TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 362, Number 3, March 2010, Pages 1547–1570
S 0002-9947(09)04804-1
Article electronically published on September 30, 2009

ALMOST ALL ELLIPTIC CURVES ARE SERRE CURVES

NATHAN JONES

Abstract. Using a multidimensional large sieve inequality, we obtain a bound for the mean-square error in the Chebotarev theorem for division fields of elliptic curves that is as strong as what is implied by the Generalized Riemann Hypothesis. As an application we prove that, according to height, almost all elliptic curves are Serre curves, where a Serre curve is an elliptic curve whose torsion subgroup, roughly speaking, has as much Galois symmetry as possible.

1. Introduction

Let E be an elliptic curve defined over $\mathbb{Q}$ and denote by

$$\phi_{N,E} : G_{\mathbb{Q}} \to \mathrm{Aut}(E[N])$$

the representation of $G_{\mathbb{Q}} := \mathrm{Gal}\,(\overline{\mathbb{Q}}/\mathbb{Q})$ on the N-torsion $E[N]$ of E. Fixing a $\mathbb{Z}/N\mathbb{Z}$-basis of $E[N]$, we identify $\mathrm{Aut}(E[N])$ with $GL_2(\mathbb{Z}/N\mathbb{Z})$ and write

$$\phi_{N,E} : G_{\mathbb{Q}} \to GL_2(\mathbb{Z}/N\mathbb{Z}).$$

The image $\phi_{N,E}(G_{\mathbb{Q}})$ is exactly the Galois group of the *Nth division field of E over* $\mathbb{Q}$, i.e. the field obtained by adjoining to $\mathbb{Q}$ the x and y coordinates of the N-torsion of a given Weierstrass model of E, which we will denote by $\mathbb{Q}(E[N])$. Taking the inverse limit over all $N \geq 1$ with the bases chosen compatibly, we obtain the full torsion representation

$$\phi_E : G_{\mathbb{Q}} \to GL_2(\hat{\mathbb{Z}}) := \lim_{\leftarrow} GL_2(\mathbb{Z}/N\mathbb{Z}).$$

It is natural to wonder how large the image of ϕ_E in $GL_2(\hat{\mathbb{Z}})$ is.

Definition 1. The integer N is said to be **exceptional** for E if $\phi_{N,E}$ is not surjective.

To wonder about the size of the image of ϕ_E in $GL_2(\hat{\mathbb{Z}})$ is simply to wonder about which numbers N are exceptional for E, and about "how exceptional each N is", i.e. about the index $[GL_2(\mathbb{Z}/N\mathbb{Z}) : \phi_{N,E}(G_{\mathbb{Q}})]$.

When E has complex multiplication, $\mathbb{Q}(E[N])$ is always an abelian extension of the CM field (Kronecker's "Jugendtraum"; see [24, Theorem 2.3, p. 108]), from which it follows that every N except possibly $N = 2$ is exceptional, so that the image $\phi_E(G_{\mathbb{Q}})$ has infinite index in $GL_2(\hat{\mathbb{Z}})$. On the other hand, when E does not have CM, Serre [20] has shown that the index $[GL_2(\hat{\mathbb{Z}}) : \phi_E(G_{\mathbb{Q}})]$ is finite. Equivalently, there exists an integer m_E so that

$$\phi_E(G_{\mathbb{Q}}) = \pi^{-1}(\phi_{m_E,E}(G_{\mathbb{Q}})), \tag{1}$$

Received by the editors May 4, 2007 and, in revised form, April 3, 2008.
2000 *Mathematics Subject Classification.* Primary 11G05, 11F80.

where $\pi : GL_2(\hat{\mathbb{Z}}) \to GL_2(\mathbb{Z}/m_E\mathbb{Z})$ is the natural projection. In particular, this implies that any fixed non-CM elliptic curve E has only finitely many exceptional primes, since any such exceptional prime must divide m_E. One might wonder how the integer m_E (chosen minimally so that (1) still holds) depends on the curve E. Various results exist which bound the largest possible exceptional prime for E. For example, Mazur [17] proves that if E is semistable, then no prime $N \geq 11$ can be exceptional for E. Other authors have bounded the largest possible exceptional prime in terms of invariants of the elliptic curve, such as the height [16] and conductor ([21], [15], and [2]).

Results also exist which count the number of elliptic curves with no exceptional primes. Let $E_{r,s}$ denote the plane curve given by the equation

$$E_{r,s} \ : \ y^2 = x^3 + rx + s. \tag{2}$$

For a varying parameter x let $R(x)$ and $S(x)$ be a given length and width (understood to grow with x) and define

$$C(x) := \{E_{r,s} : (r,s) \in \mathbb{Z}^2, |r| \leq R(x), |s| \leq S(x) \text{ and } 4r^3 + 27s^2 \neq 0\}.$$

Duke [8] takes $R(x) = x^2$ and $S(x) = x^3$ (which are the choices defining naive height) and shows that if $\varepsilon(x)$ is the set of $E_{r,s} \in C(x)$ which have at least one exceptional prime, then

$$\lim_{x\to\infty} \frac{|\varepsilon(x)/\simeq|}{|C(x)/\simeq|} = 0, \tag{3}$$

where $E_{r,s} \simeq E_{r',s'}$ if the two models are isomorphic over $\mathbb{Q}$. Using a two-dimensional large sieve inequality, he shows that

$$|\varepsilon(x)/\simeq| \ll x^4 \log^B x,$$

with an absolute (but ineffective) constant. Since

$$|C(x)/\simeq| = \frac{4}{\zeta(10)}x^5 + O(x^3)$$

(see [1]), this implies (3). Cojocaru and Hall [3] prove a similar result for elliptic curves in one-parameter families.

In [12], Grant obtains an asymptotic formula for $|\varepsilon(x)/\simeq|$. He shows that the curves which are exceptional at the primes 2 and 3 contribute the main term of $|\varepsilon(x)/\simeq|$ and that, for an explicit constant C,

$$|\varepsilon(x)/\simeq| = Cx^3 + O_\epsilon(x^{2+\epsilon})$$

for all $\epsilon > 0$.

This paper gives a different generalization. The statement that an elliptic curve E has no exceptional primes may be viewed as saying that the Galois representation ϕ_E has "large image". In this paper we extend (3) to a result that almost all elliptic curves have $\phi_E(G_{\mathbb{Q}})$ "as large as possible".

2. Statement of results

Our main result is a theorem bounding the mean-square error in the Chebotarev theorem for division fields of elliptic curves. Fix a positive integer level N and a subset

$$\mathcal{A} \subset GL_2(\mathbb{Z}/N\mathbb{Z})$$

which is closed under conjugation by $GL_2(\mathbb{Z}/N\mathbb{Z})$ and represents only one determinant value, i.e. which satisfies

$$\forall g \in GL_2(\mathbb{Z}/N\mathbb{Z}),\ g\mathcal{A}g^{-1} = \mathcal{A} \quad \text{and} \quad \forall a, b \in \mathcal{A},\ \det a = \det b \tag{4}$$

(for instance, we could take $\mathcal{A}$ to be a conjugacy class). Denote by

$$\pi_E(x; N, \mathcal{A}) := |\{p \le x : \phi_{N,E}(\mathrm{Frob}_p) \subseteq \mathcal{A}\}|$$

the function which counts the number of primes up to x which are unramified in $\mathbb{Q}(E[N])$ and whose Frobenius class is contained in $\mathcal{A}$, and as usual let

$$\pi(x; N, d) := |\{p \le x : p \equiv d \mod N\}|.$$

Theorem 2. *For $x \ge 1$ and $\min\{R(x), S(x)\} \ge x^2$, one has*

$$\frac{1}{|C(x)|} \sum_{E \in C(x)} \left(\pi_E(x; N, \mathcal{A}) - \frac{|\mathcal{A}|\varphi(N)}{|GL_2(\mathbb{Z}/N\mathbb{Z})|} \pi(x; N, d) \right)^2 \ll |\mathcal{A}|^2 x,$$

where $d := \det \mathcal{A}$, $\varphi(N)$ denotes the Euler-phi function, and the implied constant is absolute.

In [8], Duke proves this (without the $|\mathcal{A}|^2$ factor) for *prime* level N and where $\mathcal{A}$ has the specific form

$$\mathcal{A} = GL_2(\mathbb{Z}/N\mathbb{Z})_{t,d} := \{A \in GL_2(\mathbb{Z}/N\mathbb{Z}) :\ \mathrm{tr}\, A = t,\ \det A = d\}.$$

Such sets are unions of conjugacy classes. For example, even when N is prime, the set $GL_2(\mathbb{Z}/N\mathbb{Z})_{2\lambda,\lambda^2}$ contains two conjugacy classes, represented by the matrices

$$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \text{ and } \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix},$$

respectively. Theorem 2 distinguishes between these two cases.

Our second result is an application of Theorem 2 to the problem of counting elliptic curves E for which $\phi_E(G_{\mathbb{Q}})$ is as large as possible. First of all, how large can this image be? Does there exist an elliptic curve E with ϕ_E surjective? In other words, is there a curve E with $m_E = 1$? Serre [20] answers no. For each elliptic curve E over $\mathbb{Q}$, there is an index two subgroup $H_E \subset GL_2(\hat{\mathbb{Z}})$ (for a precise definition, see Section 4) such that

$$\phi_E(G_{\mathbb{Q}}) \subseteq H_E. \tag{5}$$

Definition 3. We call an elliptic curve E a **Serre curve** when equality holds in (5).

Our second theorem is

Theorem 4. *Let $C_{Serre}(x)$ denote the set*

$$\{E_{r,s} \in C(x) : E_{r,s} \text{ is a Serre curve }\}.$$

Assuming that $\min\{R(x), S(x)\} \ge x^2$, one has

$$|C(x) - C_{Serre}(x)| \ll \frac{|C(x)| \log^B x}{x},$$

where B is an explicit constant. Thus, in particular,

$$\lim_{x \to \infty} \frac{|C_{Serre}(x)|}{|C(x)|} = 1.$$

In order to obtain this result that "almost all elliptic curves are Serre curves", we prove an algebraic lemma which gives a sufficient condition for an elliptic curve E to be a Serre curve.

Lemma 5. *Suppose E over $\mathbb{Q}$ is an elliptic curve such that:*

1. *E has no exceptional primes; and*
2. *E is not exceptional at 72.*

Then, E is a Serre curve.

This lemma is used together with Theorem 2 to give Theorem 4. In [14], we use Theorem 4 to compute the average value over elliptic curves of the Lang-Trotter constants, answering a question of David and Pappalardi [6].

We remark that there are differences between authors as to how to count "elliptic curves over $\mathbb{Q}$". Some authors count isomorphism classes of elliptic curves over $\mathbb{Q}$, while others count models $E_{r,s}$. We choose to count models, but in our results the distinction is only technical: the statements of Theorem 2 and Theorem 4 are seen without difficulty to hold if one replaces "$C(x)$" with "$C(x)/\simeq$", and vice versa with the results we have quoted. In particular the work of [8] shows also that

$$\min\{R(x), S(x)\} \geq x^2 \quad \Longrightarrow \quad |\varepsilon(x)| \ll \frac{|C(x)| \log^B x}{x}. \tag{6}$$

It is interesting to consider what the situation might look like for elliptic curves over a general number field K, as well as to refine the upper bound of Theorem 4 to an asymptotic in the style of Grant. The study of these problems is recent doctoral work in progress by D. Zywina and V. Radhakrishnan, respectively.

The paper is organized as follows. In Section 3 we prove Theorem 2. Section 4 gives the complete definition of a Serre curve, and Section 5 is devoted to a proof of Lemma 5. Finally in Section 6 we prove Theorem 4, and in Section 7 we produce an example of a one-parameter family of elliptic curves which are exceptional at $N = 4$ but not at $N = 2$.

3. Bounding mean-square Chebotarev error

In this section we prove Theorem 2. We first remark that although the result gives a bound as strong as the appropriate Generalized Riemann Hypothesis would, its proof is unconditional. It employs the following large sieve inequality of Gallagher [11, Lemma A] and proceeds along the same lines as the proof of [8, Theorem 2].

Lemma 6. *Let k be a positive integer and for each prime number p let $\Omega(p) \subseteq (\mathbb{Z}/p\mathbb{Z})^k$ be any subset. For each fixed $m \in \mathbb{Z}^k$ we define*

$$P(x; m) = |\{p \leq x : m \mod p \in \Omega(p)\}|$$

and

$$P(x) = \sum_{p \leq x} |\Omega(p)| p^{-k}.$$

Let B be a box in $\mathbb{R}^k$ whose sides are parallel to the coordinate planes and which has minimum width $W(B)$ and volume $V(B)$. If $W(B) \geq x^2$, then

$$\sum_{m \in B \cap \mathbb{Z}^k} (P(x; m) - P(x))^2 \ll V(B) P(x),$$

where the implied constant depends only on k.

We recall the setting of Theorem 2: for a pair of integers (r, s) let $E_{r,s}$ be the curve defined by (2). Let N be a positive integer and fix a subset

$$\mathcal{A} \subset GL_2(\mathbb{Z}/N\mathbb{Z})$$

satisfying (4). We will proceed to define the set $\Omega(p) = \Omega_{\mathcal{A}}(p)$ in such a way that $P(x; (r, s))$ and $P(x)$ will satisfy

$$P(x; (r, s)) = \pi_{E_{r,s}}(x; N, \mathcal{A}) + O(1) \tag{7}$$

and

$$P(x) = \frac{|\mathcal{A}|\varphi(N)}{|GL_2(\mathbb{Z}/N\mathbb{Z})|}\pi(x; N, d) + O(|\mathcal{A}|x^{1/2}), \tag{8}$$

respectively, where the implied constants are absolute. Applying Lemma 6 and observing that $(A + B)^2 \ll A^2 + B^2$, we conclude the result of Theorem 2.

3.1. **Defining the set $\Omega_{\mathcal{A}}(p) \subseteq (\mathbb{Z}/p\mathbb{Z})^2$.** We begin by quoting a result of Duke and Tóth [9] which describes explicitly the conjugacy class in $\mathrm{Gal}\,(\mathbb{Q}(E[N])/\mathbb{Q})$ of the Frobenius automorphism at a prime p which is unramified in $\mathbb{Q}(E[N])$. The description is given purely in terms of data attached to E_p, the reduction of E modulo p.

In our context, their result may be stated as follows: let $\mathbb{F}_p$ denote the finite field of p elements and E_p any elliptic curve defined over $\mathbb{F}_p$. Let

$$a = a(E_p) := p + 1 - |E_p(\mathbb{F}_p)|$$

be the trace of the Frobenius endomorphism ϕ_p of E_p and $b = b(E_p)$ the index in the ring of $\mathbb{F}_p$-endomorphisms of E_p of the subring generated by the Frobenius endomorphism, i.e.

$$b = [\mathrm{End}_{\mathbb{F}_p}(E_p) : \mathbb{Z}[\phi_p]].$$

In any case (including the supersingular case), the ring $\mathrm{End}_{\mathbb{F}_p}(E_p)$ is isomorphic to an imaginary quadratic order (see [25, Theorem 4.2]), whose discriminant we denote by $\Delta = \Delta(E_p)$. The comparison of discriminants yields

$$\Delta b^2 = a^2 - 4p. \tag{9}$$

We associate to E_p the following matrix of trace a and determinant p:

$$\sigma(E_p) = \begin{pmatrix} (a + b\delta)/2 & b \\ b(\Delta - \delta)/4 & (a - b\delta)/2 \end{pmatrix}, \tag{10}$$

where for a discriminant Δ we have $\delta = 0$ or 1 according to whether $\Delta \equiv 0$ or $1 \bmod 4$. Note that, because of (9), σ has integer entries.

Theorem 7. *Let E be an elliptic curve over $\mathbb{Q}$ and let N be any positive integer. If p is a prime of good reduction for E which does not divide N, then p is unramified in $\mathbb{Q}(E[N])$. Furthermore, denoting by E_p the reduction of E modulo p, the integral matrix $\sigma(E_p)$ defined in* (10)*, when reduced modulo N, represents the class of the Frobenius automorphism at p in* $\mathrm{Gal}\,(\mathbb{Q}(E[N])/\mathbb{Q})$.

Now suppose $p > 3$ is a prime number. For $(r, s) \in \mathbb{F}_p^2$, let $E_{r,s}$ denote the curve over $\mathbb{F}_p$ given by equation (2) and $\Delta_{r,s} = -16(4r^3 + 27s^2)$ its associated discriminant. We define $\Omega_{\mathcal{A}}(p) = \emptyset$ if $p \mid 6N$, and for $p \nmid 6N$,

$$\Omega_{\mathcal{A}}(p) := \{(r, s) \in \mathbb{F}_p^2 : \Delta_{r,s} \neq 0 \text{ and } \sigma(E_{r,s}) \mod N \in \mathcal{A}\}.$$

Observe that for $(r,s) \in \mathbb{Z}^2$, the discriminant $\Delta_{r,s}$ of an elliptic curve $E_{r,s}$ over $\mathbb{Q}$ is related to its minimal discriminant Δ by

$$\Delta_{r,s} = e^{12}\Delta$$

for some e dividing 6. Thus, Theorem 7 implies that (7) holds. We now turn to verifying (8).

3.2. **The asymptotic in p of $|\Omega_{\mathcal{A}}(p)|$.** The goal of this section is to give the asymptotic of $|\Omega_{\mathcal{A}}(p)|$ as p ranges through the set of prime numbers for which $\Omega_{\mathcal{A}}(p) \neq \emptyset$. Our proof will show that, in fact,

$$\Omega_{\mathcal{A}}(p) \neq \emptyset \Longleftrightarrow p \equiv \det \mathcal{A} \mod N.$$

Theorem 8. *For p prime congruent to $\det \mathcal{A}$ modulo N we have*

$$|\Omega_{\mathcal{A}}(p)| = \frac{|\mathcal{A}|\varphi(N)}{|GL_2(\mathbb{Z}/N\mathbb{Z})|}p^2 + O(|\mathcal{A}|p^{3/2}),$$

where the implied constant is absolute.

We observe that (8) follows from this asymptotic. Thus, Theorem 2 will follow from Theorem 8 together with Lemma 6.

In order to prove Theorem 8, we first express $|\Omega_{\mathcal{A}}(p)|$ in terms of a weighted class number. Define the set

$$\mathcal{T}_{\mathcal{A}}(p) := \{A \in M_{2\times 2}(\mathbb{Z}) : \det A = p,\ A \mod N \in \mathcal{A}\}$$

and the subset of elliptic matrices

$$\mathcal{T}^e_{\mathcal{A}}(p) := \{A \in \mathcal{T}_{\mathcal{A}}(p) : (\operatorname{tr} A)^2 - 4\det A < 0\}.$$

Since $\mathcal{A}$ is stable by $SL_2(\mathbb{Z}/N\mathbb{Z})$-conjugation, both of the above sets are stable by $SL_2(\mathbb{Z})$-conjugation.

Note: Throughout the rest of this paper we will use the standard notation

$$\Gamma(1) := SL_2(\mathbb{Z}).$$

The following two auxiliary results will imply Theorem 8.

Proposition 9.

$$|\Omega_{\mathcal{A}}(p)| = \frac{p-1}{2}\sum_{\alpha \in \mathcal{T}^e_{\mathcal{A}}(p)\,//\,\Gamma(1)} \frac{1}{|\Gamma(1)_\alpha|},$$

where $\mathcal{T}^e_{\mathcal{A}}(p)\,//\,\Gamma(1)$ is the set of $\Gamma(1)$-conjugation orbits in $\mathcal{T}^e_{\mathcal{A}}(p)$ and

$$\Gamma(1)_\alpha := \{\gamma \in \Gamma(1) : \gamma\alpha = \alpha\gamma\}.$$

Lemma 10. *If $p \equiv \det \mathcal{A} \mod N$, then*

$$\sum_{\alpha \in \mathcal{T}^e_{\mathcal{A}}(p)\,//\,\Gamma(1)} \frac{1}{|\Gamma(1)_\alpha|} = \frac{2|\mathcal{A}|}{|SL_2(\mathbb{Z}/N\mathbb{Z})|}p + O(|\mathcal{A}|p^{1/2}),$$

with an absolute constant.

Proof of Lemma 10. This is [13, Corollary 5]. □

The remainder of this section is devoted to proving Proposition 9. First, by writing $\mathcal{A}$ as a disjoint union of $GL_2(\mathbb{Z}/N\mathbb{Z})$-conjugacy classes,

$$\mathcal{A} = \bigsqcup_{i=1}^{\eta} \mathcal{A}_i,$$

and observing that

$$\Omega_{\mathcal{A}}(p) = \bigsqcup_{i=1}^{\eta} \Omega_{\mathcal{A}_i}(p) \quad \text{and} \quad \mathcal{T}^e_{\mathcal{A}}(p) = \bigsqcup_{i=1}^{\eta} \mathcal{T}^e_{\mathcal{A}_i}(p),$$

we may (and will henceforth) assume that $\mathcal{A}$ *is* a $GL_2(\mathbb{Z}/N\mathbb{Z})$-conjugacy class. Note that

$$\Omega_{\mathcal{A}}(p) = \{(r,s) \in (\mathbb{Z}/p\mathbb{Z})^2 : \ \Delta_{r,s} \neq 0 \ \text{ and } \ \sigma(E_{r,s}) \in \mathcal{T}^e_{\mathcal{A}}(p)\}.$$

At this point we must give a finer description of the conjugacy class $\mathcal{A}$. For any divisor M of N and integers $\overline{T}$, $\overline{D}$ modulo N/M, define

$$\mathcal{T}_{N/M}(\overline{T}, \overline{D}) = \{A \in M_{2\times 2}(\mathbb{Z}/(N/M)\mathbb{Z}) : \ (\operatorname{tr} A, \det A) \equiv (\overline{T}, \overline{D}) \mod N/M\}$$

and

$$\mathcal{T}^*_{N/M}(\overline{T}, \overline{D}) = \{A \in \mathcal{T}_{N/M}(\overline{T}, \overline{D}) : \ A \text{ is non-scalar mod each prime } l \mid N/M\}.$$

The following lemma describes the structure of conjugacy classes in the group $GL_2(\mathbb{Z}/N\mathbb{Z})$.

Lemma 11. *Any conjugacy class*

$$\mathcal{A} \subset GL_2(\mathbb{Z}/N\mathbb{Z})$$

has the form

$$\mathcal{A} = \lambda I + M\mathcal{T}^*_{N/M}(\overline{T}, \overline{D}), \tag{11}$$

where M divides N and λ is an integer satisfying $0 \leq \lambda < M$.

Proof of Lemma 11. Since the set $\lambda I + M\mathcal{T}^*_{N/M}(\overline{T}, \overline{D})$ is stable by $GL_2(\mathbb{Z}/N\mathbb{Z})$-conjugation, it suffices to show that for any matrix $A \in \mathcal{T}^*_{N/M}(\overline{T}, \overline{D})$, we can find $B \in GL_2(\mathbb{Z}/(N/M)\mathbb{Z})$ with

$$BAB^{-1} = \begin{pmatrix} 0 & -\overline{D} \\ 1 & \overline{T} \end{pmatrix}.$$

To this end, let $v = \begin{pmatrix} x \\ y \end{pmatrix} \in (\mathbb{Z}/(N/M)\mathbb{Z})^2$ be a variable vector and notice that the linear transformation L_A on $(\mathbb{Z}/(N/M)\mathbb{Z})^2$ given by left multiplication by A has the form

$$[L_A]_{\{v, Av\}} = \begin{pmatrix} 0 & -\overline{D} \\ 1 & \overline{T} \end{pmatrix}$$

when written with respect to the ordered basis $\{v, Av\}$ of $(\mathbb{Z}/(N/M)\mathbb{Z})^2$. This verifies the claim, provided that we can find a vector v so that the change of basis matrix

$$B = \begin{pmatrix} x & ax+by \\ y & cx+dy \end{pmatrix} \qquad \left(A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\right)$$

belongs to $GL_2(\mathbb{Z}/(N/M)\mathbb{Z})$, i.e. so that

$$\det B = cx^2 + (d-a)xy - by^2 \in (\mathbb{Z}/(N/M)\mathbb{Z})^*.$$

By the Chinese Remainder Theorem and the fact that A is non-scalar modulo each prime l dividing N/M, we may take

$$(x,y) \equiv \begin{cases} (1,0) & \text{if } l \nmid c \\ (0,1) & \text{if } l \nmid b \\ (1,1) & \text{if } l \mid b \text{ and } l \mid c, \end{cases}$$

which finishes the proof of Lemma 11. □

Let us henceforth assume that our conjugacy class $\mathcal{A}$ is of the form (11). We would like to partition $\mathcal{T}_{\mathcal{A}}^e(p)$ into subsets which are stable by $\Gamma(1)$-conjugation. Let $\mathcal{T}^*(T,D,f)$ denote the set

$$\left\{ A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_{2\times 2}(\mathbb{Z}) : \ \operatorname{tr} A = T, \det A = D, \gcd(b, d-a, c) = f \right\}.$$

We note then that the trace t and determinant d of any matrix in the set $\lambda I + M \cdot \mathcal{T}^*(T,D,f)$ satisfy

$$t = 2\lambda + MT, \quad d = \lambda^2 + M\lambda T + M^2 D, \quad \text{and} \quad t^2 - 4d = M^2\left(T^2 - 4D\right). \tag{12}$$

Thus, from Lemma 11 we see that

$$\mathcal{T}_{\mathcal{A}}^e(p) = \bigsqcup_{(T,D)} \bigsqcup_{\substack{f \geq 1 \\ \gcd(f, N/M) = 1}} \left(\lambda I + M \cdot \mathcal{T}^*(T,D,f)\right),$$

where (T,D) runs over integer pairs satisfying

$$(T,D) \equiv (\overline{T}, \overline{D}) \mod N/M, \ p = \lambda^2 + M\lambda T + M^2 D, \text{ and } (2\lambda + MT)^2 < 4p.$$

Defining $\Omega^*(\lambda, M, T, D, f)$ by

$$\{(r,s) \in (\mathbb{Z}/p\mathbb{Z})^2 : \ \Delta_{r,s} \neq 0 \text{ and } \sigma(E_{r,s}) \in \lambda I + M \cdot \mathcal{T}^*(T,D,f)\},$$

Proposition 9 is reduced to showing that

$$|\Omega^*(\lambda, M, T, D, f)| = \frac{p-1}{2} \sum_{\alpha \in (\lambda I + M \cdot \mathcal{T}^*(T,D,f)) \,//\, \Gamma(1)} \frac{1}{|\Gamma(1)_\alpha|}. \tag{13}$$

Lemma 12. *We have that $\Omega^*(\lambda, M, T, D, f)$ is equal to*

$$\{(r,s) \in (\mathbb{Z}/p\mathbb{Z})^2 : \ \Delta_{r,s} \neq 0, \ b(E_{r,s}) = Mf \ \text{ and } \ a(E_{r,s}) = 2\lambda + MT\}.$$

Proof. The containment "$\Omega^*(\lambda, M, T, D, f) \subseteq \ldots$" is immediate from (10) and (12). The reverse containment comes from the fact that, for fixed t and p, the two equations

$$t = 2\lambda + MT \quad \text{and} \quad p = \lambda^2 + M\lambda T + M^2 D$$

have a unique solution $(\lambda, T, D) \in \{0, 1, \ldots, M-1\} \times \mathbb{Z}^2$, if they have one at all. This fact is immediate when M is odd. If M is even, we see from the first equation that the only way two distinct solutions can exist is if one solution looks like (λ, T, D) with $\lambda \in \{0, 1, \ldots, M/2 - 1\}$ and the other solution has the form $(\lambda + M/2, T - 1, D')$ for some integer D'. But then the second equation gives us the contradiction that

$$\lambda^2 + M\lambda T - p \equiv 0 \mod M^2 \quad \text{and} \quad \lambda^2 + M\lambda T - p \equiv \frac{M^2}{4}(1 - 2T) \mod M^2.$$

□

We now summarize some fundamental facts about imaginary quadratic orders. More details may be found, for example, in [4, §7]. An **imaginary quadratic order** $\mathcal{O}$ is a subring (containing 1) of an imaginary quadratic field K which contains a basis of K over $\mathbb{Q}$ and has rank 2 as a free abelian group. For each negative integer Δ satisfying

$$\Delta \equiv 0 \text{ or } 1 \mod 4,$$

there is a unique imaginary quadratic order of discriminant Δ, which we will denote by $\mathcal{O}(\Delta)$. Orders $\mathcal{O}(\Delta')$ which contain a given order $\mathcal{O}(\Delta)$ are exactly those orders whose discriminant Δ' satisfies

$$f^2\Delta' = \Delta, \qquad f = [\mathcal{O}(\Delta') : \mathcal{O}(\Delta)].$$

Every imaginary quadratic order $\mathcal{O}$ is contained in a unique maximal imaginary quadratic order,

$$\mathcal{O} \subseteq \mathcal{O}_{\max} = \mathcal{O}_K \subset K,$$

which is the ring of integers of K. The ideal class group $\mathcal{C}(\mathcal{O})$ is the group of invertible fractional ideals of $\mathcal{O}$ modulo the subgroup of principal fractional ideals. This is a finite group whose size we denote by $h(\mathcal{O})$.

Lemma 13. *Suppose $p \geq 5$ is prime and t is any integer satisfying $t^2 < 4p$. Let $\mathcal{O}$ be any imaginary quadratic order containing the order of discriminant $t^2 - 4p$. The number of elliptic curves $E_{r,s}$ over $\mathbb{F}_p$ of the form (2) which satisfy*

$$a(E_{r,s}) = t \quad \text{and} \quad \operatorname{End}_{\mathbb{F}_p}(E_{r,s}) = \mathcal{O}$$

is given by

$$\frac{p-1}{|\mathcal{O}^*|} h(\mathcal{O}),$$

where $\mathcal{O}^*$ is the group of units of $\mathcal{O}$.

Proof. The following theorem restates [25, Theorems 4.2 and 4.5], specialized to our situation. See also [18], which corrects a small error in the proof. The original work is due to Deuring [7].

Theorem 14. *Let t be any integer satisfying $t^2 < 4p$. Then the following are precisely the rings which occur as rings of $\mathbb{F}_p$-endomorphisms of some elliptic curve E_p over $\mathbb{F}_p$ satisfying $a(E_p) = t$:*

- *if $t \neq 0$, all complex quadratic orders containing $\mathcal{O}(t^2 - 4p)$;*
- *if $t = 0$, all complex quadratic orders $\mathcal{O}$ satisfying*

$$\mathcal{O}(-4p) \subset \mathcal{O} \quad \text{and} \quad p \nmid [\mathcal{O}_{\max} : \mathcal{O}].$$

Furthermore, given such an order $\mathcal{O}$, the number of $\mathbb{F}_p$-isomorphism classes of elliptic curves E_p over $\mathbb{F}_p$ satisfying

$$a(E_p) = t \quad \text{and} \quad \operatorname{End}_{\mathbb{F}_p}(E_p) = \mathcal{O}$$

is equal to $h(\mathcal{O})$.

Note that, since $p \geq 5$, every $\mathbb{F}_p$-isomorphism class contains an elliptic curve of the form (2). By the theorem, the proof of Lemma 13 is reduced to showing that whenever $E_{r,s}$ is of the form (2) with $a(E_{r,s}) = t$ and $\operatorname{End}_{\mathbb{F}_p}(E_{r,s}) = \mathcal{O}$, the number of elliptic curves of the same form which are isomorphic over $\mathbb{F}_p$ to $E_{r,s}$ is $(p-1)/|\mathcal{O}^*|$. Such elliptic curves are exactly those given by the equations

$$E_{ru^4, su^6} : \; y^2 = x^3 + ru^4x + su^6, \quad u \in (\mathbb{Z}/p\mathbb{Z})^*.$$

In the case where $|\mathcal{O}^*| = 2$, neither r nor s can be equal to zero (see [23, Theorem 10.1, p. 103]). In this case, $E_{ru^4,su^6} = E_{r(u')^4,s(u')^6}$ if and only if $u = \pm u'$ and we count exactly $(p-1)/2$ distinct E_{ru^4,su^6}'s. The case of $|\mathcal{O}^*| = 4$ occurs exactly when $\mathcal{O} = \mathcal{O}(-4) = \mathbb{Z}[i]$ is the ring of Gaussian integers, and this happens only if $s = 0$. Since

$$\mathcal{O}(t^2 - 4p) \subset \mathcal{O}(-4),$$

we see by relating the discriminants that t must be even and that $p \equiv 1 \mod 4$. Choosing $i_p \in (\mathbb{Z}/p\mathbb{Z})^*$ satisfying $i_p^2 = -1$, we note that in this case $E_{ru^4,su^6} = E_{r(u')^4,s(u')^6}$ if and only if $u/u' \in \{\pm i_p, \pm 1\}$, and so there are again exactly $(p-1)/|\mathcal{O}^*|$ elliptic curves of the form (2) isomorphic over $\mathbb{F}_p$ to $E_{r,s}$. The $|\mathcal{O}^*| = 6$ case is quite similar, so we omit it. This finishes the proof of Lemma 13. □

Returning to the verification of (13), we see by the two lemmas and (12) that

$$|\Omega^*(\lambda, M, T, D, f)| = \frac{p-1}{\left|\mathcal{O}\left(\frac{T^2-4D}{f^2}\right)^*\right|} h\left(\mathcal{O}\left(\frac{T^2-4D}{f^2}\right)\right).$$

Now we use a theorem which equates the counting of weighted $\Gamma(1)$-orbits of matrices of a fixed trace and determinant (of negative discriminant) with the counting of weighted ideal classes in the imaginary quadratic order of the same discriminant. We denote by $Q^*(\Delta)$ the set of primitive integral binary quadratic forms of discriminant Δ (for a definition, see [4]) and by $Q_+^*(\Delta)$ the subset of positive definite forms, both acted on by the classical $\Gamma(1)$-action

$$f \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix}(x, y) = f(ax + by, cx + dy).$$

By $Q^*(\Delta)\,//\,\Gamma(1)$ and $Q_+^*(\Delta)\,//\,\Gamma(1)$ we denote the corresponding orbit spaces under this action.

Theorem 15. *Let T and D be integers and f a positive integer satisfying*

$$T^2 - 4D < 0 \quad \text{and} \quad \frac{T^2-4D}{f^2} \in \mathbb{Z}, \quad \frac{T^2-4D}{f^2} \equiv 0 \text{ or } 1 \mod 4.$$

Then there are set bijections

$$\mathcal{T}^*(T, D, f)\,//\,\Gamma(1) \longleftrightarrow Q^*\left(\frac{T^2-4D}{f^2}\right)\,//\,\Gamma(1)$$

and

$$Q_+^*\left(\frac{T^2-4D}{f^2}\right)\,//\,\Gamma(1) \longleftrightarrow \mathcal{C}\left(\mathcal{O}\left(\frac{T^2-4D}{f^2}\right)\right).$$

Proof. We first observe that whenever $\mathcal{T}^*(T, D, f) \neq \emptyset$ (which is equivalent to the second two given conditions), there are unique integers T', D' and $\lambda \in \{0, 1, \ldots, f-1\}$ such that

$$\mathcal{T}^*(T, D, f) = \lambda I + f\mathcal{T}^*(T', D', 1).$$

Since $T^2 - 4D = f^2((T')^2 - 4D')$, the first bijection in the theorem is induced by the bijection

$$\mathcal{T}^*(T', D', 1) \longleftrightarrow Q^*\left((T')^2 - 4D'\right)$$

given by sending the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ to the form $cx^2+(d-a)xy-by^2$ and the form $\alpha x^2+\beta xy+\gamma y^2$ to the matrix $\begin{pmatrix} (T-\beta)/2 & -\gamma \\ \alpha & (T-\beta)/2 \end{pmatrix}$. The second bijection is classical (see e.g. [4, Theorem 7.7]). □

We observe that for any matrix $\alpha \in \mathcal{T}^*(T,D,f)$, we have

$$|\Gamma(1)_\alpha| = \left| \mathcal{O}\left(\frac{T^2-4D}{f^2}\right)^* \right|,$$

and the common value can be greater than 2 only when $\frac{T^2-4D}{f^2} \in \{-3,-4\}$, in which case $h\left(\mathcal{O}\left(\frac{T^2-4D}{f^2}\right)\right) = 1$. We conclude the following:

Corollary 16.

$$\frac{2}{\left|\mathcal{O}\left(\frac{T^2-4D}{f^2}\right)^*\right|} h\left(\mathcal{O}\left(\frac{T^2-4D}{f^2}\right)\right) = \sum_{\alpha \in (\lambda I + M\cdot\mathcal{T}^*(T,D,f))\,//\,\Gamma(1)} \frac{1}{|\Gamma(1)_\alpha|}.$$

From the corollary, (13) follows and we have proved Proposition 9, from which Theorem 8 follows.

4. The definition of a Serre curve

We now describe the subgroup H_E mentioned in Definition 3, following the proof of [20, Proposition 22]. Suppose that E is given by the equation

$$y^2 = x^3 + rx + s = (x-e_1)(x-e_2)(x-e_3).$$

Then $\{e_1,e_2,e_3\}$ is the set of x-coordinates of the nontrivial 2-torsion of E. The discriminant Δ of this model of E is given by

$$\Delta = ((e_1-e_2)(e_1-e_3)(e_2-e_3))^2. \tag{14}$$

Thus, one has

$$\mathbb{Q}(\sqrt{\Delta}) \subset \mathbb{Q}(E[2]).$$

Because of the action of $\operatorname{Aut} E[2] \simeq GL_2(\mathbb{Z}/2\mathbb{Z})$ on the e_i's, we have a group isomorphism between $GL_2(\mathbb{Z}/2\mathbb{Z})$ and the symmetric group on three letters:

$$GL_2(\mathbb{Z}/2\mathbb{Z}) \simeq S_3.$$

By (14) we see that for any Galois automorphism $\tau \in \operatorname{Gal}(\mathbb{Q}(E[2])/\mathbb{Q}) \subset S_3$,

$$\tau : \sqrt{\Delta} \mapsto \varepsilon(\tau)\sqrt{\Delta}, \tag{15}$$

where ε denotes the signature character on S_3. If $\sqrt{\Delta} \in \mathbb{Q}$, then

$$\operatorname{Gal}(\mathbb{Q}(E[2])/\mathbb{Q}) \subset A_3 = \text{ the alternating group on 3 letters.}$$

In this case, we define the **Serre number** M_1 and the **Serre subgroup** H_{M_1} by

$$M_1 := 2 \quad \text{and} \quad H_{M_1} := A_3 \subset GL_2(\mathbb{Z}/2\mathbb{Z}).$$

Suppose now that $\mathbb{Q}(\sqrt{\Delta}) \neq \mathbb{Q}$ is a quadratic extension, which in particular is abelian. Since each abelian extension of $\mathbb{Q}$ is contained in a cyclotomic extension, one may choose a positive integer D so that

$$\mathbb{Q}(\sqrt{\Delta}) \subset \mathbb{Q}(\zeta_D) \subset \mathbb{Q}(E[D]),$$

where as usual ζ_D denotes a primitive D-th root of unity and the second containment comes from the Weil pairing (see [23, III §8], for example).

Lemma 17. *Let W be any square-free integer and define the positive integer D_W by*

$$D_W = \begin{cases} |W| & \text{if } W \equiv 1 \mod 4 \\ 4|W| & \text{otherwise.} \end{cases}$$

Then we have

$$\mathbb{Q}(\sqrt{W}) \subset \mathbb{Q}(\zeta_D) \Leftrightarrow D_W \text{ divides } D.$$

Furthermore, for such a D and $\tau \in \mathrm{Gal}\,(\mathbb{Q}(E[D])/\mathbb{Q}) \subseteq GL_2(\mathbb{Z}/D\mathbb{Z})$, we have

$$\tau : \sqrt{W} \mapsto \left(\frac{W}{\det \tau}\right)\sqrt{W}. \tag{16}$$

Here we use the Kronecker symbol $\left(\frac{W}{\cdot}\right)$.

Proof. These are standard results from algebraic number theory. The assertion (16) follows from [22, Proposition 6.3, p. 135]. □

By the lemma we see that

$$\mathbb{Q}(\sqrt{\Delta}) \subset \mathbb{Q}(\zeta_D) \Leftrightarrow D_{\Delta_{sf}} \text{ divides } D, \tag{17}$$

where $\Delta_{sf} = \Delta_{sf}(E)$ is the square-free part of the discriminant Δ of E. For any square-free integer W we define the **Serre number**

$$M_W = \begin{cases} 2|W| & \text{if } W \equiv 1 \mod 4 \\ 4|W| & \text{otherwise,} \end{cases}$$

to be the least common multiple of 2 and D_W. Thus, in particular, $\mathbb{Q}(E[M_{\Delta_{sf}}])$ is the compositum of $\mathbb{Q}(E[2])$ and $\mathbb{Q}(E[D_{\Delta_{sf}}])$. We furthermore define the **Serre subgroup** H_{M_W} by

$$H_{M_W} = \ker\left(\left(\frac{W}{\det(\cdot)}\right)\varepsilon(\cdot)\right) \subset GL_2(\mathbb{Z}/M_W\mathbb{Z}),$$

where here we have extended the definition of the signature character ε in the natural way to any even level:

$$\varepsilon : GL_2(\mathbb{Z}/2m\mathbb{Z}) \longrightarrow GL_2(\mathbb{Z}/2\mathbb{Z}) \longrightarrow \{\pm 1\}. \tag{18}$$

Later in the paper we will casually refer to "$\ker \varepsilon$", hoping that in each instance its domain will be clear from context.

By virtue of (15) and (16), we see that

$$\mathrm{Gal}\,(\mathbb{Q}(E[M_{\Delta_{sf}}]/\mathbb{Q}) \subseteq H_{M_{\Delta_{sf}}}.$$

In either case ($\sqrt{\Delta} \in \mathbb{Q}$ or $\sqrt{\Delta} \notin \mathbb{Q}$), the subgroup H_E of $GL_2(\hat{\mathbb{Z}})$ referred to in (5) is simply

$$H_E = \pi^{-1}_{M_{\Delta_{sf}}}(H_{M_{\Delta_{sf}}}),$$

where $\pi_{M_{\Delta_{sf}}} : GL_2(\hat{\mathbb{Z}}) \longrightarrow GL_2(\mathbb{Z}/M_{\Delta_{sf}}\mathbb{Z})$ is the natural projection. H_E is evidently an index 2 subgroup of $GL_2(\hat{\mathbb{Z}})$ and

$$\phi_E(G_{\mathbb{Q}}) \subseteq H_E.$$

An elliptic curve E is a Serre curve if $\phi_E(G_{\mathbb{Q}}) = H_E$. In other words, an elliptic curve is a Serre curve exactly when, for every integer m, we have

$$[GL_2(\mathbb{Z}/m\mathbb{Z}) : \phi_{m,E}(G_{\mathbb{Q}})] = \begin{cases} 2 & \text{if } M_{\Delta_{sf}(E)} \mid m \\ 1 & \text{otherwise.} \end{cases}$$

We will refer to $H_{M_{\Delta_{sf}(E)}} \subset GL_2(\mathbb{Z}/M_{\Delta_{sf}(E)}\mathbb{Z})$ (and by abuse of notation, also to $H_E \subset GL_2(\hat{\mathbb{Z}})$) as the **Serre subgroup** associated to E.

Note that if $\sqrt{\Delta_E} \in \mathbb{Q}$, one may replace the field $\mathbb{Q}(\sqrt{\Delta_E})$ in the preceding argument with the abelian extension $\mathbb{Q}(E[2])$ and conclude that E is *not* a Serre curve in this case.

5. A criterion for distinguishing Serre curves

If N is exceptional for E (see Definition 1), then so is any multiple of N.

Definition 18. We call a positive integer N **minimal exceptional** for E if it is exceptional for E and none of its proper divisors are exceptional for E.

For example, if E is a Serre curve, then the Serre number $M_{\Delta_{sf}(E)}$ (see Section 4) is a minimal exceptional number for E. Also, any exceptional prime p of E is minimal exceptional.

The proof of Lemma 5 uses only the theory of the groups $GL_2(\mathbb{Z}/N\mathbb{Z})$ (especially for N divisible by 2 and 3, complementing [19]) as well as a few facts about cyclotomic fields. The arguments are similar to those given in Kani's appendix to [2]. Two separate issues arise: (1) which numbers N can actually occur as minimal exceptional numbers for an elliptic curve and (2) the stability of the Serre number $M_{\Delta_{sf}(E)}$. We treat them in that order.

5.1. **Lemmas from group theory.** In this section we state and prove several technical lemmas needed for the proof of Lemma 5. First, we will need

Lemma 19. *The commutator subgroup $(GL_2(\mathbb{Z}/p^n\mathbb{Z}))'$ of $GL_2(\mathbb{Z}/p^n\mathbb{Z})$ is given by*

$$(GL_2(\mathbb{Z}/p^n\mathbb{Z}))' = \begin{cases} SL_2(\mathbb{Z}/p^n\mathbb{Z}) & \text{if } p \neq 2 \\ \ker(\varepsilon) \cap SL_2(\mathbb{Z}/2^n\mathbb{Z}) & \text{if } p = 2 \end{cases}$$

(see (18)). For $p \geq 5$, the group $SL_2(\mathbb{Z}/p^n\mathbb{Z})$ is equal to its own commutator:

$$(SL_2(\mathbb{Z}/p^n\mathbb{Z}))' = SL_2(\mathbb{Z}/p^n\mathbb{Z}) \qquad (p \geq 5).$$

Proof. First, since

$$\left\langle \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \right\rangle \quad = \quad SL_2(\mathbb{Z}),$$

we see that for any $x \in (\mathbb{Z}/p^n\mathbb{Z})^*$,

$$\left\langle \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix} \right\rangle \quad = \quad SL_2(\mathbb{Z}/p^n\mathbb{Z}).$$

Denoting by 2^* any inverse of 2 modulo p^n, we compute

$$\begin{pmatrix} 2 & 0 \\ 0 & 2^* \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2^* & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \equiv \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} \mod p^n.$$

We see that

$$\begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} \in (SL_2(\mathbb{Z}/p^n\mathbb{Z}))',$$

and similarly for $\begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix}$, which proves the last assertion in the lemma. If $p = 3$, we compute

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \tag{19}$$

and similarly for $\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$, which proves that

$$p \geq 3 \Longrightarrow (GL_2(\mathbb{Z}/p^n\mathbb{Z}))' = SL_2(\mathbb{Z}/p^n\mathbb{Z}).$$

Finally, if $p = 2$, the equation

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix},$$

taken together with (19), implies that

$$(GL_2(\mathbb{Z}/2^n\mathbb{Z}))' \sqcup \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} (GL_2(\mathbb{Z}/2^n\mathbb{Z}))' = SL_2(\mathbb{Z}/2^n\mathbb{Z}).$$

(Note that $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \notin (GL_2(\mathbb{Z}/2^n\mathbb{Z}))'$ since $\varepsilon\left(\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\right) = -1$). We are then finished, since clearly

$$(GL_2(\mathbb{Z}/2^n\mathbb{Z}))' \subseteq \ker(\varepsilon) \cap SL_2(\mathbb{Z}/2^n\mathbb{Z})$$

and the indices inside $SL_2(\mathbb{Z}/2^n\mathbb{Z})$ match. □

We will also use

Lemma 20. *If N_1 and N_2 are relatively prime positive integers, then the groups $GL_2(\mathbb{Z}/N_1\mathbb{Z})$ and $GL_2(\mathbb{Z}/N_2\mathbb{Z})$ have no common simple nonabelian quotient, and neither do the groups $SL_2(\mathbb{Z}/N_1\mathbb{Z})$ and $SL_2(\mathbb{Z}/N_2\mathbb{Z})$.*

Proof. Any simple nonabelian quotient of a group occurs as a factor in its Jordan-Hölder series. Whenever

$$1 \to G_1 \to G \to G_2 \to 1$$

is an exact sequence of abelian groups, we have

$$\{ \text{ Jordan-Hölder factors of } G \} = \bigcup_{i=1,2} \{ \text{ Jordan-Hölder factors of } G_i \}.$$

Applying this observation to the exact sequences

$$1 \to GL_2(\mathbb{Z}/p_i^{n_i}\mathbb{Z}) \to GL_2(\mathbb{Z}/N_j\mathbb{Z}) \to GL_2(\mathbb{Z}/(N_j/(p_i^{n_i}))\mathbb{Z}) \to 1$$

(where $j = 1, 2$ and $N_j =: \prod_i p_i^{n_i}$),

$$1 \to I + p^{n-1}M_{2\times 2}(\mathbb{Z}/p\mathbb{Z}) \to GL_2(\mathbb{Z}/p^n\mathbb{Z}) \to GL_2(\mathbb{Z}/p^{n-1}\mathbb{Z}) \to 1,$$
$$1 \to SL_2(\mathbb{Z}/p\mathbb{Z}) \to GL_2(\mathbb{Z}/p\mathbb{Z}) \to (\mathbb{Z}/p\mathbb{Z})^* \to 1,$$

and

$$1 \to \{\pm I\} \to SL_2(\mathbb{Z}/p\mathbb{Z}) \to PSL_2(\mathbb{Z}/p\mathbb{Z}) \to 1,$$

and using the fact that

$$PSL_2(\mathbb{Z}/p\mathbb{Z}) \text{ is } \begin{cases} \text{simple} & \text{if } p \geq 5 \\ \text{solvable} & \text{otherwise} \end{cases}$$

and that $I + p^{n-1}M_{2\times 2}(\mathbb{Z}/p\mathbb{Z}) \subset GL_2(\mathbb{Z}/p^n\mathbb{Z})$ is an abelian subgroup ($n \geq 2$), we see that

$$\{ \text{ simple nonabelian quotients of } GL_2(\mathbb{Z}/N_j\mathbb{Z}) \} \subseteq \{PSL_2(\mathbb{Z}/p\mathbb{Z})\}_{p|N_j, p\geq 5}$$

(and likewise with $SL_2(\mathbb{Z}/N_j\mathbb{Z})$), finishing the proof. □

Finally, we will have need of

Lemma 21. *Let $N > 1$ be any even integer which is divisible by some prime $p > 3$. Write*

$$N = N_1 \cdot N_2$$

where $N_1 > 1$ is not divisible by any prime $p > 3$ and $N_2 > 1$ is not divisible by any prime $p \leq 3$. Suppose that $G_a \subset GL_2(\mathbb{Z}/N\mathbb{Z})$ is a subgroup such that

$$G_a \cap SL_2(\mathbb{Z}/N\mathbb{Z}) = (GL_2(\mathbb{Z}/N\mathbb{Z}))'.$$

Finally, assume $G_b \subset G_a$ is a subgroup for which the canonical maps

$$G_b \twoheadrightarrow GL_2(\mathbb{Z}/N_1\mathbb{Z}) \quad \text{and} \quad G_b \twoheadrightarrow GL_2(\mathbb{Z}/N_2\mathbb{Z}) \tag{20}$$

as well as the determinant map

$$\det : G_b \twoheadrightarrow (\mathbb{Z}/N\mathbb{Z})^*$$

are surjections. Then $G_b = G_a$.

Proof. By (20), we find by taking commutators that

$$G_b' \twoheadrightarrow (GL_2(\mathbb{Z}/N_1\mathbb{Z}))' = \ker(\varepsilon) \cap SL_2(\mathbb{Z}/N_1\mathbb{Z})$$

and

$$G_b' \twoheadrightarrow (GL_2(\mathbb{Z}/N_2\mathbb{Z}))' = SL_2(\mathbb{Z}/N_2\mathbb{Z})$$

are also surjections. We are now in a position to apply the Goursat Lemma:

Lemma 22. *Let G_1 and G_2 be groups. Denote by $\pi_i : G_1 \times G_2 \to G_i$ $(i = 1, 2)$ the projection map. Suppose that $G \subseteq G_1 \times G_2$ is a subgroup such that $\pi_i(G) = G_i$ for $i = 1, 2$ and define*

$$H_1 = \pi_1(G \cap (G_1 \times \{e_2\})) \quad \text{and} \quad H_2 = \pi_2(G \cap (\{e_1\} \times G_2)).$$

Then,

$$G_1/H_1 \simeq G_2/H_2$$

and the graph of this isomorphism is induced by G.

We apply the lemma with $G_1 = \ker(\varepsilon) \cap SL_2(\mathbb{Z}/N_1\mathbb{Z})$, $G_2 = SL_2(\mathbb{Z}/N_2\mathbb{Z})$, and $G = G_b'$ and conclude that $\ker(\varepsilon) \cap SL_2(\mathbb{Z}/N_1\mathbb{Z})$ and $SL_2(\mathbb{Z}/N_2\mathbb{Z})$ have a common quotient group Q. If Q is nontrivial, then it has a nontrivial simple quotient Q_s. By Lemma 20, Q_s must be abelian. By Lemma 19, $(SL_2(\mathbb{Z}/N_2\mathbb{Z}))' = SL_2(\mathbb{Z}/N_2\mathbb{Z})$, and so we must have $Q_s = 1$. This shows that Q was trivial to begin with. We conclude that $G_1 = H_1$ and $G_2 = H_2$, i.e. that

$$(GL_2(\mathbb{Z}/N_1\mathbb{Z}))' \times \{1\} \subset G_b' \quad \text{and} \quad \{1\} \times (GL_2(\mathbb{Z}/N_2\mathbb{Z}))' \subset G_b',$$

which implies that

$$G_b' = (GL_2(\mathbb{Z}/N\mathbb{Z}))'.$$

But now from the exact sequence

$$1 \to (GL_2(\mathbb{Z}/N\mathbb{Z}))' \to G_a \to (\mathbb{Z}/N\mathbb{Z})^* \to 1$$

and

$$\det : G_b \twoheadrightarrow (\mathbb{Z}/N\mathbb{Z})^*$$

we conclude that $(GL_2(\mathbb{Z}/N\mathbb{Z}))'G_b = G_a$. So since $(GL_2(\mathbb{Z}/N\mathbb{Z}))' \subset G_b$, we have $G_b = G_a$. □

5.2. **Minimal exceptional numbers of elliptic curves.** The following lemma gives us a restriction on which positive integers N can occur as a minimal exceptional number of an elliptic curve. Throughout the remainder of the paper we will sometimes use the abbreviation

$$G_N := \mathrm{Gal}\left(\mathbb{Q}(E[N])/\mathbb{Q}\right),$$

suppressing the dependence on the elliptic curve E.

Lemma 23. *Let E be an elliptic curve over $\mathbb{Q}$. Suppose that $N \in \mathbb{N}$ is minimal exceptional for E. Then,*

$$N \in \{ \text{ prime numbers } \} \cup \{M_{\Delta_{sf}(E)}\} \cup \{4, 8, 9\}.$$

If 8 is a minimal exceptional number for E, then there exists a real primitive character $\delta : (\mathbb{Z}/8\mathbb{Z})^* \to \{\pm 1\}$ and

$$G_8(E) = \ker\left(\varepsilon \cdot (\delta \circ \det)\right).$$

Proof. Let us assume that N is not prime. If N is exceptional for E, then we have

$$G_N \subsetneq GL_2(\mathbb{Z}/N\mathbb{Z}).$$

If N is *minimal* exceptional, we have $G_d = GL_2(\mathbb{Z}/d\mathbb{Z})$ for each proper divisor d of N. Therefore the canonical map

$$G_N \twoheadrightarrow GL_2(\mathbb{Z}/d\mathbb{Z}) \tag{21}$$

is a surjection for each d dividing N. By the surjectivity of the Weil pairing, we also see that the determinant map

$$\det : G_N \twoheadrightarrow (\mathbb{Z}/N\mathbb{Z})^* \tag{22}$$

is surjective. We consider the question: for which composite numbers N does there exist a proper subgroup G_N of $GL_2(\mathbb{Z}/N\mathbb{Z})$ satisfying conditions (21) and (22)? We divide the investigation into cases according to whether N is a prime power or not. We tackle the latter case first.

Case 1. N is not a prime power. Let p be the smallest prime divisor of N. Suppose that $p^n \,||\, N$ (i.e. that $p^n \mid N$ and $p^{n+1} \nmid N$) and write $M := N/p^n$ ($\neq 1$). By Galois theory we must have

$$\mathbb{Q} \subsetneq \mathbb{Q}(E[p^n]) \cap \mathbb{Q}(E[M]).$$

Let $F := \mathbb{Q}(E[p^n]) \cap \mathbb{Q}(E[M])$ and $H := \mathrm{Gal}\left(F/\mathbb{Q}\right)$. If H is not simple, replace it by any nontrivial simple quotient, and replace F by the corresponding field. Since H is a common simple quotient of the groups

$$\mathrm{Gal}\left(\mathbb{Q}(E[M])/\mathbb{Q}\right) = GL_2(\mathbb{Z}/M\mathbb{Z}) \quad \text{and} \quad \mathrm{Gal}\left(\mathbb{Q}(E[p^n])/\mathbb{Q}\right) = GL_2(\mathbb{Z}/p^n\mathbb{Z}),$$

we conclude by Lemma 20 that H is abelian. From this and Lemma 19 it follows that

$$F \subset \mathbb{Q}(\zeta_M).$$

If $p > 2$, then we must similarly have $F \subset \mathbb{Q}(\zeta_{p^n})$. Since

$$\mathbb{Q}(\zeta_M) \cap \mathbb{Q}(\zeta_{p^n}) = \mathbb{Q}, \tag{23}$$

we conclude that $F = \mathbb{Q}$, contradicting the assumption that H is nontrivial. Therefore we must have $p = 2$. But then using Lemma 19 we similarly conclude that

$$\mathbb{Q} \neq F \subset \mathbb{Q}(\sqrt{\Delta_E}, \zeta_{2^n}) \cap \mathbb{Q}(\zeta_M).$$

If $n \leq 1$, then we must have $F = \mathbb{Q}(\sqrt{\Delta_E})$, and we see that N is a multiple of the Serre number $M_{\Delta_{sf}(E)}$. If $n \geq 2$, then we reason as follows: since the Galois group $\mathrm{Gal}(\mathbb{Q}(\sqrt{\Delta_E}, \zeta_{2^n})/\mathbb{Q})$ has order a power of two, F must be a quadratic field. By (23), we conclude that if $n = 2$, then F must be one of the fields

$$\mathbb{Q}(\sqrt{\Delta_E}),\ \mathbb{Q}(\sqrt{-\Delta_E}),$$

and if $n \geq 3$, then F must be one of the fields

$$\mathbb{Q}(\sqrt{\Delta_E}),\ \mathbb{Q}(\sqrt{-\Delta_E}),\ \mathbb{Q}(\sqrt{2\Delta_E}),\ \mathbb{Q}(\sqrt{-2\Delta_E}).$$

Thus in any case, by (17), N is a multiple of the Serre number of E, which implies that N *is* the Serre number of E, since N is assumed to be minimal exceptional. We have shown that the Serre number of E is the only possible minimal exceptional number which is not a prime power.

Case 2. $N = p^n$ is a prime power. If p is odd, then we reason as follows. Suppose, for the sake of contradiction, that

$$n \geq \begin{cases} 2 & \text{if } p \geq 5 \\ 3 & \text{if } p = 3. \end{cases}$$

Taking commutators of (21), we have a surjection

$$(G_{p^n}(E))' \twoheadrightarrow SL_2(\mathbb{Z}/p^{n-1}\mathbb{Z}) = (GL_2(\mathbb{Z}/p^{n-1}\mathbb{Z}))'.$$

By [19, Lemma 3, p. IV-23] and [19, Exercise 1, p. IV-27], this implies that $(G_{p^n}(E))' = SL_2(\mathbb{Z}/p^n\mathbb{Z})$. But now since

$$SL_2(\mathbb{Z}/p^n\mathbb{Z}) \subset G_{p^n}(E)$$

we conclude by (22) that $G_{p^n}(E) = GL_2(\mathbb{Z}/p^n\mathbb{Z})$, contradicting the fact that p^n is exceptional. Thus, the only composite odd prime power which could possibly occur as a minimal exceptional number is 9.

Suppose now that $N = 2^n$ $(n \geq 2)$ and consider the exact sequence

$$1 \to K \cap G_{2^n} \to G_{2^n} \to GL_2(\mathbb{Z}/2^{n-1}\mathbb{Z}) \to 1,$$

where $K = I + 2^{n-1}M_{2\times 2}(\mathbb{Z}/2\mathbb{Z})$. First we show that if $n \geq 3$, then

$$I + 2^{n-1}\{A \in M_{2\times 2}(\mathbb{Z}/2\mathbb{Z}) : \operatorname{tr} A = 0\} \subseteq K \cap G_{2^n}. \tag{24}$$

This is seen by choosing any preimage

$$\begin{pmatrix} 1 & 2^{n-2} \\ 0 & 1 \end{pmatrix} + 2^{n-1}A \in G_{2^n}$$

of the matrix $\begin{pmatrix} 1 & 2^{n-2} \\ 0 & 1 \end{pmatrix} \in GL_2(\mathbb{Z}/2^{n-1}\mathbb{Z})$ and observing that, if $n \geq 3$,

$$\left(\begin{pmatrix} 1 & 2^{n-2} \\ 0 & 1 \end{pmatrix} + 2^{n-1}A\right)^2 \equiv \begin{pmatrix} 1 & 2^{n-1} \\ 0 & 1 \end{pmatrix} \mod 2^n,$$

which shows that the matrix $I + 2^{n-1}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in K \cap G_{2^n}$. Now let $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be any matrix in $GL_2(\mathbb{Z}/2\mathbb{Z})$ and choose a matrix $A \in G_{2^n}(E)$ with

$$A \equiv \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mod 2.$$

We then have

$$A\left(I + 2^{n-1}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right)A^{-1} = I + \frac{1}{ad-bc}2^{n-1}\begin{pmatrix} -ac & a^2 \\ -c^2 & ac \end{pmatrix} \in K \cap G_{2^n}(E).$$

Letting the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ vary modulo 2, we see that (24) holds. From this we see that G_{2^n} must be an index 2 subgroup of $GL_2(\mathbb{Z}/2^n\mathbb{Z})$. Thus, there is a character

$$\chi : GL_2(\mathbb{Z}/2^n\mathbb{Z}) \to \{\pm 1\} \quad \text{with} \quad G_{2^n}(E) = \ker \chi. \tag{25}$$

Lemma 19 says that either χ or $\varepsilon \cdot \chi$ restricted to $SL_2(\mathbb{Z}/2^n\mathbb{Z})$ must be trivial. But if χ is trivial on $SL_2(\mathbb{Z}/2^n\mathbb{Z})$, then $SL_2(\mathbb{Z}/2^n\mathbb{Z}) \subseteq G_{2^n}$, a contradiction. Thus we must have

$$\chi = \varepsilon \cdot (\delta \circ \det), \tag{26}$$

where $\delta : (\mathbb{Z}/2^n\mathbb{Z})^* \to \{\pm 1\}$ is a *primitive* character (or else 2^n is not *minimal* exceptional). Now pick $X = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} + 2^{n-1}A \in G_{2^n}$. We have $\det X = 1$ or $1 + 2^{n-1}$. One checks that for $n \geq 3$,

$$1 + 2^{n-1} \equiv 5^{2^{n-3}} \mod 2^n,$$

so for $n > 3$ we must have $\delta(\det X) = 1$, contradicting (25). Thus, the only composite powers of 2 which could possibly occur as minimal exceptional numbers are 4 and 8. This concludes the proof of Lemma 23. □

5.3. **Stability of the Serre number $M_{\Delta_{sf}(E)}$.** We will now finish the proof of Lemma 5 by showing that under the assumptions stated therein and for each positive integer N, we have

$$G_N(E) = \begin{cases} \pi^{-1}_{N, M_{\Delta_{sf}}}(H_{M_{\Delta_{sf}}}) & \text{if } M_{\Delta_{sf}} \mid N \\ GL_2(\mathbb{Z}/N\mathbb{Z}) & \text{otherwise,} \end{cases} \tag{27}$$

where $\pi_{N, M_{\Delta_{sf}}}$ denotes the natural projection

$$GL_2(\mathbb{Z}/N\mathbb{Z}) \longrightarrow GL_2(\mathbb{Z}/M_{\Delta_{sf}}\mathbb{Z}).$$

To see this, first suppose that $M_{\Delta_{sf}} \nmid N$. If $G_N(E) \subsetneq GL_2(\mathbb{Z}/N\mathbb{Z})$, then E has some minimal exceptional number d dividing N. Clearly d cannot be equal to the Serre number $M_{\Delta_{sf}}$, so again by Lemma 23 and the assumptions on E from Lemma 5 we arrive at a contradiction. Thus, if $M_{\Delta_{sf}} \nmid N$ we have

$$G_N(E) = GL_2(\mathbb{Z}/N\mathbb{Z}).$$

Now suppose $M_{\Delta_{sf}} \mid N$. We will apply Lemma 21 with $G_a = \pi^{-1}_{N, M_{\Delta_{sf}}}(H_{M_{\Delta_{sf}}})$ and $G_b = G_N(E)$. To verify the hypotheses of the lemma, note that under the assumptions stated in Lemma 5, we have that

$$\exists\, p > 3 \text{ which divides } M_{\Delta_{sf}(E)}. \tag{28}$$

Thus we may write $N = N_1 \cdot N_2$ as in Lemma 21. The condition

$$\pi_{N,M_{\Delta_{sf}}}^{-1}(H_{M_{\Delta_{sf}}}) \cap SL_2(\mathbb{Z}/N\mathbb{Z}) = \ker \varepsilon \cap SL_2(\mathbb{Z}/N\mathbb{Z})$$

follows immediately from the definition of $H_{M_{\Delta_{sf}}}$. We next verify the surjectivity conditions

$$G_N(E) \twoheadrightarrow GL_2(\mathbb{Z}/N_1\mathbb{Z}) \text{ and } G_N(E) \twoheadrightarrow GL_2(\mathbb{Z}/N_2\mathbb{Z}). \tag{29}$$

If the first map is not surjective, then E has some minimal exceptional number d which divides N_1. By Lemma 23, we conclude that $d \in \{2, 3, 4, 6, 8, 9, 12, 24\}$, contradicting the assumptions on E in Lemma 5. Therefore $G_N(E) \twoheadrightarrow GL_2(\mathbb{Z}/N_1\mathbb{Z})$ is surjective. Similarly, if $G_N(E) \to GL_2(\mathbb{Z}/N_2\mathbb{Z})$ is not surjective, then E has some minimal exceptional d dividing N_2. By Lemma 23, we must have that d is an odd prime number greater than 3, again contradicting the assumptions of Lemma 5. We have verified the conditions (29). Finally, the surjectivity of

$$\det : G_N \twoheadrightarrow (\mathbb{Z}/N\mathbb{Z})^*$$

follows from the surjectivity of the Weil pairing. By Lemma 21, we conclude that

$$G_N(E) = \pi_{N,M_{\Delta_{sf}}}^{-1}(H_{M_{\Delta_{sf}}}),$$

and our proof of Lemma 5 is now complete.

6. Almost all elliptic curves are Serre curves

We now show how Lemma 5 and Theorem 2 together imply Theorem 4. For $N \in \{4, 6, 8, 9, 12, 24\}$ define

$$\varepsilon_N(X) := \begin{cases} \{E \in C(X) : E \text{ is minimal exceptional at } N\} & \text{if } N \in \{4, 9\} \\ \{E \in C(X) : G_8(E) = \ker \chi, \text{ for } \chi \text{ as in (26) }\} & \text{if } N = 8 \\ \{E \in C(X) : G_N(E) \subseteq H_N\} & \text{if } N \in \{6, 12, 24\}. \end{cases}$$

By Lemmas 5 and 23, we have that the set of non-Serre curves satisfies

$$C(x) - C_{\text{Serre}}(x) \subseteq \varepsilon(x) \cup \left(\bigcup_{N \in \{4,6,8,9,12,24\}} \varepsilon_N(x) \right),$$

where $\varepsilon(x)$ is as in (3). Thus, to prove Theorem 4 it suffices to estimate the sets $\varepsilon_N(x)$.

Definition 24. Let W be any integer and let $(t, d) \in (\mathbb{Z}/W\mathbb{Z})^2$ be any pair of integers modulo W with $d \in (\mathbb{Z}/W\mathbb{Z})^*$. Suppose that $G \subseteq GL_2(\mathbb{Z}/W\mathbb{Z})$ is any subgroup. We say that G **represents the pair** (t, d) if there is a matrix $g \in G$ satisfying

$$\operatorname{tr}(g) = t, \quad \det(g) = d.$$

The next two lemmas guarantee that when an elliptic curve fails to be a Serre curve by being minimal exceptional at N, there must be some pair (t, d) not represented by $G_N(E)$.

Lemma 25. *Let $W > 1$ be any even integer,*

$$\delta : (\mathbb{Z}/W\mathbb{Z})^* \to \{\pm 1\}$$

any nonprincipal real character and

$$G \subseteq \ker(\varepsilon \cdot (\delta \circ \det)) \subseteq GL_2(\mathbb{Z}/W\mathbb{Z})$$

any subgroup. Then there exists a pair of integers $(t,d) \in \mathbb{Z}/W\mathbb{Z} \times (\mathbb{Z}/W\mathbb{Z})^*$ *which is not represented by* G.

Proof. Choose any $f \in (\mathbb{Z}/W\mathbb{Z})^*$ satisfying

$$\delta(f) = -1$$

and set $(t,d) = (1,f)$. □

Lemma 26. *Let* $p = 2$ *or* 3 *and suppose* $G \subseteq GL_2(\mathbb{Z}/p^2\mathbb{Z})$ *is a subgroup which represents every trace-determinant pair* $(t,d) \in (\mathbb{Z}/p^2\mathbb{Z}) \times (\mathbb{Z}/p^2\mathbb{Z})^*$ *and which surjects onto* $GL_2(\mathbb{Z}/p\mathbb{Z})$. *Then,* $G = GL_2(\mathbb{Z}/p^2\mathbb{Z})$.

Proof. We consider the intersection

$$G \cap K$$

of G with K, the kernel of the projection

$$GL_2(\mathbb{Z}/p^2\mathbb{Z}) \longrightarrow GL_2(\mathbb{Z}/p\mathbb{Z}).$$

Our goal is to show that G actually contains K. From here we divide the argument into cases, according to whether p is 2 or 3.

Case $p = 3$. Under the given hypothesis, we may find a matrix $g \in G$ with $\operatorname{tr} g = 3$ and $\det g = 1$. Such a matrix must have the form

$$X + 3Y, \quad X \in \left\{ \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 2 & 2 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 2 \\ 2 & 1 \end{pmatrix} \right\},$$

with the (mod 3) coefficients of the matrix $Y = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ satisfying the conditions

$$\begin{aligned}
a + d = 1,\ b - c = 1 \text{ if } X &= \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix} \\
a + d = 1,\ b - c = 2 \text{ if } X &= \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} \\
a + d = 0,\ a + b + c - d = 0 \text{ if } X &= \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \\
a + d = 0,\ a - b - c - d = 2 \text{ if } X &= \begin{pmatrix} 1 & 2 \\ 2 & 2 \end{pmatrix} \\
a + d = 0,\ a - b - c - d = 0 \text{ if } X &= \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \\
a + d = 0,\ a + b + c - d = 1 \text{ if } X &= \begin{pmatrix} 2 & 2 \\ 2 & 1 \end{pmatrix}.
\end{aligned}$$

In each case, the first equation comes from the trace condition on g and the second one comes from the determinant condition. One computes that

$$(X + 3Y)^4 \equiv I + 3X \mod 9.$$

Since in this case the discriminant $t^2 - 4d = 5$ is nonzero modulo 3 we see by Lemma 11 that all six of the matrices X, when reduced modulo 3, are $GL_2(\mathbb{Z}/3\mathbb{Z})$-conjugate to one another. Proceeding as in the argument which showed (24) and using the

fact that the various X span the $\mathbb{Z}/3\mathbb{Z}$-vector space $M_{2\times 2}(\mathbb{Z}/3\mathbb{Z})$, we conclude that

$$G \cap K = I + 3M_{2\times 2}(\mathbb{Z}/3\mathbb{Z}).$$

Thus we have $K \subseteq G$, and so $G = GL_2(\mathbb{Z}/9\mathbb{Z})$ in this case.

Case $p = 2$. The proof in this case is similar. Pick $g \in G$ with $\operatorname{tr} g = 2$ and $\det g = -1$. Then g must have the form

$$g = X + 2Y, \quad X \in \left\{ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \right\}, \quad Y = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

where the (mod 2) coefficients of the matrix Y satisfy the conditions

$$a + d = 1,\ b + c = 0 \text{ if } X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$a + d = 0,\ a + c + d = 1 \text{ if } X = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$a + d = 0,\ a + b + d = 1 \text{ if } X = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

(The possibility $X = I + 2Y$ is eliminated since the conditions on the coefficients of Y in that case read $a + d = 0$, $a + d = 1$.) One computes that

$$(X + 2Y)^2 \equiv I + 2X \mod 4.$$

After conjugating by preimages of elements of $GL_2(\mathbb{Z}/2\mathbb{Z})$, one concludes that

$$G \cap K \supseteq \left\{ I + 2 \cdot \text{ span } \left\{ \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \right\} \right\}.$$

Proceeding in the same way with $t = 0$ and $d = 1$, one sees that in fact

$$G \cap K \supseteq \{A \in M_{2\times 2}(\mathbb{Z}/2\mathbb{Z}) : \operatorname{tr} A = 0\}.$$

Now G is a subgroup of $GL_2(\mathbb{Z}/4\mathbb{Z})$ of index ≤ 2. However, if G is indeed a subgroup of index 2, we may apply Lemma 25 and arrive at a contradiction. This concludes the proof in this case. □

Lemmas 25 and 26 imply the following corollary.

Corollary 27. *For* $N \in \{4, 6, 8, 9, 12, 24\}$, *we have*

$$\varepsilon_N(x) = \bigcup_{(t,d) \in (\mathbb{Z}/N\mathbb{Z}) \times (\mathbb{Z}/N\mathbb{Z})^*} \varepsilon_{N,(t,d)}(x),$$

where

$$\varepsilon_{N,(t,d)} := \{E \in \varepsilon_N(x) : \ (t, d) \text{ is not represented by } G_N(E)\}.$$

Lemma 28. *For each* $N \in \{4, 6, 8, 9, 12, 24\}$, *we have*

$$|\varepsilon_N(x)| \ll N^8 x |C(x)| \max_d \pi(x; N, d)^{-2},$$

with an absolute implied constant.

Proof. This lemma and its proof are analogous to [8, Lemma 5], the statement of which contains a typo: the "$\ll X^6 \pi(X; N, d)^{-2}$" should be replaced by "$\ll N^4 X^6 \max_d \pi(X; N, d)^{-2}$". Theorem 2 implies that

$$\left(\frac{|GL_2(\mathbb{Z}/N\mathbb{Z})_{t,d}| \varphi(N)}{|GL_2(\mathbb{Z}/N\mathbb{Z})|} \right)^2 \pi(x; N, d)^2 |\varepsilon_{N,(t,d)}(x)| \ll |GL_2(\mathbb{Z}/N\mathbb{Z})_{t,d}|^2 x |C(x)|,$$

where $GL_2(\mathbb{Z}/N\mathbb{Z})_{t,d} := \{g \in GL_2(\mathbb{Z}/N\mathbb{Z}) : \operatorname{tr} g = t, \det g = d\}$. Summing over (t, d) proves the lemma. □

By the prime number theorem in arithmetic progressions, we see that Lemma 28, together with (6), implies Theorem 4. We note that because the N in Lemma 28 belongs to a finite set, there is no need here to use the Siegel-Walfisz theorem, as was necessary in [8]. However, since Theorem 4 depends on [8, Theorem 1], which does use the Siegel-Walfisz theorem, the constant implied by the $\ll$ symbol in Theorem 4 is ineffective.

7. $N = 4$ occurs as a minimal exceptional number

If $N = 4$ or 9, the argument given in Section 5.2 is invalid since we may not conclude that (24) holds. In fact, there is a subgroup $H \subset GL_2(\mathbb{Z}/4\mathbb{Z})$ of index four which satisfies conditions (21) and (22). We now describe H and demonstrate an infinite family of non-isomorphic elliptic curves E for which $G_4(E) = H$. Elkies [10] has recently exhibited similar examples for $N = 9$.

First, we give a geometric description of H: Let L be a complex lattice and let $L[4]$ denote the 4-torsion of $\mathbb{C}/L$. By choosing a basis, we may identify $L[4]$ with $(\mathbb{Z}/4\mathbb{Z})^2$. Define

$$L[4]^* := \{x \in L[4] - L[2]\}$$

and let $l_1, l_2 \ldots, l_6$ denote the lines through the origin in $L[4]^*$. More precisely, define the equivalence relation on $L[4]^*$ by declaring $u \sim u'$ exactly if $u' = \lambda u$ for some $\lambda \in (\mathbb{Z}/4\mathbb{Z})^* = \{\pm 1\}$, and denote the resulting equivalence classes by $l_1, l_2 \ldots, l_6$. Since the Weierstrass $\wp$-function is even, the association $l_i = [u] \mapsto \wp(u)$ identifies $\mathbb{P}^1_{\mathbb{Z}}(\mathbb{Z}/4\mathbb{Z}) := \{l_1, l_2, \ldots, l_6\}$ with

$$E[4]^*_x := \{x(P) : P \in E[4] - E[2]\},$$

the set of x-coordinates of the set of 4-torsion points of $E = E_L$, the elliptic curve associated to the lattice L, which are not 2-torsion points. This identification allows one to view the Galois group of $\mathbb{Q}(E[4]_x)$ over $\mathbb{Q}$ as a subgroup of $PGL_2(\mathbb{Z}/4\mathbb{Z})$.

We may extend the natural action of $PGL_2(\mathbb{Z}/4\mathbb{Z})$ on $\mathbb{P}^1_{\mathbb{Z}}(\mathbb{Z}/4\mathbb{Z})$ to obtain a $PGL_2(\mathbb{Z}/4\mathbb{Z})$ action on the set

$$S := \{\{\{l_{i_1}, l_{i_2}, l_{i_3}\}, \{l_{i_4}, l_{i_5}, l_{i_6}\}\} : \text{ all } i_j \in \{1, 2, \ldots, 6\} \text{ are distinct}\}.$$

This action is not transitive. The size 10 set S decomposes into two orbits S_1 and S_2 of sizes 4 and 6, respectively. To describe these sets, we will define an "addition relation" on $\mathbb{P}^1_{\mathbb{Z}}(\mathbb{Z}/4\mathbb{Z})$. If l_1, l_2, and l_3 are lines in $\mathbb{P}^1_{\mathbb{Z}}(\mathbb{Z}/4\mathbb{Z})$, we say that

$$l_1 + l_2 = l_3$$

exactly when, for some choice of representatives $u_i \in l_i$, we have

$$u_1 + u_2 = u_3.$$

(Note: This is a *relation*, not a well-defined operation. For example, $[(1, 0)] + [(0, 1)] = [(1, 1)]$ and $[(1, -1)]$.) Then the two orbits are defined by

$$S_1 := \{\{\{l_{i_1}, l_{i_2}, l_{i_3}\}, \{l_{i_4}, l_{i_5}, l_{i_6}\}\} \in S : l_{i_1} + l_{i_2} = l_{i_3}\}$$

and $S_2 = S - S_1$. Fixing any element $r \in S_1$, we define $H_x = H_x(r) \subset PGL_2(\mathbb{Z}/4\mathbb{Z})$ to be the stabilizer of r. Finally, we define $H = H(r)$ to be the preimage of H_x under the natural projection $GL_2(\mathbb{Z}/4\mathbb{Z}) \to PGL_2(\mathbb{Z}/4\mathbb{Z})$.

To find elliptic curves E with $G_4(E) = H$, we reason as follows: let $x_1, x_2, \ldots, x_6$ be the elements of $E[4]^*_x$. If E is given in the form

$$E: \quad y^2 = 4x^3 - g_2x - g_3,$$

then the minimal polynomial for $x_1, x_2, \ldots, x_6$ is given by

$$f_E(t) = t^6 - \frac{5g_2}{4}t^4 - 5g_3t^3 - \frac{5g_2^2}{16}t^2 - \frac{g_2g_3}{4}t + \frac{g_2^3 - 32g_3^2}{64}.$$

The set S_1 defined above corresponds to the set of numbers

$$X_1 := \{(x_{i_1} + x_{i_2} + x_{i_3})(x_{i_4} + x_{i_5} + x_{i_6}) : (x_{i_1}, y_{i_1}) \oplus (x_{i_2}, y_{i_2}) = (x_{i_3}, y_{i_3})\},$$

where $\oplus$ refers to the addition law on E. X_1 is a set of four complex numbers which satisfy the (generically irreducible) polynomial

$$f_{1,E}(t) = t^4 + 3g_2t^3 + \frac{27g_2^2}{8}t^2 + (\frac{-37g_2^3}{16} + 108g_3^2)t + \frac{81g_2^4}{256}.$$

We note that $G_4(E) \subseteq$ some $H(r)$ whenever $f_{1,E}(t)$ has a linear factor over $\mathbb{Q}$. Let $s \in \mathbb{Q}$ and denote by E_s the elliptic curve given by the equation

$$y^2 := 4x^3 + \frac{16s^2 + 56s + 81}{3s}x + \frac{(16s^2 + 56s + 81)^2(-1 + 4s)}{864s^2}.$$

It may be checked that for each s, $f_{1,E_s}(t)$ is divisible by $t + 27 + \frac{56}{3}s + \frac{16}{3}s^2$ and that

$$\mathrm{Gal}(\mathbb{Q}(s)(E_s[4])/\mathbb{Q}(s)) \simeq H.$$

The discriminant is computed to be

$$\Delta(E_s) = -\frac{(16s^2 + 56s + 81)^3(4s + 3)^4}{27648s^4},$$

and the j-invariant is

$$j(E_s) = \frac{1769472s}{(4s + 3)^4}.$$

In particular, if we apply the Hilbert irreducibility criterion, we see that there are infinitely many non-isomorphic curves E_s over $\mathbb{Q}$, each with Galois group $G_4 \simeq H$.

Acknowledgment

This paper contains results of the author's Ph.D. dissertation. The author is grateful to his advisor, William Duke, for his guidance.

References

[1] A. Brumer, *The average rank of elliptic curves I*, Invent. Math. **109** (1992), 445–472. MR1176198 (93g:11057)

[2] A. C. Cojocaru, *On the surjectivity of the Galois representations associated to non-CM elliptic curves*, with an appendix by Ernst Kani, Canad. Math. Bull. **48** (2005), no. 1, 16–31. MR2118760 (2005k:11109)

[3] A. C. Cojocaru and C. Hall, *Uniform results for Serre's theorem for elliptic curves*, Int. Math. Res. Not. **50** (2005), 3065–3080. MR2189500 (2006g:11107)

[4] D. A. Cox, *Primes of the Form $x^2 + ny^2$*, Wiley, New York, 1989. MR1028322 (90m:11016)

[5] H. Davenport, *Multiplicative Number Theory*, Springer, New York-Berlin, 1980. MR606931 (82m:10001)

[6] C. David and F. Pappalardi, *Average Frobenius distributions of elliptic curves*, Int. Math. Res. Not. **4** (1999), 165–183. MR1677267 (2000g:11045)

[7] M. Deuring, *Die typen der Multiplikationenringe der elliptischen Funktionenkörper*, Abh. Math. Sem. Univ. Hamburg **14** (1941), 197–272. MR0005125 (3:104f)

[8] W. D. Duke, *Elliptic curves with no exceptional primes*, C. R. Math. Acad. Sci. Paris Sér. I **325** (1997), 813–818. MR1485897 (99b:11059)
[9] W. Duke and A. Tóth, *The splitting of primes in division fields of elliptic curves*, Experiment. Math. **11** (2003), 555–565. MR1969646 (2004c:11087)
[10] N. Elkies, *Elliptic curves with* 3*-adic Galois representation surjective mod* 3 *but not mod* 9, preprint (2006). Available at `http://arxiv.org/abs/math/0612734`
[11] P. X. Gallagher, *The large sieve inequality and probabilistic Galois theory*, in: Analytic Number Theory (St. Louis Univ., St. Louis, 1972), Proc. Sympos. Pure Math., Vol. IV, 91–101. Amer. Math. Soc., Providence, 1973. MR0332694 (48:11020)
[12] D. Grant, *A formula for the number of elliptic curves with exceptional primes*, Compos. Math. **122** (2000), 151–164. MR1775416 (2001j:11033)
[13] N. Jones, *Trace formulas and class number sums*, Acta Arith. **132** (2008), no. 4, 301–313. MR2413354 (2009g:11149)
[14] N. Jones, *Averages of elliptic curve constants*, to appear in Mathematische Annalen.
[15] A. Kraus, *Une remarque sur les points de torsion des courbes elliptiques*, C. R. Acad. Sci. Paris, Sér. I, **321** (1995), 1143–1146. MR1360773 (97a:11085)
[16] D. Masser and G. Wüstholz, *Galois properties of division fields of elliptic curves*, Bull. London Math. Soc. **25** (1993), 247–254. MR1209248 (94d:11036)
[17] B. Mazur, *Rational isogenies of prime degree*, Invent. Math. **44** (1978), 129–162. MR482230 (80h:14022)
[18] R. Schoof, *Nonsingular plane cubic curves over finite fields*, J. Combin. Theory Ser. A **46** (1986), 183–211. MR914657 (88k:14013)
[19] J. P. Serre, *Abelian l-Adic Representations and Elliptic Curves*, Benjamin, New York-Amsterdam, 1968. MR0263823 (41:8422)
[20] ———, *Propriétés galoisiennes des points d'ordre fini des courbes elliptiques*, Invent. Math. **15** (1972), 259–331. MR0387283 (52:8126)
[21] ———, *Quelques applications du théorème de densité de Chebotarev*, Inst. Hautes Études Sci. Publ. Math. **54** (1981), 123–201
[22] G. Shimura, *Introduction to the Arithmetic Theory of Automorphic Functions*, Princeton University Press, Princeton, 1971. MR0314766 (47:3318)
[23] J. H. Silverman, *The Arithmetic of Elliptic Curves*, Springer-Verlag, New York, 1986. MR817210 (87g:11070)
[24] ———, *Advanced Topics in the Arithmetic of Elliptic Curves*, Springer-Verlag, New York, 1994. MR1312368 (96b:11074)
[25] E. Waterhouse, *Abelian varieties over finite fields*, Ann. Sci. École Norm. Sup. **2** (1969), 521–560. MR0265369 (42:279)

Centre de Recherches Mathématiques, Université de Montréal, P.O. Box 6128, Centreville Station, Montréal, Québec, Canada H3C 3J7

Current address: Department of Mathematics, University of Mississippi, Hume Hall 305, P.O. Box 1848, University, Mississippi 33677-1848

E-mail address: `ncjones@olemiss.edu`