x'))$. Because x and x' are in the same residue class, the same is true of y and y'. It follows from conditions (i) and (ii) that $\mathrm{red}_\ell : t(J_1(p)(\mathbb{Q})) \to A(\mathbb{F}_\ell)$ is injective, which implies that $y = y'$. By condition (iii) and Lemma 4.2, $t \circ \iota$ is injective on the residue class of $\bar{x}$, which finally shows that $x = x'$. □

Remark. If $\ell = 2$ and we take commuting $t_1, t_2 \in \mathrm{End}_{\mathbb{Q}}(J_1(p))$ such that $t_1(J_1(p)(\mathbb{Q}))$ is finite and t_2 kills the 2-torsion subgroup of $J_1(p)(\mathbb{Q})$, then the conclusion of Corollary 4.3 holds for $t = t_1 t_2$ also when $\#t(J_1(p)(\mathbb{Q}))$ is even (assuming condition (iii) is satisfied); see [Parent 2000, Theorem 1.10].

Writing $A_1 = \mathrm{im}(t_1)$ and taking $A = \mathrm{im}(t)$ without loss of generality, we have the following commutative diagram:

$$\begin{array}{ccccccc} X_1(p)^{(d)}(\mathbb{Q}) & \xrightarrow{\iota} & J_1(p)(\mathbb{Q}) & \xrightarrow{t_1} & A_1(\mathbb{Q}) & \xrightarrow{t_2} & A(\mathbb{Q}) \\ \downarrow{\scriptstyle \mathrm{red}_2} & & \downarrow{\scriptstyle \mathrm{red}_2} & & \downarrow{\scriptstyle \mathrm{red}_2} & & \downarrow{\scriptstyle \mathrm{red}_2} \\ X_1(p)^{(d)}(\mathbb{F}_2) & \xrightarrow{\iota} & J_1(p)(\mathbb{F}_2) & \xrightarrow{t_1} & A_1(\mathbb{F}_2) & \xrightarrow{t_2} & A(\mathbb{F}_2) \end{array}$$

Take $x, x' \in X_1(p)^{(d)}(\mathbb{Q})$ with the same reduction $\bar{x}$ mod 2, such that $\bar{x}$ is a sum of images of rational cusps. Then $t_1(\iota(x') - \iota(x))$ is in the kernel of reduction mod 2 of $A_1(\mathbb{Q})$, which (since $A_1(\mathbb{Q})$ is finite) consists of 2-torsion points, so $t(\iota(x')) = t(\iota(x))$ by the assumption on t_2. We can then conclude as in the proof above.

In our intended application, the set $X_1(p)^{(d)}(\mathbb{F}_\ell)$ can be quite large: the curve $X_1(p)$ has $(p-1)/2$ $\mathbb{Q}$-rational cusps; assuming that they account for all of $X_1(p)^{(d)}(\mathbb{F}_\ell)$, the latter set has $\binom{(p-1)/2+d-1}{d}$ elements. Corollary 4.3 requires us to check that $t \circ \iota$ is a formal immersion at each of these points. To reduce the necessary computational effort, we now show how we can use curves intermediate between $X_1(p)$ and $X_0(p)$ that have fewer cusps.

Corollary 4.4. *Let $d \in \mathbb{Z}_{\geq 1}$ and let $\ell \neq p$ be primes. Let X_H be an intermediate curve between $X_1(p)$ and $X_0(p)$. Fix $x_0 \in X_H^{(d)}(\mathbb{Q})$ and define $\iota_H : X_H^{(d)} \to J_H$ using x_0 as basepoint. Let $t : J_H \to A$ be a morphism of abelian schemes over $\mathbb{Z}_{(\ell)}$ such that*

(i) *$t(J_H(\mathbb{Q}))$ is finite,*

(ii) *$\ell > 2$ or $\#t(J_H(\mathbb{Q}))$ is odd,*

(iii) *$t \circ \iota_H$ is a formal immersion at all $\bar{x}_H \in X_H^{(d)}(\mathbb{F}_\ell)$ that are sums of images of rational cusps on $X_1(p)$.*

Then assumption (a) *of Lemma 1.7 is satisfied.*

Proof. Let $\bar{x} \in X_1(p)^{(d)}(\mathbb{F}_\ell)$ be a sum of images of rational cusps and take two points $x, x' \in X_1(p)^{(d)}(\mathbb{Q})$ in the residue class of $\bar{x}$, where we take x to be the unique sum of rational cusps in this residue class. Write x_H, x'_H for their images in $X_H^{(d)}(\mathbb{Q})$. Then $\bar{x}_H := \mathrm{red}_\ell(x_H) = \mathrm{red}_\ell(x'_H)$ is a sum of images of rational cusps on $X_1(p)$. Arguing as in the proof of Corollary 4.3, we see that $x'_H = x_H$; in particular, x'_H is a sum of images of rational cusps on $X_1(p)$, since this is true for x_H. The set of rational cusps on $X_1(p)$ is the full preimage of the cusp $\infty \in X_0(p)(\mathbb{Q})$. This implies that all points in $X_1(p)^{(d)}(\mathbb{Q})$ that are preimages of x_H under the obvious map are sums of rational cusps. So x' is a sum of rational cusps as well. But red_ℓ is injective on sums of rational cusps (since reduction mod ℓ is injective on cusps; see [Deligne and Rapoport 1973, Theorem IV.3.4]), and hence $x' = x$. □

5. Computational verification of assumption (a)

We use Corollary 4.4 to show that assumption (a) of Lemma 1.7 holds for the relevant pairs (d, p). To verify the assumptions of Corollary 4.4, we need to do essentially two things: we have to find a suitable

morphism t of abelian schemes that satisfies conditions (i) and (ii), and we have to check that $t \circ \iota$ is a formal immersion at all points in $\bar{x}_H \in X_H^{(d)}(\mathbb{F}_\ell)$ that are sums of images of rational cusps on $X_1(p)$.

To satisfy condition (i), we take a morphism t that factors through the winding quotient $J_H^{\boldsymbol{e}}$; then $t(J_H(\mathbb{Q}))$ is contained in the image of $J_H^{\boldsymbol{e}}(\mathbb{Q})$ under a morphism of abelian varieties. Since, by Theorem 2.2, $J_H^{\boldsymbol{e}}(\mathbb{Q})$ is finite, $t(J_H(\mathbb{Q}))$ is finite as well. One possibility is to take the projection $J_H \to J_H^{\boldsymbol{e}}$. If we choose $\ell \geq 3$, then condition (ii) is also satisfied. This was 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 $J_0(p^n)$ with $\ell = 3$ or 5 in the argument of [Parent 1999]. The proof of Oesterlé's bound uses $\ell = 3$; see Section 6.

If we take for t an element of the Hecke algebra $\mathbb{T} \subseteq \mathrm{End}_{\mathbb{Q}}(J_H)$, then the condition for t to factor via the winding quotient is that $t \cdot \mathrm{Ann}(\boldsymbol{e}) = 0$ in $\mathbb{T}$. We obtain such t as follows. This is essentially [Parent 2000, Lemma 1.9]; we extend the statement slightly by removing the condition that the characteristic polynomial of t_0 (acting on the space of cusp forms) is squarefree.

Proposition 5.1. *Let $t_0 \in \mathbb{T}$ with factored characteristic polynomial $P(X) = \prod_{i=1}^{n} P_i(X)^{e_i}$ with respect to its action on $H^0(X_H, \Omega^1)$. Set*

$$I := \big\{ i \in \{1, \ldots, n\} \mid (P/P_i)(t_0) \cdot \boldsymbol{e} = 0 \text{ or } e_i \geq 2 \big\};$$

then $t_1(t_0) := \prod_{i \in I} P_i^{e_i}(t_0)$ is such that $t_1(t_0) \cdot \mathrm{Ann}(\boldsymbol{e}) = 0$.

Proof. The proof is basically the same as that in [Parent 2000, § 2.5], noting that the factors $P_i^{e_i}(t_0)$ with $e_i \geq 2$ in the product defining $t_1(t_0)$ are used to kill any factor of the Hecke algebra for which we cannot simply decide whether it is contained in $\mathrm{Ann}(\boldsymbol{e})$. □

We note that we can compute $P(X)$ and test the condition $(P/P_i)(t_0) \cdot \boldsymbol{e} = 0$ explicitly using modular symbols, so we can determine $t_1(t_0)$ explicitly for any given t_0. We see that $t_1(t_0)$ satisfies condition (i) for every $t_0 \in \mathbb{T}$.

To satisfy condition (ii) when $\ell = 2$, we use Proposition 2.3, which implies that for q an odd prime not dividing N, $T_q - \langle q \rangle - q$ kills the rational torsion subgroup of J_H. Combining this with Proposition 5.1 gives the following version of "Parent's trick".

Corollary 5.2. *Let X_H be an intermediate curve between $X_1(p)$ and $X_0(p)$. Let $t_0 \in \mathbb{T}$ and let $q \neq p$ be an odd prime. Then*

$$t := t_1(t_0) \cdot (T_q - \langle q \rangle - q) \in \mathbb{T},$$

considered as an element of $\mathrm{End}_{\mathbb{Q}}(J_H)$, satisfies conditions (i) *and* (ii) *of Corollary 4.4 for $\ell = 2$. If $X_H = X_0(p)$ and $p \not\equiv 1 \bmod 8$, then $t := t_1(t_0)$ satisfies both conditions.*

Proof. By Proposition 5.1 and the discussion preceding it, $t_1(t_0)$ satisfies condition (i). Obviously this condition still holds after composing $t_1(t_0)$ with some further morphism. By Proposition 2.3, the factor $T_q - \langle q \rangle - q$ kills the torsion in $t_1(t_0)(J_H(\mathbb{Q})) \subseteq J_H(\mathbb{Q})_{\mathrm{tors}}$, which implies that $t(J_H(\mathbb{Q})) = 0$, so that condition (ii) also holds.

It is known that $J_0(p)(\mathbb{Q})_{\mathrm{tors}}$ is cyclic of order $(p-1)/\gcd(p-1,12)$, generated by the difference of the two (rational) cusps; see [Mazur 1977, Theorem 1]. This implies that the rational torsion group of $J_0(p)$ has odd order when $p \not\equiv 1 \bmod 8$, and so condition (ii) is automatically satisfied. □

We still need a way of verifying condition (iii) of Corollary 4.4. This is provided by the following version of "Kamienny's criterion" as given in [Parent 2000, Theorem 1.10, Proposition 2.7]. Parent states this criterion for $X_1(p)$ in place of X_H, but the generalization is immediate.

Proposition 5.3. *Let $H \subseteq (\mathbb{Z}/p\mathbb{Z})^\times/\{\pm 1\}$ be a subgroup. Let $\ell \neq p$ be a prime and consider $t = t_1(t_0)$ as in Proposition 5.1 when ℓ is odd, or t as in Corollary 5.2 when $\ell = 2$. Then $t \circ \iota$ is a formal immersion at all $\bar{x}_H \in X_H^{(d)}(\mathbb{F}_\ell)$ that are sums of images of rational cusps on $X_1(p)$ if for all partitions $d = n_1 + \ldots + n_m$ with $n_1 \geq \cdots \geq n_m$ and all m-tuples $(d_1 = 1, d_2, \ldots, d_m)$ of integers representing pairwise distinct elements of H, the d Hecke operators*

$$(T_i\langle d_j\rangle t)_{\substack{j=1,\ldots,m,\\ i=1,\ldots,n_j}} \tag{5-1}$$

are $\mathbb{F}_\ell$-linearly independent in $\mathbb{T} \otimes \mathbb{F}_\ell$, where $\mathbb{T}$ is considered as a subalgebra of $\mathrm{End}_{\mathbb{Q}}(J_H)$.

We note that we can check the criterion for any given t by a computation with modular symbols.

This criterion was first established by Kamienny [1992b] for $X_0(p)$. In this case, the condition simplifies to:

The d Hecke operators $T_1 t, T_2 t, \ldots, T_d t$ are $\mathbb{F}_\ell$-linearly independent in $\mathbb{T} \otimes \mathbb{F}_\ell$.

Implementing the criterion implied by Corollary 5.2 and Proposition 5.3 for $X_0(p)$ and running the resulting code gives the following. We take as basepoint for ι the point given by d times the cusp ∞ on $X_0(p)$. Note that $2281 = \lfloor (1+3^{7/2})^2 \rfloor$; larger primes will be dealt with using Oesterlé's bound.

Lemma 5.4. *For each of the following choices of $3 \leq d \leq 7$ and a prime p, there is $t \in \mathrm{End}_{\mathbb{Q}}(J_0(p))$ as in Corollary 5.2 for $\ell = 2$ such that $t \circ \iota : X_0(p)^{(d)}_{\mathbb{Z}_{(2)}} \to J_0(p)_{\mathbb{Z}_{(2)}}$ is a formal immersion at the point of $X_0(p)^{(d)}(\mathbb{F}_2)$ corresponding to d times the cusp ∞:*

$$\begin{aligned}
&d = 3 \quad \textit{and} \quad && 47 \leq p \leq 2281,\ p \neq 73, 79;\\
&d = 4 \quad \textit{and} \quad && p \in \{47, 59, 71, 83, 89\} \quad \textit{or} \quad 103 \leq p \leq 2281;\\
&d = 5 \quad \textit{and} \quad && p \in \{59, 71, 83\} \quad \textit{or} \quad 103 \leq p \leq 2281;\\
&d = 6 \quad \textit{and} \quad && p \in \{71, 107\} \quad \textit{or} \quad 127 \leq p \leq 2281,\ p \neq 193;\\
&d = 7 \quad \textit{and} \quad && p = 131 \quad \textit{or} \quad 139 \leq p \leq 2281,\ p \neq 157, 193.
\end{aligned}$$

Proof. We try $t_0 = T_n$ for $2 \leq n \leq 60$, and when $p \equiv 1 \bmod 8$, we try for each t_0 the additional factor $T_q - (q+1)$ for primes $3 \leq q \leq 20$ until either the criterion is satisfied or else all combinations are exhausted. (Actually, $n \leq 14$ and $q \in \{3, 5\}$ would be enough, as the computation reveals.) The computation took about 1.5 hours. We note that to exclude $p = 163$ for $d = 7$, we actually needed the statement of Corollary 5.2 that $t = t_1(t_0)$ is sufficient when $p \not\equiv 1 \bmod 8$ (which also helps to speed up the computation, since it eliminates the inner loop over q). For $p = 431$ and $d = 7$, taking $t_0 = T_n$ does not

seem to work. We tried random linear combinations of the first few Hecke operators and were successful with $t_0 = T_2 + T_3 - T_7$. □

For the remaining primes p of interest for any given degree d, we use the criterion on an intermediate curve X_H; we try the various groups H ordered by increasing index in $(\mathbb{Z}/p\mathbb{Z})^\times/\{\pm 1\}$, since smaller index means that we have to deal with smaller objects, leading to a faster computation.

If we were to use the criterion of Proposition 5.3 literally, then we would have to run through a potentially very large number of partitions of d combined with choices of d_j. We use the following trick to speed up the computation.

Lemma 5.5. *Let $H \subseteq (\mathbb{Z}/p\mathbb{Z})^\times/\{\pm 1\}$ be a subgroup. Let $\ell \neq p$ be a prime, d an integer and $t \in \mathbb{T}$, viewed as an endomorphism of J_H. Let $D \subseteq \mathbb{Z}$ be a set of representatives of the cosets of H with $1 \in D$. Define the set*

$$I := \big\{(1,i) \mid 1 \le i \le d\big\} \cup \big\{(k,i) \mid 1 \le i \le \lfloor \tfrac{1}{2}d \rfloor, 1 \neq k \in D\big\}.$$

Suppose that there is no $\mathbb{F}_\ell$-linear dependence among at most d of the images in $\mathbb{T} \otimes \mathbb{F}_\ell$ of the elements $t_{(k,i)} := T_i\langle k\rangle t$ for $(k,i) \in I$, where we consider $\mathbb{T}$ as a subalgebra of $\mathrm{End}_{\mathbb{Q}}(J_H)$. Then the criterion of Proposition 5.3 is satisfied.

Proof. Assume the criterion fails. Then there is a partition $d = n_1 + \cdots + n_m$ with $n_1 \ge \cdots \ge n_m$ and there are $d_1 = 1, d_2, \ldots, d_m \in D$ pairwise distinct such that the d operators $T_i\langle d_j\rangle t$ for $1 \le j \le m$ and $1 \le i \le n_j$ are linearly dependent in $\mathbb{T} \otimes \mathbb{F}_\ell$. But these operators are all of the form $t_{(k,i)}$ (note $n_j \le \lfloor \frac{1}{2}d \rfloor$ for $j \ge 2$), so this would produce a linear dependence mod ℓ among d of the $t_{(k,i)}$; this is a contradiction. □

When implementing this, we can in addition look at each linear relation of weight at most d between the elements in the lemma and check if it is indeed of the "forbidden" form as given in Proposition 5.3. In the cases of interest, the relation space has low enough dimension to allow for the enumeration of all relations and performing this check. We use algorithms for binary linear codes that are included in Magma to do this efficiently.

We obtain the following result.

Lemma 5.6. *For each of the following choices of $3 \le d \le 7$ and a prime p, there is a subgroup H of $(\mathbb{Z}/p\mathbb{Z})^\times/\{\pm 1\}$ and $t \in \mathrm{End}_{\mathbb{Q}}(J_H)$ as in Corollary 5.2 for $\ell = 2$ such that $t \circ \iota : X^{(d)}_{H,\mathbb{Z}_{(2)}} \to J_{H,\mathbb{Z}_{(2)}}$ is a formal immersion at all points of $X^{(d)}_H(\mathbb{F}_2)$ that are sums of images of rational cusps on $X_1(p)$:*

$$\begin{aligned}
&d = 3 \quad \text{and} \quad 19 \le p \le 2281;\\
&d = 4 \quad \text{and} \quad 19 \le p \le 2281,\ p \neq 29;\\
&d = 5 \quad \text{and} \quad 23 \le p \le 2281,\ p \neq 29;\\
&d = 6 \quad \text{and} \quad 23 \le p \le 2281,\ p \neq 29;\\
&d = 7 \quad \text{and} \quad 37 \le p \le 2281.
\end{aligned}$$

Proof. For each pair (d, p) that is not covered by Lemma 5.4, we check the criterion of Lemma 5.5 for subgroups H by increasing index. For each H, we again try $t_0 = T_n$ for $2 \le n \le 60$ and the second factor given by primes $3 \le q \le 20$. The most involved computation is for $d = 7$ and $p = 107$, where we have to take the trivial subgroup H corresponding to $J_1(107)$; this computation took about 35 minutes. Most of the other cases just take a few seconds, a small number of them a few minutes. □

Proposition 1.8 now follows from Lemma 5.6 and Corollary 3.6.

6. A proof of Oesterlé's bound

The purpose of this section is to provide a proof of Oesterlé's bound (1-1) and thus close a gap in the literature. Oesterlé gives a proof in his notes [1994], which have been available to the people working in the field, but a proof has never appeared in print. The proof below is based on these notes, which Oesterlé kindly provided to us; in particular, we do not claim originality for anything in this section: the ideas are all Oesterlé's. We will use results that are available in the literature by now to simplify the exposition in some places. We state the result of this section as a theorem.

Theorem 6.1 (Oesterlé). *Let $d \ge 3$. If $p > (3^{d/2} + 1)^2$ is a prime, then $p \notin S(d)$.*

We can restrict to $d \ge 3$ here, since the cases $d = 1$ and $d = 2$ have been dealt with by Mazur and Kamienny, respectively.

We will work with $\ell = 3$. By Lemma 1.9, assumption (b) of Lemma 1.7 is always satisfied when $p > (3^{d/2} + 1)^2$. So it is sufficient to show that assumption (a) of Lemma 1.7 holds. This in turn is done by using the formal immersion criterion via the winding quotient of $J_0(p)$. For sufficiently large d, this follows from the following result.

Proposition 6.2. *If $d \ge 3$ and $p \ge 65(2d)^6$ is a prime, then the map*

$$f_{d,p} : X_0(p)^{(d)} \xrightarrow{\iota} J_0(p) \to J_0^e(p)$$

is a formal immersion at the point $\bar{x} \in X_0(p)^{(d)}(\mathbb{F}_3)$ that is the reduction mod 3 of d times the cusp ∞ on $X_0(p)$.

In particular, Theorem 6.1 holds for $d \ge 26$.

Proof. The first statement is a consequence of [Parent 1999, Theorem 1.8 and Proposition 1.9]. Since $65(2d)^6 < (3^{d/2} + 1)^2$ when $d \ge 26$, the statement of Theorem 6.1 follows for such d by the discussion above. □

Oesterlé proves a similar statement with a slightly worse bound on p; Parent uses the same underlying approach.

In principle, Proposition 6.2 reduces the proof of Theorem 6.1 to a finite problem: for each $3 \le d \le 25$ and each prime p such that $(3^{d/2}+1)^2 < p < 65(2d)^6$, we have to check that the map in Proposition 6.2 is a formal immersion at the relevant point, which can be done via Kamienny's criterion given in Proposition 5.3. However, the primes we would have to deal with in this way get much too large and there are way too

many of them to make this practical. So instead, we need a criterion that allows us to deal with all (or many) of these primes at the same time.

One idea that Oesterlé uses here (and also to prove a statement similar to Proposition 6.2 above) is to make use of the intersection pairing on $H_1(X_0(p)(\mathbb{C}), \mathbb{Z})$, which is an alternating perfect pairing into $\mathbb{Z}$. We will denote this pairing by $\bullet$.

We will use the following version of Kamienny's criterion. Recall from Definition 2.1 the winding element $\boldsymbol{e} \in H_1(X_0(p)(\mathbb{C}), \mathbb{Q})$.

Proposition 6.3. *The map $f_{d,p}$ as in Proposition 6.2 is a formal immersion at $\bar{x}$ if (and only if) the images of $T_1\boldsymbol{e}, \ldots, T_d\boldsymbol{e}$ in $\mathbb{T}\boldsymbol{e}/3\mathbb{T}\boldsymbol{e}$ are linearly independent over $\mathbb{F}_3$.*

Proof. This is [Parent 1999, Theorem 4.18] for $l = 3$. □

To make use of the intersection pairing, we have to move the elements $T_n\boldsymbol{e} \in H_1(X_0(p)(\mathbb{C}), \mathbb{Q})$ into $H_1(X_0(p)(\mathbb{C}), \mathbb{Z})$. The Hecke operator $T_2 - 3$ sends $\boldsymbol{e}$ into $H_1(X_0(p)(\mathbb{C}), \mathbb{Z})$, since the action of T_2, viewed as a correspondence on $X_0(p)$, multiplies the cusps 0 and ∞ by 3, so that the boundary of $-(T_2-3)\cdot\{0, \infty\}$ is zero. In the same way, we see that $(T_n - \sigma_1(n))\boldsymbol{e} \in H_1(X_0(p)(\mathbb{C}), \mathbb{Z})$ when $n < p$; here $\sigma_1(n)$ denotes the sum of (positive) divisors of n. (This is true in general when $p \nmid n$; when $p \mid n$, one has to replace $\sigma_1(n)$ with the sum of divisors not divisible by p.)

Corollary 6.4. *If $p > (3^{d/2} + 1)^2$ and the images of*

$$(T_2 - 3)T_1\boldsymbol{e}, \ldots, (T_2 - 3)T_d\boldsymbol{e}$$

in $H_1(X_0(p)(\mathbb{C}), \mathbb{F}_3)$ are linearly independent over $\mathbb{F}_3$, then $p \notin S(d)$.

Proof. We show that $f_{d,p}$ is a formal immersion at $\bar{x}$, which implies the claim. Assume that this is not the case. By Proposition 6.3, there are integers $\lambda_1, \ldots, \lambda_d$, not all divisible by 3, such that $\lambda_1 T_1\boldsymbol{e} + \cdots + \lambda_d T_d\boldsymbol{e} \in 3\mathbb{T}\boldsymbol{e}$. Multiplying by $T_2 - 3$, this gives

$$\lambda_1(T_2 - 3)T_1\boldsymbol{e} + \cdots + \lambda_d(T_2 - 3)T_d\boldsymbol{e} \in 3(T_2 - 3)\mathbb{T}\boldsymbol{e} \subset 3H_1(X_0(p)(\mathbb{C}), \mathbb{Z}),$$

with all terms on the left contained in $H_1(X_0(p)(\mathbb{C}), \mathbb{Z})$. Reducing this relation mod 3 shows that the images of $(T_2 - 3)T_1\boldsymbol{e}, \ldots, (T_2 - 3)T_d\boldsymbol{e}$ in $H_1(X_0(p)(\mathbb{C}), \mathbb{F}_3)$ are linearly dependent. □

We now define the following Hecke operators.

Definition 6.5. Let $n \geq 1$. We set

$$T'_n = \sum_{m|n} \mu\Big(\frac{n}{m}\Big) T_m,$$

where μ is the Möbius function, and

$$L_n = T'_{2n} - 2T'_n.$$

Then $T_n - \sigma_1(n) = \sum_{m|n}(T'_m - m)$. Using the relations

$$T_2 T_m = \begin{cases} T_{2m} & \text{if } m \text{ is odd,} \\ T_{2m} + 2T_{m/2} & \text{if } m \text{ is even,} \end{cases}$$

we find that

$$(T_2 - 3)T_n' = \begin{cases} L_n & \text{if } n \text{ is odd,} \\ L_n - L_{n/2} & \text{if } n \text{ is even.} \end{cases}$$

Corollary 6.6. *If $p > (3^{d/2} + 1)^2$ and the images of*

$$L_1\boldsymbol{e}, \ldots, L_d\boldsymbol{e}$$

in $H_1(X_0(p)(\mathbb{C}), \mathbb{F}_3)$ are linearly independent over $\mathbb{F}_3$, then $p \notin S(d)$.

Proof. The relations deduced above show that the $\mathbb{Z}$-submodule of $\mathbb{T}$ generated by $L_1, \ldots, L_d$ is the same as the $\mathbb{Z}$-submodule generated by $(T_2 - 3)T_1, \ldots, (T_2 - 3)T_d$. Now use Corollary 6.4. □

We now introduce notation for certain modular symbols, following [Merel 1996, Section 2]. If $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$, then the modular symbol $\{\gamma 0, \gamma\infty\}$ depends only on the coset $\Gamma_0(p)\gamma$, which in turn depends only on the image of c/d in $\mathbb{P}^1(\mathbb{F}_p)$. We denote this modular symbol by $\xi(c/d)$. If k is an integer coprime with p, then $\xi(k) = \{0, 1/k\} \in H_1(X_0(p)(\mathbb{C}), \mathbb{Z})$, since the cusp $1/k$ is $\Gamma_0(p)$-equivalent to 0.

The following result is crucial; we defer its proof until later and first show how Theorem 6.1 can be deduced from it with some computation. For $M \geq 3$ an odd integer, we define

$$\varepsilon_M : (\mathbb{Z}/M\mathbb{Z})^\times \to \{0, 1\}$$

so that $\varepsilon_M(a + M\mathbb{Z}) = 0$ if $1 \leq a < \frac{1}{2}M$ and $\varepsilon_M(a + M\mathbb{Z}) = 1$ if $\frac{1}{2}M < a < M$. We extend ε_M to a map on all rational numbers a/b with numerator and denominator coprime to M by applying it to the image of a/b in $(\mathbb{Z}/M\mathbb{Z})^\times$.

Lemma 6.7. *Let $d \geq 1$ be an integer, let $M \geq 3$ be an odd integer and let $p > 2dM$ be a prime. Let $u \in \mathbb{Z}$ be such that $pu \equiv 1 \bmod M$. Then for a coprime to M and $1 \leq n \leq d$, we have*

$$L_n\boldsymbol{e} \bullet \left\{0, \frac{a}{M}\right\} = \varepsilon_M(na) - \varepsilon_M(nu/a).$$

Corollary 6.8 [Oesterlé 1994, Proposition 8]. *Let d and M be as in Lemma 6.7 and fix $u \in \mathbb{Z}$ coprime with M. If the matrix*

$$\big(\varepsilon_M(na) - \varepsilon_M(nu/a)\big)_{1 \leq n \leq d,\, a \in (\mathbb{Z}/M\mathbb{Z})^\times},$$

with entries taken in $\mathbb{F}_3$, has rank d, then $p \notin S(d)$ for all primes

$$p > \max\{2dM, (3^{d/2} + 1)^2\} \quad \textit{such that } pu \equiv 1 \bmod M.$$

Proof. By Lemma 6.7, the matrix entries are the intersection numbers, taken mod 3, between $L_n\boldsymbol{e}$ and $\{0, a/M\}$ in $H_1(X_0(p)(\mathbb{C}), \mathbb{Z})$, when p is a prime as in the statement. So when the matrix has rank d, this implies that $L_1\boldsymbol{e}, \ldots, L_d\boldsymbol{e}$ are linearly independent mod 3, and the claim follows from Corollary 6.6. □

Proof of Theorem 6.1. The following table gives, for each $3 \le d \le 25$, a value of M as in Corollary 6.8 such that the matrix above has rank d for all $u \in (\mathbb{Z}/M\mathbb{Z})^\times$. By Corollary 6.8, this proves Theorem 6.1 for all $p > \max\{2dM, (3^{d/2}+1)^2\}$.

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

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

Note that $2dM < (3^{d/2}+1)^2$ for $d \ge 6$. We have already verified the formal immersion criterion (with $\ell = 2$) for the primes between $(3^{d/2}+1)^2$ and $2dM$ for $3 \le d \le 5$ in Lemma 5.6, which implies $p \notin S(d)$ for these primes as well. □

Remark. Oesterlé deals with the remaining primes p for $3 \le d \le 5$ by computing the intersection products $I_n \boldsymbol{e} \bullet \xi(k)$ for $1 \le k \le p-1$ and $1 \le n \le d$, where $I_1 = (p-1)/\gcd(p-1, 12)$ is the order of $J_0(p)(\mathbb{Q})_{\mathrm{tors}}$ and $I_n = T_n' - n$ for $n \ge 2$, and verifying that the resulting matrix has rank d (even when reduced modulo any prime $\ell \ge 3$). This works for all cases except $p = 43$ and $p = 73$ for $d = 3$. For $p = 73$, he has a separate argument, whereas he does not mention $p = 43$ further, even though the maximal d for which the rank condition is satisfied is $d = 2$ according to the table at the end of [Oesterlé 1994, Section 7].

From now on until the end of this section, the degree d of the field of definition of the elliptic curves will be irrelevant. We will therefore feel free to use "d" as a local variable as in the definition of M_n below, and hope that this will not lead to confusion.

It remains to prove Lemma 6.7. We follow Oesterlé's notes quite closely here (modulo some changes of notation). We remark that Corollaries 6.12 and 6.13 are in a separate file that Oesterlé made available to Bas Edixhoven and the first author of this paper.

We begin with a result that expresses $(T_n - \sigma_1(n))\boldsymbol{e}$ for $n < p$ in terms of modular symbols.

Lemma 6.9 [Oesterlé 1994, Corollary 2 of Proposition 10]. *For $n < p$, we have in $H_1(X_0(p)(\mathbb{C}), \mathbb{Z})$ that*

$$(T_n - \sigma_1(n))\boldsymbol{e} = -\sum_{(a,b,c,d)\in M_n} \xi\Big(\frac{c}{d}\Big),$$

where

$$M_n = \{(a,b,c,d) \in \mathbb{Z}^4 : a > b \ge 0,\ d > c > 0,\ ad - bc = n\}.$$

Proof. This is [Merel 1996, Lemma 2], using that both sides are contained in $H_1(X_0(p)(\mathbb{C}), \mathbb{Z})$. □

We need a formula for the intersection product. We define the following function on $\mathbb{R}$:

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

Lemma 6.10 [Oesterlé 1994, Equation (37)]. *Let p be a prime and $k, k' \in \{1, \dots, p-1\}$. We write k_* for the unique element of $\{1, \dots, p-1\}$ such that $kk_* \equiv -1 \bmod p$. Then*

$$\xi(k) \bullet \xi(k') = -H(k'-k) + H(k'-k_*) + H(k'_*-k) - H(k'_*-k_*).$$

Proof. By [Merel 1996, Lemma 4], for $k' \notin \{k, k_*\}$, $\xi(k) \bullet \xi(k')$ is the intersection number (-1, 0, or 1) of the oriented line segment joining $e^{2\pi i k'_*/p}$ to $e^{2\pi i k'/p}$ and that joining $e^{2\pi i k_*/p}$ to $e^{2\pi i k/p}$. Otherwise the intersection number is zero, since the pairing is alternating and $\xi(k_*) = -\xi(k)$. The formula we have given follows by considering the various possible cyclic orderings of the four points on the unit circle connected by the two line segments. □

We enlarge M_n slightly and set

$$B_n = \{(a, b, c, d) \in \mathbb{Z}^4 : a > b \geq 0,\ d > c \geq 0,\ ad - bc = n\}$$

(so we allow $c = 0$ here) and write $B_n^{b=0}$, $B_n^{b>0}$, $B_n^{c=0}$ and $B_n^{c>0} = M_n$ for the subsets satisfying the indicated extra condition.

We define, for $n \geq 1$, a prime $p > n$ and $k \in \{1, \dots, p-1\}$, the following two quantities:

$$v_{p,n}(k) = \#\{(a, b, c, d) \in \mathbb{Z}_{>0} : ad + bc = n,\ c \equiv dk \bmod p\},$$
$$v'_{p,n}(k) = \#\{(a, b, c, d) \in \mathbb{Z}_{>0} : ad + bc = n,\ \gcd(c, d) = 1,\ c \equiv dk \bmod p\}.$$

We now give an explicit formula for the intersection number $(T_n - \sigma_1(n))\boldsymbol{e} \bullet \xi(k)$. Its proof by Oesterlé is quite ingenious.

Proposition 6.11 [Oesterlé 1994, Proposition 12]. *Let p be a prime and $k, n \in \{1, \dots, p-1\}$. Then*:

(i) $$(T_n - \sigma_1(n))\boldsymbol{e} \bullet \xi(k) = \sum_{m|n} \left(\left\lfloor \frac{mk}{p} \right\rfloor - \left\lfloor \frac{mk_*}{p} \right\rfloor \right) + v_{p,n}(k) - v_{p,n}(k_*).$$

(ii) $$(T'_n - n)\boldsymbol{e} \bullet \xi(k) = \left\lfloor \frac{nk}{p} \right\rfloor - \left\lfloor \frac{nk_*}{p} \right\rfloor + v'_{p,n}(k) - v'_{p,n}(k_*).$$

Proof. Claim (ii) follows from claim (i) by Möbius inversion. So it suffices to show (i).

We write $k_{c/d}$ for the integer $k \in \{1, \dots, p-1\}$ such that $c \equiv dk \bmod p$, where c and d are integers coprime to p. We extend this to all remaining elements $x \in \mathbb{P}^1(\mathbb{Q})$ by setting $k_x = p$. Then Lemmas 6.9 and 6.10 imply that

$$\begin{aligned}(T_n - \sigma_1(n))\boldsymbol{e} \bullet \xi(k) &= -\sum_{(a,b,c,d) \in M_n} \xi\left(\frac{c}{d}\right) \bullet \xi(k) \\ &= \sum_{(a,b,c,d) \in M_n} \big(H(k - k_{c/d}) - H(k - k_{-d/c}) - H(k_* - k_{c/d}) + H(k_* - k_{-d/c}) \big).\end{aligned}$$

We note that when $c = 0$, all terms under the summation sign are zero (since $k_0 = k_\infty = p$ and $H(k - p) = H(k_* - p) = 0$), so that we can replace the summation over $M_n = B_n^{c>0}$ by a summation over B_n without changing the value of the sum.

We now observe that there is a bijection

$$\phi_n : B_n^{b>0} \to B_n^{c>0}, \quad (a,b,c,d) \mapsto (b, -a+mb, d, -c+mb),$$

where $m = \lceil a/b \rceil \geq 2$ is the unique integer such that $0 \leq -a+mb < b$; its inverse is given by

$$(a,b,c,d) \mapsto (-b+m'a, a, -d+m'c, c) \quad \text{with } m' = \lceil d/c \rceil.$$

We split the sum as follows:

$$\begin{aligned}
&\sum_{(a,b,c,d)\in B_n} \big(H(k-k_{c/d}) - H(k-k_{-d/c})) - H(k_*-k_{c/d}) + H(k_*-k_{-d/c})\big) \\
&= \sum_{(a,b,c,d)\in B_n^{b=0}} \big(H(k-k_{c/d}) - H(k_*-k_{c/d})\big) + \sum_{(a,b,c,d)\in B_n^{b>0}} \big(H(k-k_{c/d}) - H(k_*-k_{c/d})\big) \\
&\quad - \sum_{(a,b,c,d)\in B_n^{c=0}} \big(H(k-k_{-d/c}) - H(k_*-k_{-d/c})\big) - \sum_{(a,b,c,d)\in B_n^{c>0}} \big(H(k-k_{-d/c}) - H(k_*-k_{-d/c})\big). \qquad (6\text{-}1)
\end{aligned}$$

Writing the quadruple in the last sum in (6-1) as $\phi_n(a,b,c,d)$, this then gives

$$\sum_{(a,b,c,d)\in B_n^{b=0}} \big(H(k-k_{c/d}) - H(k_*-k_{c/d})\big) \qquad (6\text{-}2)$$

$$- \sum_{(a,b,c,d)\in B_n^{c=0}} \big(H(k-k_{-d/c}) - H(k_*-k_{-d/c})\big) \qquad (6\text{-}3)$$

$$+ \sum_{(a,b,c,d)\in B_n^{b>0}} \big(H(k-k_{c/d}) - H(k-k_{c/d-\lceil a/b\rceil}) - H(k_*-k_{c/d}) + H(k_*-k_{c/d-\lceil a/b\rceil})\big). \qquad (6\text{-}4)$$

We evaluate the three sums in the last expression separately. First note that the second sum (6-3) is zero, since $k_{-d/c} = p$ for $c = 0$ and $H(k-p) = H(k_*-p) = 0$ for all relevant k. We now look at the first sum (6-2), which is the following expression minus the same expression with k replaced by k_*:

$$\sum_{(a,b,c,d)\in B_n^{b=0}} H(k-k_{c/d}) = \sum_{d|n} \sum_{c=0}^{d-1} H(k-k_{c/d}) = \sum_{d|n} \sum_{c=1}^{d-1} H(k-k_{c/d}).$$

We set

$$s(k) = \#\{(c,d) \in \mathbb{Z}^2 : d \mid n,\ 0 < c < d,\ c \equiv dk \bmod p\}; \qquad (6\text{-}5)$$

then the sum above is

$$\sum_{d|n} \#\{c \in \mathbb{Z} : 0 < c < d,\ k_{c/d} \leq k\} - \tfrac{1}{2}s(k).$$

Now $dk_{c/d} = up + c$, where $1 \leq u < d$ satisfies $up \equiv -c \bmod d$, and so $k_{c/d} = \lceil up/d \rceil$. The map that sends u to c is a permutation of $\{1, \ldots, d-1\}$, which implies that

$$\#\{c \in \mathbb{Z} : 0 < c < d,\ k_{c/d} \leq k\} = \#\{u \in \mathbb{Z} : 0 < u < d,\ up \leq dk\} = \left\lfloor \frac{dk}{p} \right\rfloor.$$

This gives the expression

$$\sum_{m|n}\Big(\Big\lfloor \frac{mk}{p}\Big\rfloor - \Big\lfloor \frac{mk_*}{p}\Big\rfloor\Big) - \frac{1}{2}(s(k)-s(k_*)) \tag{6-6}$$

for the sum in (6-2).

Now we look at the third sum (6-4). Let $x = c/d$ for some $(a,b,c,d) \in B_n^{b>0}$; then $p > n \geq d > 0$, so $p \nmid d$. If $\mathcal{A}$ is a statement, we set $[\mathcal{A}] = 0$ if $\mathcal{A}$ is false and $[\mathcal{A}] = 1$ if $\mathcal{A}$ is true. Then, by checking the various cases and using that $k_{x-1} = k_x - 1$ when $k_x \neq 1$, we find that

$$H(k-k_x) - H(k-k_{x-1}) = [k_x = 1] - \tfrac{1}{2}[k = k_x] - \tfrac{1}{2}[k = k_{x-1}].$$

This implies that

$$H(k-k_x) - H(k-k_{x-1}) - H(k_*-k_x) + H(k_*-k_{x-1}) = \tfrac{1}{2}[k_* \in \{k_x, k_{x-1}\}] - \tfrac{1}{2}[k \in \{k_x, k_{x-1}\}].$$

We obtain the following expression for (6-4):

$$\begin{aligned}
&\frac{1}{2}\sum_{(a,b,c,d)\in B_n^{b>0}} \sum_{j=0}^{\lceil a/b\rceil - 1} \big([k_* \in \{k_{c/d-j}, k_{c/d-j-1}\}] - [k \in \{k_{c/d-j}, k_{c/d-j-1}\}]\big) \\
&\quad = \frac{1}{2}\big(\#\{(a,b,c,d) \in B_n^{b>0} : k_{c/d} = k_*\} - \#\{(a,b,c,d) \in B_n^{b>0} : k_{c/d} = k\}\big) && (6\text{-}7)\\
&\qquad + \frac{1}{2}\big(\#\{(a,b,c,d) \in B_n^{b>0} : k_{c/d-\lceil a/b\rceil} = k_*\} - \#\{(a,b,c,d) \in B_n^{b>0} : k_{c/d-\lceil a/b\rceil} = k\}\big) && (6\text{-}8)\\
&\qquad\quad + \#\{(a,b,c,d,j) \in U_n : k_* = k_{c/d-j}\} - \#\{(a,b,c,d,j) \in U_n : k = k_{c/d-j}\}, && (6\text{-}9)
\end{aligned}$$

where we have set

$$U_n = \Big\{(a,b,c,d,j) : (a,b,c,d) \in B_n^{b>0},\ 1 \leq j < \Big\lceil \frac{a}{b}\Big\rceil\Big\}.$$

Now we observe that there is a bijection

$$\psi_n : U_n \to \{(a,b,c,d) \in \mathbb{Z}_{>0}^4 : ad+bc = n\}, \quad (a,b,c,d,j) \mapsto (b, a-jb, d, -c+jd)$$

(its inverse maps (a,b,c,d) to $(b+ja, a, -d+jc, c, j)$ with $j = \lceil d/c\rceil$). Writing $\psi_n(a,b,c,d,j) = (a',b',c',d')$, we see that $k = k_{c/d-j}$ is equivalent to $k = k_{-d'/c'}$, which is the same as saying that $k_* = k_{c'/d'}$, or that $c' \equiv k_* d' \bmod p$. This shows that the terms in line (6-9) above are equal to

$$v_{p,n}(k) - v_{p,n}(k_*).$$

Using the bijection ϕ_n between $B_n^{b>0}$ and $B_n^{c>0}$, we see that the terms in line (6-8) can be written as

$$\begin{aligned}
&\tfrac{1}{2}\big(\#\{(a,b,c,d) \in B_n^{c>0} : k_{-d/c} = k_*\} - \#\{(a,b,c,d) \in B_n^{c>0} : k_{-d/c} = k\}\big)\\
&\qquad = \tfrac{1}{2}\big(\#\{(a,b,c,d) \in B_n^{c>0} : k_{c/d} = k\} - \#\{(a,b,c,d) \in B_n^{c>0} : k_{c/d} = k_*\}\big). \qquad (6\text{-}10)
\end{aligned}$$

This cancels the part of the terms in line (6-7) in which c is strictly positive, and the terms with $c = 0$ do not contribute anything. What remains is the part with $b = 0$ in (6-10), which is

$$\tfrac{1}{2}\big(\#\{(c,d) \in \mathbb{Z}^2_{>0} : d > c > 0,\ d \mid n,\ c \equiv dk \bmod p\} - \#\{(c,d) \in \mathbb{Z}^2_{>0} : d > c > 0,\ d \mid n,\ c \equiv dk_* \bmod p\}\big) = \tfrac{1}{2}(s(k) - s(k_*))$$

with $s(k)$ as in (6-5). This cancels the contribution coming from $s(k)$ and $s(k_*)$ in (6-6), and we obtain the desired result. □

Corollary 6.12. *Let $n \geq 1$ be an integer, let c and d be coprime integers such that $c > d > 0$, and let $p > nc$ be a prime. Let a and b be the integers satisfying $0 \leq a < c$, $0 \leq b < d$, and $ad - bc = 1$. Let $k, k_* \in \{1, \dots, p-1\}$ be such that $c \equiv dk \bmod p$ and $-d \equiv ck_* \bmod p$. Further, let the integers u and u_* satisfy $dk = up + c$ and $ck_* = u_* p - d$.*

Then $0 \leq u < d$, $0 \leq u_* < c$, and

$$(T'_n - n)\boldsymbol{e} \bullet \xi(k) = \left\lfloor \frac{nu}{d} \right\rfloor - \left\lfloor \frac{nb}{d} \right\rfloor + \left\lfloor \frac{na}{c} \right\rfloor - \left\lfloor \frac{nu_*}{c} \right\rfloor.$$

Proof. Since $dk - c > -p$ and $dk - c < dp$, we see that $0 \leq u < d$. Since $ck_* + d > 0$ and $ck_* + d < c(k_* + 1) \leq cp$, we also see that $0 \leq u_* < c$.

By Proposition 6.11,

$$(T'_n - n)\boldsymbol{e} \bullet \xi(k) = \left\lfloor \frac{nk}{p} \right\rfloor - \left\lfloor \frac{nk_*}{p} \right\rfloor + v'_{p,n}(k) - v'_{p,n}(k_*).$$

We evaluate each of the terms.

We have that $nk/p = nu/d + nc/(pd)$ and $p > nc$, so $0 < nc/(pd) < 1/d$, which implies that $\lfloor nk/p \rfloor = \lfloor nu/d \rfloor$.

Similarly, we have that $nk_*/p = nu_*/c - nd/(cp)$ and $p > nd$, so $0 < nd/(pc) < 1/c$, which implies that $\lfloor nk_*/p \rfloor = \lfloor (nu_* - 1)/c \rfloor$.

The third term counts the quadruples (a', b', c', d') of positive integers such that c' and d' are coprime, $a'd' + b'c' = n$, and $c' \equiv d'k \bmod p$. The latter implies that $c'd \equiv cd' \bmod p$. Since $0 < c'd < nd < p$ and $0 < cd' < cn < p$, we must have equality; then the coprimality of c' and d' and of c and d forces $(c', d') = (c, d)$. We have that $nad - nbc = n = a'd' + b'c' = a'd + b'c$, which implies that there is some $t \in \mathbb{Z}$ such that $na - a' = tc$ and $nb + b' = td$. The conditions $a', b' > 0$ then translate into $t < na/c$ and $t > nb/d$. Since $a/c > b/d$, this gives

$$v'_{p,n}(k) = \#\left\{t \in \mathbb{Z} : \frac{nb}{d} < t < \frac{na}{c}\right\} = \left\lfloor \frac{na-1}{c} \right\rfloor - \left\lfloor \frac{nb}{d} \right\rfloor.$$

The fourth term similarly counts quadruples (a', b', c', d') of positive integers such that c' and d' are coprime, $a'd' + b'c' = n$, and $p \mid c'k + d'$. The latter implies that $p \mid cc' + dd'$. But $0 < cc' + dd' < c(c' + d') \leq cn < p$, so there are no such quadruples, and the fourth term is zero.

Finally, note that

$$\left\lfloor \frac{na-1}{c} \right\rfloor - \left\lfloor \frac{nu_*-1}{c} \right\rfloor = \left\lfloor \frac{na}{c} \right\rfloor - \left\lfloor \frac{nu_*}{c} \right\rfloor,$$

as can be seen by considering the cases $c \mid n$ and $c \nmid n$ separately, taking into account that c is coprime with a and u_*. □

Corollary 6.13. *Let $M \geq 2$ be an integer, let $1 \leq a < M$ be coprime with M, let $n \geq 1$ be an integer, and let $p > nM$ be a prime. We let w denote the integer such that $1 \leq w < M$ and $apw \equiv 1 \bmod M$. Then*

$$(T'_n - n)\boldsymbol{e} \bullet \left\{0, \frac{a}{M}\right\} = \left\lfloor \frac{na}{M} \right\rfloor - \left\lfloor \frac{nw}{M} \right\rfloor.$$

Proof. We prove this by induction on M. If $M = 2$, then $a = 1$. We show the claim more generally for $a = 1$ and $M \geq 2$ arbitrary. We then have $\{0, a/M\} = \xi(M)$. The claim follows by taking $(c, d) = (M, 1)$ (then $(a, b) = (1, 0)$ and $(u, u_*) = (0, w)$) in Corollary 6.12.

Now assume that $M > 2$ and that the claim holds for smaller M. We can then find integers b and d such that $ad - bM = 1$ and $1 \leq d < M$. Then $0 \leq b < d$. If $d = 1$, then $b = 0$ and therefore $a = 1$; this case was already dealt with above. So we can assume that $d \geq 2$.

Let $1 \leq k < p$ be such that $M \equiv dk \bmod p$. Then

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

The matrix is in $\Gamma_0(p)$, so $\{b/d, a/M\} = \xi(k)$, and hence

$$(T'_n - n)\boldsymbol{e} \bullet \left\{0, \frac{a}{M}\right\} = (T'_n - n)\boldsymbol{e} \bullet \left\{0, \frac{b}{d}\right\} + (T'_n - n)\boldsymbol{e} \bullet \xi(k).$$

We use the induction hypothesis for the first term in the sum and Corollary 6.12 for the second term, where we take $(a, b, c, d) \leftarrow (a, b, M, d)$. Then

$$bpu = bdk - bc \equiv 1 \bmod d,$$

so u corresponds to w in the induction hypothesis, and $u_* = w$. This gives

$$\begin{aligned}(T'_n - n)\boldsymbol{e} \bullet \left\{0, \frac{a}{M}\right\} &= \left(\left\lfloor \frac{nb}{d} \right\rfloor - \left\lfloor \frac{nu}{d} \right\rfloor\right) + \left(\left\lfloor \frac{nu}{d} \right\rfloor - \left\lfloor \frac{nb}{d} \right\rfloor + \left\lfloor \frac{na}{M} \right\rfloor - \left\lfloor \frac{nw}{M} \right\rfloor\right) \\ &= \left\lfloor \frac{na}{M} \right\rfloor - \left\lfloor \frac{nw}{M} \right\rfloor.\end{aligned}$$

□

Proof of Lemma 6.7. Using that $L_n = T'_{2n} - 2T'_n = (T'_{2n} - 2n) - 2(T'_n - n)$, Corollary 6.13 gives (note that w does not depend on n)

$$L_n \boldsymbol{e} \bullet \left\{0, \frac{a}{M}\right\} = \left\lfloor \frac{2na}{M} \right\rfloor - \left\lfloor \frac{2nw}{M} \right\rfloor - 2\left\lfloor \frac{na}{M} \right\rfloor + 2\left\lfloor \frac{nw}{M} \right\rfloor = \varepsilon_M(na) - \varepsilon_M(nw),$$

and we can replace w with u/a, where u is as in Lemma 6.7. □

7. A criterion for ruling out moderately large primes

To exclude some of the larger primes for $d = 7$, we make use of the following criterion, which is due to the first author of this paper.

Proposition 7.1 (Derickx). *Let $d \geq 1$ and let p be a prime. We assume that either*

(i) *$J_1(p)(\mathbb{Q})$ is finite, or*

(ii) *there is $a \in (\mathbb{Z}/p\mathbb{Z})^\times/\{\pm 1\}$ such that $\mathrm{ord}(a) > 3d$ and $A = (\langle a\rangle - 1)(J_1(p)(\mathbb{Q}))$ is finite.*

In case (ii), *we say that "$(*)$ holds" when $\#A$ is odd or, more generally, the 2-primary part of A is contained in the subgroup of $J_1(p)(\mathbb{Q})$ generated by differences of rational cusps. We then set $n = 3$ in case* (i) *and*

$$n = \begin{cases} 5 & \text{if } (*) \text{ holds and } a \in \{2, 2^{-1}\}, \\ 6 & \text{if } (*) \text{ holds and } a \notin \{2, 2^{-1}\}, \\ 7 & \text{if } (*) \text{ does not hold and } a \in \{3, 3^{-1}\}, \\ 8 & \text{if } (*) \text{ does not hold and } a \notin \{3, 3^{-1}\} \end{cases}$$

in case (ii). *Then $nd < \mathrm{gon}_{\mathbb{Q}}(X_1(p))$ implies that $p \notin S(d)$. This holds in particular when*

$$d < \frac{325}{2^{16}} \frac{p^2 - 1}{n}.$$

Proof. If $c \in X_1(p)$ is a rational cusp, which we consider as an effective divisor of degree 1, and q is any prime, then $(T_q - \langle q\rangle - q)(c) = 0$. This can be deduced from the modular interpretation of the cusps. (See also [Parent 2000, end of Section 2.4] and note that the rational cusps are those mapping to the cusp ∞ on $X_0(p)$.)

We first consider case (i). Then, by Corollary 3.3, $J_1(p)(\mathbb{Q})$ is generated by differences of rational cusps. By the preceding paragraph, $T_q - \langle q\rangle - q$ kills $J_1(p)(\mathbb{Q})$ for all primes q (including $q = 2$; this improves Proposition 2.3 in this case). In case (ii), $T_q - \langle q\rangle - q$ kills the 2-primary part of A when this is contained in the subgroup generated by differences of rational cusps and kills the odd part of A by Proposition 2.3. So when $(*)$ holds, $T_q - \langle q\rangle - q$ kills A for arbitrary primes q. When $(*)$ does not hold, the statement is true for $q \geq 3$.

We let $x \in X_1(p)^{(d)}(\mathbb{Q})$, considered as an effective divisor of degree d on $X_1(p)$, and fix a rational cusp $c \in X_1(p)$. Then the linear equivalence class $[x - d \cdot c]$ of the divisor $x - d \cdot c$ is a rational point on $J_1(p)$.

Going back to the case (i), set $t = T_2 - \langle 2\rangle - 2$. Then

$$t(x - d \cdot c) = t(x) - dt(c) = t(x)$$

is a principal divisor, since $t([x - d \cdot c]) = 0$. This implies that the divisors $T_2(x)$ and $\langle 2\rangle(x) + 2x$ of degree $3d = nd$ are linearly equivalent. But $\mathrm{gon}_{\mathbb{Q}}(X_1(p)) > nd$ by assumption, so the divisors must in fact be equal, and $t(x) = 0$. Now Proposition 2.4 shows that x is a sum of cusps. This implies $p \notin S(d)$.

In case (ii), we set $q = 2$ when $(*)$ holds and otherwise $q = 3$, so that $T_q - \langle q\rangle - q$ kills A. Then $t(J_1(p)(\mathbb{Q})) = \{0\}$, where

$$t = (\langle a\rangle - 1)(T_q - \langle q\rangle - q) = (\langle a\rangle T_q + \langle q\rangle + q) - (T_q + \langle qa\rangle + q\langle a\rangle). \tag{7-1}$$

If $qa = 1$, this simplifies to

$$t = (\langle a\rangle T_q + \langle q\rangle + (q-1)) - (T_q + q\langle a\rangle), \tag{7-2}$$

and if $a = q$, we obtain

$$t = (\langle a\rangle T_q + q) - (T_q + \langle qa\rangle + (q-1)\langle a\rangle). \tag{7-3}$$

We write t_1 for the first term and t_2 for the second in the difference (7-1), (7-2) or (7-3). Since the diamond operators are automorphisms of $X_1(p)$ and T_q multiplies degrees by $q+1$, we see that applying t_1 or t_2, considered as a correspondence on $X_1(p)$, to an effective divisor of degree d results in an effective divisor of degree nd.

As before, $t(x - d\cdot c) = t(x) - dt(c) = t(x)$ is a principal divisor, and from $\mathrm{gon}_{\mathbb{Q}}(X_1(p)) > nd$, we conclude that

$$t(x) = (T_q - \langle q\rangle - q)(\langle a\rangle - 1)(x) = 0.$$

By Proposition 2.4 again, this implies that $\langle a\rangle(x) - x$ is supported on cusps. Since the diamond operators permute the cusps among themselves, this then implies that $x = x_0 + x_1$, where x_0 is supported in cusps and x_1 does not have cusps in its support and satisfies $\langle a\rangle(x_1) = x_1$. Now the diamond operators act freely on the noncuspidal points of $X_1(p)$ with the exception of points corresponding to elliptic curves with j-invariant 0 or 1728, which can have stabilizers of orders 3 and 2, respectively. The condition $\langle a\rangle(x_1) = x_1$ implies that x_1 is a sum of (sums over) orbits of $\langle a\rangle$, which have length at least $\mathrm{ord}(a)/3$. Since $\mathrm{ord}(a) > 3d$ by assumption, this forces $x_1 = 0$, and we conclude that x is supported in cusps. This again implies that $p \notin S(d)$.

For the last statement, note that

$$\mathrm{gon}_{\mathbb{Q}}(X_1(p)) \geq \mathrm{gon}_{\mathbb{C}}(X_1(p)) \geq \tfrac{1}{48}\lambda_1(p^2-1)$$

by [Abramovich 1996] (using that $\Gamma_1(p)$ has index $(p^2-1)/2$ in $\mathrm{PSL}(2,\mathbb{Z})$), where λ_1 is the smallest positive eigenvalue of the Laplace operator on $X_1(p)(\mathbb{C})$, which satisfies $\lambda_1 \geq \frac{975}{4096}$ by [Kim 2003]. □

Remark. Without the condition $\mathrm{ord}(a) > 3d$ in the case that $J_1(p)(\mathbb{Q})$ has positive rank, the proof shows that any rational point on $X_1(p)^{(d)}$ whose support consists of noncuspidal points must be a sum of orbits of $\langle a\rangle$. This is impossible when d cannot be written as a sum of the possible orbit lengths ($\mathrm{ord}(a)$, together with $\mathrm{ord}(a)/2$ when $\mathrm{ord}(a)$ is even and $\mathrm{ord}(a)/3$ when $\mathrm{ord}(a)$ is divisible by 3). But even when d can be written in this way, this gives strong restrictions. For example, when $\mathrm{ord}(a) = d$ and d is coprime to 6, then such a point must be obtained by pulling back a rational point on X_H, where H is generated by a.

We plan to explore this further in a follow-up paper.

Corollary 7.2. $p \notin S(7)$ *for* $p \in \{71, 113, 127\}$.

Proof. We check that for the two primes $p \in \{113, 127\}$, the positive-rank simple factors of $J_1(p)$ already occur in $J_0(p)$. We can thus take any $a \in (\mathbb{Z}/p\mathbb{Z})^\times/\{\pm 1\}$; we use $a = 3$, which generates $(\mathbb{Z}/p\mathbb{Z})^\times/\{\pm 1\}$

in both cases. In particular, $\operatorname{ord}(a) = (p-1)/2 > 3 \cdot 7$. We then have $n = 7$ in Proposition 7.1. Since

$$\frac{325}{2^{16}} \frac{p^2-1}{7} > 9,$$

all assumptions in Proposition 7.1 are satisfied.

To deal with $p = 71$, we recall that by Proposition 3.1, 71 is a rank-zero prime, so we can apply Proposition 7.1 with $n = 3$. Since

$$\frac{325}{2^{16}} \frac{71^2-1}{3} > 8,$$

the claim follows also in this case. □

Remark. For $p = 73$, the best we can do is use $a = 2$ and $n = 5$ in Proposition 7.1 (by [Conrad et al. 2003, Section 6.2], the torsion subgroup of $J_1(73)(\mathbb{Q})$ is generated by differences of rational cusps). However, the gonality lower bound works only for $d \le 5$. We would need $\operatorname{gon}_{\mathbb{Q}}(X_1(73)) > 35$. From Table 1 in [Derickx and van Hoeij 2014], it appears that this is very likely the case, but it is also very likely hard to prove. (Note that $\operatorname{ord}(a) = 9 \le 3d$, but the argument would still work; see the remark following Proposition 7.1.)

8. Verification of assumption (b) of Lemma 1.7

We now discuss assumption (b) of Lemma 1.7 for the remaining pairs of degrees d and primes p. Recall that the assumption is always satisfied (with $\ell = 2$) when $p > (2^{d/2}+1)^2$; see Lemma 1.9. The following table tells us which primes we still have to consider for each d.

d	3	4	5	6	7
$\lfloor (2^{d/2}+1)^2 \rfloor$	14	25	44	81	151

In some cases, we can show that all points in $X_1(p)^{(d)}(\mathbb{F}_2)$ are sums of images of rational cusps, even when p is below this bound. The result of [Waterhouse 1969, Theorem 4.1] tells us precisely what the possible orders of $E(\mathbb{F}_{2^d})$ are for elliptic curves E defined over $\mathbb{F}_{2^d}$. Using this (or a brute-force enumeration of all such curves up to isomorphism), we obtain the following extension of Lemma 1.9.

Lemma 8.1. *The set $X_1(p)^{(d)}(\mathbb{F}_2)$ consists of sums of images of rational cusps for the following pairs of an integer $3 \le d \le 7$ and a prime p:*

$$\begin{aligned}
&d = 3 \quad \text{and} \quad p = 11 \quad \text{or} \quad p \ge 17,\\
&d = 4 \quad \text{and} \quad p \ge 19,\\
&d = 5 \quad \text{and} \quad p \ge 23 \quad \text{and} \quad p \notin \{31, 41\},\\
&d = 6 \quad \text{and} \quad p = 23 \quad \text{or} \quad (p \ge 43 \quad \text{and} \quad p \ne 73),\\
&d = 7 \quad \text{and} \quad p \in \{47, 53\} \quad \text{or} \quad (p \ge 79 \quad \text{and} \quad p \notin \{113, 127\}).
\end{aligned}$$

Proof. According to [Waterhouse 1969, Theorem 4.1], $\#E(\mathbb{F}_{2^d})$ can take all even values in the Hasse interval $\left[\lceil (2^{d/2}-1)^2 \rceil, \lfloor (2^{d/2}+1)^2 \rfloor\right]$ and, in addition, the values

$$\begin{aligned} 2^d + m2^{d/2} + 1 \quad &\text{for } m \in \{-2,-1,0,1,2\} \qquad &&\text{if } d \text{ is even;} \\ 2^d + m2^{(d+1)/2} + 1 \quad &\text{for } m \in \{-1,0,1\} &&\text{if } d \text{ is odd.} \end{aligned}$$

This allows us to determine the set of primes p such that there are no noncuspidal points of degree $\le d$ on $X_1(p)_{\mathbb{F}_2}$. The condition that there are no cusps of degree $\le d$ that are not images of rational cusps excludes in addition $p = 31$ for $d \ge 5$ and $p = 127$ for $d \ge 7$. □

We note that for the primes not in the list above for a given d, there are indeed points $\bar{x}$ as in assumption (b). If we want to show that $p \notin S(d)$ for one of these primes, we have to do some work to show that there are no rational points in the corresponding residue classes. For $p \in \{29, 31, 41\}$ and $d \ge 5$, we already did this in Lemma 3.7. Taking into account Corollary 7.2, this leaves the primes $p \in \{37, 43, 59, 61, 67\}$ for $d = 7$ and $p = 73$ for $d = 6, 7$.

We can deal with $(d, p) \in \{(6, 73), (7, 43)\}$ in the following way.

Lemma 8.2. *Let $d \ge 1$ be an integer and let p be a prime. Let $\bar{x} \in X_1(p)^{(d)}(\mathbb{F}_2)$ be a point that is not a sum of images of rational cusps. Let $H \subseteq (\mathbb{Z}/p\mathbb{Z})^\times/\{\pm 1\}$ be a subgroup and denote the image of $\bar{x}$ in $X_H^{(d)}(\mathbb{F}_2)$ by $\bar{x}_H$. Assume that the following conditions are satisfied*:

(1) *There is $t : J_{H,\mathbb{Z}_{(2)}} \to A_{\mathbb{Z}_{(2)}}$ such that $t(J_H(\mathbb{Q}))$ is finite of odd order and $t \circ \iota$ (with $\iota : X_H^{(d)} \to J_H$) is a formal immersion at $\bar{x}_H$.*

(2) *There is a rational point $x_H \in X_H^{(d)}(\mathbb{Q})$ such that* $\mathrm{red}_2(x_H) = \bar{x}_H$.

Let $x \in X_1(p)^{(d)}(\mathbb{Q})$ be such that $\mathrm{red}_2(x) = \bar{x}$. *Then x maps to x_H under the canonical map $X_1(p)^{(d)} \to X_H^{(d)}$.*

Proof. Let x'_H be the image of x in $X_H^{(d)}(\mathbb{Q})$; then $\mathrm{red}_2(x'_H) = \bar{x}_H = \mathrm{red}_2(x_H)$. Since $t(J_H(\mathbb{Q}))$ is finite of odd order, this implies that $t(\iota(x'_H) - \iota(x_H)) = 0$. Since $t \circ \iota$ is a formal immersion at $\bar{x}_H$, it follows that $x'_H = x_H$. □

If, in the situation of Lemma 8.2, x_H does not lift to a rational point on $X_1(p)^{(d)}$, then it follows that no rational point on $X_1(p)^{(d)}$ can reduce mod 2 to $\bar{x}$. We have to carry this out for all $\bar{x}$ as in assumption (b). To do this, we formulate a criterion that allows us to verify the formal immersion condition in Lemma 8.2 also for points whose support does not consist of cusps.

Lemma 8.3. *Fix a prime ℓ and an integer $d \ge 1$. Let X be a curve over $\mathbb{Q}$ with good reduction at ℓ, with Jacobian variety J. Fix $b \in X(\mathbb{Q})$ and use it to define embeddings $\iota : X \to J$ and $\iota_d : X^{(d)} \to J$. Let A be another abelian variety (with good reduction at ℓ) such that there is a homomorphism $t : J \to A$. Let $L \subseteq H^0(X_{\mathbb{F}_\ell}, \Omega^1)$ be the pullback of $H^0(A_{\mathbb{F}_\ell}, \Omega^1)$ under $t \circ \iota$, and let $\varphi : X_{\mathbb{F}_\ell} \to \mathbb{P}\,\mathrm{Tan}_0(A_{\mathbb{F}_\ell}) \cong \mathbb{P}^{\dim A - 1}_{\mathbb{F}_\ell}$ be the morphism determined by the linear system corresponding to L. Let $\bar{x} \in X^{(d)}(\mathbb{F}_\ell)$ be a point that is the sum of d distinct geometric points $\bar{x}_1, \ldots, \bar{x}_d \in X(\overline{\mathbb{F}}_\ell)$. Assume that*

(i) *the differentials in L do not vanish simultaneously at any point $\bar{x}_j$, and that*

(ii) *the points $\varphi(\bar{x}_1), \dots, \varphi(\bar{x}_d) \in \mathbb{P}^{\dim A-1}(\overline{\mathbb{F}}_\ell)$ span a linear subspace of dimension $d-1$.*

Then $t \circ \iota_d$ is a formal immersion at $\bar{x}$.

Proof. To show that $t \circ \iota_d$ is a formal immersion, it is sufficient to show that the induced map on tangent spaces $\mathrm{Tan}_{\bar{x}}(X^{(d)}_{\overline{\mathbb{F}}_\ell}) \to \mathrm{Tan}_{t(\iota(\bar{x}))}(A_{\overline{\mathbb{F}}_\ell})$ is injective; see [Parent 1999, Theorem 4.18]. We can equivalently consider this condition over $\overline{\mathbb{F}}_\ell$.

Since the regular 1-forms on A are invariant under translation, we have a canonical identification of all tangent spaces $\mathrm{Tan}_{\bar{a}}(A_{\overline{\mathbb{F}}_\ell})$ with the tangent space at the origin, whose projectivization is the codomain of φ. Since the differentials in L do not vanish simultaneously at $\bar{x}_j$, the map φ sends a point $\bar{x}_j \in X(\overline{\mathbb{F}}_\ell)$ to the image in $\mathbb{P}\,\mathrm{Tan}_0(A_{\overline{\mathbb{F}}_\ell})$ of the tangent space $\mathrm{Tan}_{\bar{x}_j}(X_{\overline{\mathbb{F}}_\ell})$ under $(t \circ \iota)_*$ followed by a suitable translation.

Since the geometric points making up $\bar{x}$ are distinct, we have a canonical isomorphism

$$\mathrm{Tan}_{\bar{x}}(X^{(d)}_{\overline{\mathbb{F}}_\ell}) \cong \bigoplus_{j=1}^{d} \mathrm{Tan}_{\bar{x}_j}(X_{\overline{\mathbb{F}}_\ell}).$$

The image of $\mathrm{Tan}_{\bar{x}}(X^{(d)}_{\overline{\mathbb{F}}_\ell})$ in $\mathbb{P}\,\mathrm{Tan}_0(A_{\overline{\mathbb{F}}_\ell})$ under $(t \circ \iota_d)$ followed by a suitable translation is then the linear span of the various images $\varphi(\bar{x}_j)$; the map on tangent spaces is injective if and only if this span has the maximal possible dimension $d-1$. □

We will apply this as follows. We use the q-expansions mod 2 of the cusp forms associated to X_H to determine equations for the canonical model of $X_{H,\mathbb{F}_2}$. We then project away from the subspace where the forms in L vanish (in practice, we compute the image of φ in a similar way and then set up the projection) and check that none of the points $\bar{x}_j$ lie in this subspace. This verifies the nonvanishing condition (i). We then check condition (ii).

Lemma 8.4. *Let $x \in X_1(73)^{(6)}(\mathbb{Q})$. Then $\mathrm{red}_2(x) \in X_1(73)^{(6)}(\mathbb{F}_2)$ is a sum of images of rational cusps.*

Proof. There are, up to isomorphism, exactly two elliptic curves over $\mathbb{F}_{2^6}$ with a point of order 73. They have zero j-invariant (they must be supersingular according to [Waterhouse 1969]) and automorphism group $\mathbb{Z}/6\mathbb{Z}$, so each of them gives rise to $(73-1)/6 = 12$ $\mathbb{F}_{2^6}$-points on $X_1(73)_{\mathbb{F}_2}$. These 24 points split into four orbits of size six under the action of Frobenius (each orbit contains three points coming from each of the two curves), so we obtain exactly four noncuspidal points in $X_1(73)^{(6)}(\mathbb{F}_2)$. There are no cuspidal points that are not sums of images of rational cusps, since the other cusps on $X_1(73)_{\mathbb{F}_2}$ are minimally defined over $\mathbb{F}_{2^9}$. So we just have to exclude these four noncuspidal points.

Let H be the subgroup of $(\mathbb{Z}/73\mathbb{Z})^\times/\{\pm 1\}$ of index 9. The canonical map $X_1(73) \to X_H$ is of degree 4 and unramified at all 24 points mentioned above. This implies that they have six distinct images on X_H; one can check that these points form one Frobenius orbit, so we get one point $\bar{x}_H \in X_H^{(6)}(\mathbb{F}_2)$ that we have to deal with. The Jacobian J_H splits into a copy of $J_{H'}$, where $H \subseteq H'$ has index 3, and a simple 30-dimensional abelian variety A. One can check that A is a factor of the winding quotient and that all isogenous (over $\mathbb{Q}$) abelian varieties have torsion subgroup of odd order (by computing orders of $A(\mathbb{F}_q)$ for suitable primes q via the Hecke eigenvalues). We take $t = \langle 7 \rangle - 1$; this kills $J_{H'}$ and projects J_H

into A. Since the nonzero eigenvalues of t are invertible mod 2 (they are of the form $\omega - 1$ with $\omega \in \mu_3$), we can work with the q-expansions mod 2 of a basis of the space of cusp forms associated to A. We check, as described above, that $t \circ \iota_6$ is a formal immersion at $\bar{x}_H$. (In practice, we check this for all Frobenius orbits of length 6 in $X_H(\overline{\mathbb{F}}_2)$, since it is not so easy to determine which point is in the support of $\bar{x}_H$.)

Note that $X_H \to X_{H'} \to X_0(73)$ is the composition of two maps of degree 3, the second of which is étale (by Riemann–Hurwitz: $X_{H'}$ is of genus 13 and $X_0(73)$ has genus 5). Let E_0 be an elliptic curve over $\mathbb{Q}$ with complex multiplication by cube roots of unity. Then E_0 has two Galois-conjugate cyclic subgroups of order 73, with each subgroup defined over $K = \mathbb{Q}(\sqrt{-3})$ (note that 73 splits in K), so E_0 gives rise to a pair of Galois-conjugate points $y_1, y_2 \in X_0(73)(K)$. The preimages of these two points on $X_{H'}$ give six geometric points that are Galois conjugate; the map $X_H \to X_{H'}$ is totally ramified at each of them, so we find a Galois orbit of size 6 of points in $X_H(\overline{\mathbb{Q}})$, giving rise to a rational point $x_H \in X_H^{(6)}(\mathbb{Q})$. This point reduces mod 2 to $\bar{x}_H$ (as one can show by writing down an explicit twist of $E_{0,K}$ for a certain number field of degree 24 that has a K-rational point of order 73 and checking that the 24 geometric points corresponding to its Galois conjugates reduce to the 24 noncuspidal points in $X_1(73)(\mathbb{F}_{2^6})$ mentioned above), but does not lift to a rational point on $X_1(73)^{(6)}$, since there are no CM elliptic curves with a 73-torsion point over number fields of degree < 24; see [Clark et al. 2013, Table 1]. By Lemma 8.2 and the discussion following it, this finishes the proof. □

Lemma 8.5. *Let $x \in X_1(43)^{(7)}(\mathbb{Q})$. Then $\mathrm{red}_2(x) \in X_1(43)^{(7)}(\mathbb{F}_2)$ is a sum of images of rational cusps.*

Proof. There is, up to isomorphism, exactly one elliptic curve over $\mathbb{F}_{2^7}$ with a point of order 43. It is supersingular; its automorphism group over $\mathbb{F}_{2^7}$ has order 2, since $\mathbb{F}_{2^7}$ does not contain primitive cube roots of unity. It therefore gives rise to 21 noncuspidal points in $X_1(43)(\mathbb{F}_{2^7})$, making up three Galois orbits. The nonrational cusps are also defined over $\mathbb{F}_{2^7}$. We obtain six points in total in $X_1(43)^{(7)}(\mathbb{F}_2)$ that are not supported in rational cusps. Take H to be the subgroup of index 7. Then the six points above map to two points in $X_H^{(7)}(\mathbb{F}_2)$. For A, we use the winding quotient of J_H; one can show that each $\mathbb{Q}$-isogenous abelian variety has odd torsion order. We show as before that $t \circ \iota$ is a formal immersion at the two points in question.

On the other hand, there is a point in $X_H^{(7)}(\mathbb{Q})$ that corresponds to the pullback of the cusp 0 on $X_0(43)$ (note that $X_H \to X_0(43)$ has degree 7). It does not lift to a rational point on $X_1(43)^{(7)}$, since the nonrational cusps on $X_1(43)$ are points of degree 21. This shows that there are no rational points on $X_1(43)^{(7)}$ whose reduction is cuspidal, but that are not supported in rational cusps.

Consider now the rational point on $X_0(43)$ that corresponds to elliptic curves over $\mathbb{Q}$ with CM by the order of discriminant -43. Its pullback to X_H again provides us with a rational point on $X_H^{(7)}$, whose reduction must be the other point we have to consider, since such curves have (potentially) good reduction at 2. Again, this point does not lift to a rational point on $X_1(43)^{(7)}$, as can be verified by consulting [Clark et al. 2013, Table 1]. This shows that there are no rational points on $X_1(43)^{(7)}$ whose reduction is noncuspidal. □

We still have to show that $p \notin S(7)$ for

$$p = 37,\ 59,\ 61,\ 67,\ 73.$$

We use the following simple observation by the first author of this paper, together with the fact that it is actually possible to check this criterion by a computation.

Lemma 8.6 (Derickx). *Let $d \geq 1$ be an integer and let $p > 2$ be a prime. Assume that $t \in \mathbb{T}$ has the property that $t(J_1(p)(\mathbb{Q})) = \{0\}$, where we consider t as an endomorphism of $J_1(p)$. Let $\bar{x}_0, \bar{x} \in X_1(p)^{(d)}(\mathbb{F}_2)$ be such that $\bar{x}_0$ is a sum of images of rational cusps. If the divisor $t(\bar{x} - \bar{x}_0)$ on $X_1(p)_{\mathbb{F}_2}$ is not principal (where we now consider t as a correspondence on $X_1(p)_{\mathbb{F}_2}$), then there is no rational point on $X_1(p)^{(d)}$ whose reduction mod 2 is $\bar{x}$.*

Remark. This result remains valid with an odd positive integer N in place of p. (We need N to be odd so that $X_1(N)$ has good reduction mod 2.)

Proof. Let $x_0 \in X_1(p)^{(d)}(\mathbb{Q})$ be the sum of rational cusps such that $\mathrm{red}_2(x_0) = \bar{x}_0$ and assume that there is some $x \in X_1(p)^{(d)}(\mathbb{Q})$ such that $\mathrm{red}_2(x) = \bar{x}$. Then the divisor $x - x_0$ represents a point in $J_1(p)(\mathbb{Q})$; it follows that $t(x - x_0)$ represents zero and is therefore principal. Applying reduction mod 2 shows that $t(\bar{x} - \bar{x}_0)$ must be principal as well. □

We can find a suitable Hecke operator t by multiplying an operator that projects $J_1(p)$ into an abelian subvariety of Mordell–Weil rank zero (this is equivalent to this operator factoring through the winding quotient) with an operator that kills rational torsion. For the computations, we will use a model of $X_1(p)$ that is derived directly from the usual modular interpretation, i.e., noncuspidal points on $X_1(p)$ correspond to pairs (E, P), where E is an elliptic curve and $P \in E$ is a point of exact order p. The effect of a Hecke operator T_n with $p \nmid 2n$ as a correspondence on $X_1(p)_{\mathbb{F}_2}$ in this interpretation is then given by mapping (E, P) to the sum of the pairs $(E', \phi(P))$, where $\phi : E \to E'$ runs through the cyclic isogenies of degree n. This switch from the "natural" modular interpretation given in Section 2 has the effect that we have to conjugate everything by the Atkin–Lehner involution. Concretely, this means that instead of $T_q - \langle q \rangle - q$ as stated in Proposition 2.3, we have to use $T_q - q\langle q \rangle - 1$ with any odd prime q to kill the rational torsion. We will work with $q = 3$.

For the projection part of t, we will use an operator of the form $\langle a \rangle - 1$, so we take

$$t = (\langle a \rangle - 1)(T_3 - 3\langle 3 \rangle - 1).$$

(This is similar to the idea used in Proposition 7.1.) We use the modular interpretation of the points on $X_1(p)$ to find the image of the divisor $\bar{x} - \bar{x}_0$ under t. Sutherland has computed planar equations for $X_1(N)$ for all $N = p$ in the relevant range, together with explicit expressions relating the coordinates in these equations to the parameters b and c in the Tate form

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

of the associated elliptic curve with point $(0, 0)$ of order N. See [Sutherland 2012]; the equations are available https://math.mit.edu/~drew/X1_altcurves.html.

We find the action of a diamond operator $\langle a \rangle$ on a point on $X_1(p)$ by multiplying the point $P = (0, 0)$ on the associated curve $E_{b,c}$ by a and then bringing the pair $(E_{b,c}, aP)$ into Tate form $(E_{b',c'}, (0, 0))$. To

get the effect of the Hecke operator T_3, we use the description of T_n given above, i.e., we find the four elliptic curves that are 3-isogenous to $E_{b,c}$ (they may be defined over an extension of the base field we are considering) and find the points corresponding to the isogenous curves together with the image of P. The sum of these four points is then the image of the original point (considered as a divisor of degree 1) under T_3.

Lemma 8.7. *Let $p \in \{59, 61, 67, 73\}$ and $x \in X_1(p)^{(7)}(\mathbb{Q})$. Then $\mathrm{red}_2(x) \in X_1(p)^{(7)}(\mathbb{F}_2)$ is a sum of images of rational cusps.*

Proof. We determine a suitable a for each of the primes p such that $\langle a\rangle - 1$ projects $J_1(p)$ into an abelian subvariety of rank zero. For $p \in \{59, 67, 73\}$, the only simple components of $J_1(p)$ that have positive rank are also components of $J_0(p)$, so we can take a to be any element of $(\mathbb{Z}/p\mathbb{Z})^\times/\{\pm 1\}$. For $p = 61$, there is a component of positive rank in J_H for the subgroup H of index 6 that does not occur in $J_0(p)$, and all components of positive rank occur in J_H, so we take $a = 3 \equiv 2^6 \bmod 61$, where 2 is a primitive root mod 61. We note that $\langle a\rangle - 1$ maps x_0 to a degree-zero divisor representing a torsion point in $J_1(p)(\mathbb{Q})$, so we just have to compute $t(\bar{x})$ and check whether this divisor is principal, where $\bar{x}$ and x_0 are as in Lemma 8.6.

We then find all the noncuspidal places of degree at most 7 on $X_1(p)_{\mathbb{F}_2}$. For the computation, it is sufficient to consider one representative in each orbit under the diamond operators. For $p < 73$, we find no such places of degree ≤ 6 and either one (for $p = 61, 67$) or two (for $p = 59$) orbits of places of degree 7. For $p = 73$, there are two orbits of places of degree 6 and one orbit of places of degree 7.

For the representatives $\bar{x}$ of orbits of places of degree 7 (which we identify with effective divisors of degree 7), we compute the divisor $t(\bar{x})$ and verify that it is not principal. This can be done by computing the Riemann–Roch space associated to the divisor; a divisor of degree zero is principal if and only if its Riemann–Roch space is nontrivial. (Magma has a built-in function for testing whether a divisor is principal.)

The places of degree 6 on $X_1(73)_{\mathbb{F}_2}$ give rise to effective divisors of degree 7 by adding one of the images of the rational cusps (which are exactly the $\mathbb{F}_2$-points on $X_1(73)$). Applying t to such a sum differs from the result of applying t to the degree 6 divisor coming from the place by a principal divisor, since the rational cusps map to principal divisors. So we only have to check that $t(\bar{x})$ is nonprincipal for the two representatives of orbits of places of degree 6. (We note that this also gives an alternative proof of Lemma 8.4.)

Finally, we note that all other points in $X_1(p)^{(7)}(\mathbb{F}_2)$ are supported in images of rational cusps, since the other cusps give rise to points of degree at least 9 over $\mathbb{F}_2$.

The computations took less than one hour each for $p = 59$ and 61, about three hours for $p = 67$ and about seven hours for $p = 73$. □

Remark. We can use this approach also to show that there are no noncuspidal points in $X_1(43)^{(7)}(\mathbb{F}_2)$ that arise as the reduction modulo 2 of a rational point. We would still have to deal with the points arising from Frobenius orbits of cusps that are not images of rational cusps, however; see the proof of Lemma 8.5.

Now Proposition 1.10 follows from Lemmas 3.7, 8.1, 8.4, 8.5, and 8.7 and Corollary 7.2.

Finally, we deal with $p = 37$.

Lemma 8.8. *Modulo the action of Frobenius and the diamond operators, there is exactly one point of degree* 6 *on* $X_1(37)_{\mathbb{F}_2}$ *such that the corresponding point* $\bar{x} \in X_1(37)^{(6)}(\mathbb{F}_2)$ *is the reduction mod* 2 *of a rational point* $x \in X_1(37)^{(6)}(\mathbb{Q})$, *and this point* x *is uniquely determined by* $\bar{x}$.

Proof. We proceed as in the proof of Lemma 8.7. The only positive-rank factor of $J_1(37)$ occurs in $J_0(37)$ (it is the "first" elliptic curve of rank 1), so we can take any a for the criterion of Lemma 8.6. The computation shows that of the two diamond orbits of places of degree 6, only one satisfies the criterion in Lemma 8.6. (It should be noted that this can be used to verify that we are correct in working with the Hecke operator $T_3 - 3\langle 3\rangle - 1$: none of the two places satisfies the criterion when using $T_3 - \langle 3\rangle - 3$ instead, but one of them has to, since there are noncuspidal rational points on $X_1(37)^{(6)}$.)

We know that there is a diamond orbit of rational points that has to reduce to our unique diamond orbit that lifts. To show that the lift is unique, we use Lemma 8.3. The Hecke operator T_{17} projects $J_1(37)$ into an abelian subvariety of rank zero. Its eigenvalues are invertible mod 2 on newforms corresponding to a subvariety of dimension 36, which has odd-order rational torsion subgroup. We then verify the formal immersion criterion (for all points of degree 6, since we work with a different model here and did not try to find an explicit birational map between the two models). □

Proof of Proposition 1.4. Let $x \in X_1(37)^{(6)}(\mathbb{Q})$ be a point whose support contains no cusps. Since (a) holds for $(d, p) = (6, 37)$ by Proposition 1.8 and there are no noncuspidal points on $X_1(37)_{\mathbb{F}_2}$ of degree ≤ 5, it follows that $\bar{x} = \mathrm{red}_2(x) \in X_1(37)^{(6)}(\mathbb{F}_2)$ is also a point whose support contains no cusps. By Lemma 8.8, $\bar{x}$ is uniquely determined up to the action of the diamond operators, and there is no other point than x that reduces mod 2 to $\bar{x}$. On the other hand, we know a point x' with this property; this is a point coming from the curve $E_{6,37}$ with some choice of point of order 37 (they are all in the same diamond orbit). It follows that $x = x'$, which implies the claim. □

We finish off the determination of $S(7)$ by excluding $p = 37$.

Lemma 8.9. $37 \notin S(7)$.

Proof. As in the proof of Lemma 8.8, we show that there is no point of degree 7 on $X_1(37)_{\mathbb{F}_2}$ such that the corresponding point in $X_1(37)^{(7)}(\mathbb{F}_2)$ is the reduction of a rational point. Now assume that $x \in X_1(37)^{(7)}(\mathbb{Q})$ and consider $\bar{x} = \mathrm{red}_2(x)$. By the preceding statement, the support of $\bar{x}$ must contain a cusp, and the noncuspidal part of $\bar{x}$ must satisfy the criterion of Lemma 8.6. By Lemma 8.8 and its proof, the noncuspidal part is then either empty or in the unique diamond orbit coming from noncuspidal rational points on $X_1(37)^{(6)}$. In the first case, x must be a sum of rational cusps, since assumption (a) holds. To deal with the second case, we verify the formal immersion criterion as in the proof of Lemma 8.8, but now for all sums of an $\mathbb{F}_2$-rational cusp and a prime divisor of degree 6. This shows that the criterion is satisfied; therefore the point x is unique in its residue class mod 2. On the other hand, there is a known point in this residue class, which comes from adding the rational cusp that lifts the unique cusp in the support of $\bar{x}$ to the degree 6 divisor lifting the remaining part (this is one of the

sporadic points in $X_1(37)^{(6)}(\mathbb{Q})$). It follows that x is this point; in particular, x has a cusp in its support. So we conclude that every rational point on $X_1(37)^{(7)}$ has a cusp in its support; this is equivalent to the statement that $37 \notin S(7)$. □

Acknowledgments

We would like to thank Bas Edixhoven, Barry Mazur, and Loïc Merel for their many valuable comments and suggestions, Pierre Parent for his idea to look at CM elliptic curves for the proof of $73 \notin S(6)$ (Lemma 8.4), Filip Najman for some helpful information on sporadic torsion points, and Tessa Schild for her proofreading of an earlier version of this paper. We thank Joseph Oesterlé for kindly allowing us to use his notes [1994] and helping the first author understand the proof of the bound (1-1). We also thank the referee of an earlier version of this paper for some valuable feedback.

References

[Abramovich 1996] D. Abramovich, "A linear lower bound on the gonality of modular curves", *Int. Math. Res. Not.* **1996**:20 (1996), 1005–1011. MR Zbl

[Bosma et al. 1997] W. Bosma, J. Cannon, and C. Playoust, "The Magma algebra system, I: The user language", *J. Symbolic Comput.* **24**:3-4 (1997), 235–265. MR Zbl

[Clark et al. 2013] P. L. Clark, B. Cook, and J. Stankewicz, "Torsion points on elliptic curves with complex multiplication", *Int. J. Number Theory* **9**:2 (2013), 447–479. MR Zbl

[Conrad et al. 2003] B. Conrad, B. Edixhoven, and W. Stein, "$J_1(p)$ has connected fibers", *Doc. Math.* **8** (2003), 331–408. MR Zbl

[Deligne and Rapoport 1973] P. Deligne and M. Rapoport, "Les schémas de modules de courbes elliptiques", pp. 143–316 in *Modular functions of one variable*, *II* (Antwerp, Belgium, 1972), edited by P. Deligne and W. Kuyk, Lecture Notes in Math. **349**, Springer, 1973. MR Zbl

[Derickx 2016] M. Derickx, *Torsion points on elliptic curves over number fields of small degree*, Ph.D. thesis, Universiteit Leiden, 2016, available at http://hdl.handle.net/1887/43186.

[Derickx 2020] M. Derickx, SageMath code for the verification of assumption (a), 2020, available at https://github.com/koffie/mdsage/blob/master/mdsage/kamiennys_criterion.py.

[Derickx and van Hoeij 2014] M. Derickx and M. van Hoeij, "Gonality of the modular curve $X_1(N)$", *J. Algebra* **417** (2014), 52–71. MR Zbl

[Derickx et al. 2017] M. Derickx, S. Kamienny, W. Stein, and M. Stoll, "Torsion points on elliptic curves over number fields of small degree", preprint, 2017. arXiv 1707.00364

[Diamond and Im 1995] F. Diamond and J. Im, "Modular forms and modular curves", pp. 39–133 in *Seminar on Fermat's last theorem* (Toronto, 1993–1994), edited by V. K. Murty, CMS Conf. Proc. **17**, Amer. Math. Soc., Providence, RI, 1995. MR Zbl

[Drinfeld 1973] V. G. Drinfeld, "Two theorems on modular curves", *Funktsional. Anal. i Prilozhen.* **7**:2 (1973), 83–84. In Russian; translated in *Funct. Anal. Appl.* **7** (1973), 155–156. MR Zbl

[Elkies 1998] N. D. Elkies, "Elliptic and modular curves over finite fields and related computational issues", pp. 21–76 in *Computational perspectives on number theory* (Chicago, 1995), edited by D. A. Buell and J. T. Teitelbaum, AMS/IP Stud. Adv. Math. **7**, Amer. Math. Soc., Providence, RI, 1998. MR Zbl

[van Hoeij 2012] M. van Hoeij, "Low degree places on the modular curve $X_1(N)$", preprint, 2012. arXiv 1202.4355

[Jeon et al. 2011a] D. Jeon, C. H. Kim, and Y. Lee, "Families of elliptic curves over cubic number fields with prescribed torsion subgroups", *Math. Comp.* **80**:273 (2011), 579–591. MR Zbl

[Jeon et al. 2011b] D. Jeon, C. H. Kim, and Y. Lee, "Families of elliptic curves over quartic number fields with prescribed torsion subgroups", *Math. Comp.* **80**:276 (2011), 2395–2410. MR Zbl

[Kamienny 1992a] S. Kamienny, "Torsion points on elliptic curves and q-coefficients of modular forms", *Invent. Math.* **109**:2 (1992), 221–229. MR Zbl

[Kamienny 1992b] S. Kamienny, "Torsion points on elliptic curves over fields of higher degree", *Int. Math. Res. Not.* **1992**:6 (1992), 129–133. MR Zbl

[Kamienny and Mazur 1995] S. Kamienny and B. Mazur, "Rational torsion of prime order in elliptic curves over number fields", pp. 81–100 in *Columbia University Number Theory Seminar* (New York, 1992), Astérisque **228**, Soc. Math. France, Paris, 1995. MR Zbl

[Kato 2004] K. Kato, "p-adic Hodge theory and values of zeta functions of modular forms", pp. 117–290 in *Cohomologies p-adiques et applications arithmétiques*, *III*, edited by P. Berthelot et al., Astérisque **295**, Soc. Math. France, Paris, 2004. MR Zbl

[Kim 2003] H. H. Kim, "Functoriality for the exterior square of GL_4 and the symmetric fourth of GL_2", *J. Amer. Math. Soc.* **16**:1 (2003), 139–183. MR Zbl

[Kolyvagin and Logachëv 1989] V. A. Kolyvagin and D. Y. Logachëv, "Finiteness of the Shafarevich–Tate group and the group of rational points for some modular abelian varieties", *Algebra i Analiz* **1**:5 (1989), 171–196. In Russian; translated in *Leningrad Math. J.* **1**:5 (1990), 1229–1253. MR Zbl

[Manin 1972] J. I. Manin, "Parabolic points and zeta functions of modular curves", *Izv. Akad. Nauk SSSR Ser. Mat.* **36** (1972), 19–66. In Russian; translated in *Math. USSR-Izv.* **6** (1972), 19–64. MR Zbl

[Mazur 1977] B. Mazur, "Modular curves and the Eisenstein ideal", *Inst. Hautes Études Sci. Publ. Math.* **47** (1977), 33–186. MR Zbl

[Mazur 1978] B. Mazur, "Rational isogenies of prime degree", *Invent. Math.* **44**:2 (1978), 129–162. MR Zbl

[Merel 1996] L. Merel, "Bornes pour la torsion des courbes elliptiques sur les corps de nombres", *Invent. Math.* **124**:1-3 (1996), 437–449. MR Zbl

[Oesterlé 1994] J. Oesterlé, "Torsion des courbes elliptiques sur les corps de nombres", unpublished notes, 1994.

[Parent 1999] P. Parent, "Bornes effectives pour la torsion des courbes elliptiques sur les corps de nombres", *J. Reine Angew. Math.* **506** (1999), 85–116. MR Zbl

[Parent 2000] P. Parent, "Torsion des courbes elliptiques sur les corps cubiques", *Ann. Inst. Fourier* (*Grenoble*) **50**:3 (2000), 723–749. MR Zbl

[Parent 2003] P. Parent, "No 17-torsion on elliptic curves over cubic number fields", *J. Théor. Nombres Bordeaux* **15**:3 (2003), 831–838. MR Zbl

[Rebolledo 2009] M. Rebolledo, "Merel's theorem on the boundedness of the torsion of elliptic curves", pp. 71–82 in *Arithmetic geometry*, edited by H. Darmon et al., Clay Math. Proc. **8**, Amer. Math. Soc., Providence, RI, 2009. MR Zbl

[SageMath] W. A. Stein et al., "Sage mathematics software", Version 9.2, available at http://www.sagemath.org.

[Stein 2007] W. Stein, *Modular forms, a computational approach*, Grad. Stud. in Math. **79**, Amer. Math. Soc., Providence, RI, 2007. MR Zbl

[Stevens 1982] G. Stevens, *Arithmetic on modular curves*, Progr. Math. **20**, Birkhäuser, Boston, 1982. MR Zbl

[Sutherland 2012] A. V. Sutherland, "Constructing elliptic curves over finite fields with prescribed torsion", *Math. Comp.* **81**:278 (2012), 1131–1147. MR Zbl

[Sutherland 2013] A. V. Sutherland, "Isogeny volcanoes", pp. 507–530 in *ANTS X*: *Proceedings of the Tenth Algorithmic Number Theory Symposium* (San Diego, CA, 2012), edited by E. W. Howe and K. S. Kedlaya, Open Book Ser. **1**, Math. Sci. Publ., Berkeley, CA, 2013. MR Zbl

[Waterhouse 1969] W. C. Waterhouse, "Abelian varieties over finite fields", *Ann. Sci. École Norm. Sup.* (4) **2** (1969), 521–560. MR Zbl

Communicated by Bjorn Poonen
Received 2017-07-18 Revised 2021-02-05 Accepted 2022-02-01

maarten@mderickx.nl *Mathematisch Instituut, Universiteit Leiden, Leiden, Netherlands*
kamienny@usc.edu *Department of Mathematics, USC Dornsife, Los Angeles, CA, United States*
wstein@gmail.com *SageMath, Inc., Renton, WA, United States*
michael.stoll@uni-bayreuth.de *Mathematisches Institut, Universität Bayreuth, Bayreuth, Germany*

Algebra & Number Theory

msp.org/ant

See inside back cover or msp.org/ant for submission instructions.

The subscription price for 2023 is US $485/year for the electronic version, and $705/year (+$65, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online.

ANT peer review and production are managed by EditFLOW® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

Algebra & Number Theory

Volume 17 No. 2 2023