Math. Ann. (2009) 345:685–710
DOI 10.1007/s00208-009-0373-1

Mathematische Annalen

Averages of elliptic curve constants

Nathan Jones

Received: 27 November 2007 / Revised: 28 January 2009 / Published online: 7 May 2009

Abstract We compute the averages over elliptic curves of the constants occurring in the Lang–Trotter conjecture, the Koblitz conjecture, and the cyclicity conjecture. The results obtained confirm the consistency of these conjectures with the corresponding "theorems on average" obtained recently by various authors.

1 Introduction

Let E be an elliptic curve defined over the rational numbers, by which we understand a Weierstrass equation of the form

$$y^2 = x^3 + ax + b \quad (a, b \in \mathbb{Q})$$

for which the discriminant $\Delta_E := -16(4a^3+27b^2)$ is non-zero. For a prime p of good reduction for E (in other words, a prime p for which a and b are integral modulo p and Δ_E is non-zero modulo p), let E_p denote the reduction of E modulo p, viewed as an elliptic curve over the finite field $\mathbb{Z}/p\mathbb{Z}$. There are various conjectured asymptotics for functions which count good primes p up to x for which the reduced curve E_p has certain properties. In this paper we will focus on three such questions, although our methods are applicable to a wider range of problems.

For a fixed integer r, let

$$\pi_{E,r}(x) = |\{p \leq x : p \text{ of good reduction for } E,\ a_p(E) = r\}|,$$

N. Jones (✉)
Montreal, Canada
e-mail: jones@dms.umontreal.ca

where $a_p(E) = p+1-|E_p(\mathbb{Z}/p\mathbb{Z})|$ is the trace of the Frobenius endomorphism of E at p. Lang and Trotter [14], using a probabilistic model consistent with the Chebotarev density theorem and the Sato–Tate conjecture, predicted an asymptotic for $\pi_{E,r}(x)$:

Conjecture 1 (Lang–Trotter) Assume that either E has no complex multiplication or that $r \neq 0$. Then

$$\pi_{E,r}(x) \sim C_{E,r} \cdot \frac{\sqrt{x}}{\log x} \quad \text{as} \quad x \to \infty,$$

where $C_{E,r}$ is a specific constant.

We will describe the constant $C_{E,r}$ in detail in Sect. 2. The second conjecture we will consider involves the counting function

$$\pi_{E,\text{prime}}(x) := |\{p \leq x : p \text{ of good reduction for } E,\ |E_p(\mathbb{Z}/p\mathbb{Z})| \text{ is prime}\}|.$$

Conjecture 2 (Koblitz)

$$\pi_{E,\text{prime}}(x) \sim C_{E,\text{prime}} \cdot \frac{x}{(\log x)^2} \quad \text{as} \quad x \to \infty,$$

where $C_{E,\text{prime}}$ is a specific constant.

Finally we will consider the cyclicity conjecture, which has been settled conditionally by Serre [18] and unconditionally in the CM case by Murty [15] and also by Cojocaru [5]. Let

$$\pi_{E,\text{cyclic}}(x) := |\{p \leq x : p \text{ of good reduction for } E,\ E_p(\mathbb{Z}/p\mathbb{Z}) \text{ is a cyclic group}\}|.$$

Conjecture 3 (Cyclicity conjecture)

$$\pi_{E,\text{cyclic}}(x) \sim C_{E,\text{cyclic}} \cdot \frac{x}{\log x} \quad \text{as} \quad x \to \infty,$$

where $C_{E,\text{cyclic}}$ is a specific constant.

We remark that, in case the constant happens to vanish, i.e. if

$$C_{E,r} = 0 \quad \left(\text{resp. } C_{E,\text{prime}} = 0 \ \text{ or } \ C_{E,\text{cyclic}} = 0\right),$$

then each conjectured asymptotic is interpreted to mean that the corresponding counting functions are finite, i.e. that

$$\lim_{x\to\infty} \pi_{E,r}(x) < \infty \quad \left(\text{resp. } \lim_{x\to\infty} \pi_{E,\text{prime}}(x) < \infty \ \text{ or } \ \lim_{x\to\infty} \pi_{E,\text{cyclic}}(x) < \infty\right).$$

Recently, various authors have proven that these conjectures “hold on average over elliptic curves.” More precisely, for non-negative parameters A and B, let $\mathcal{F} =$

$\mathcal{F}(A, B)$ denote the set of elliptic curves $Y^2 = X^3 + aX + b$ with $(a, b) \in ([-A, A] \times [-B, B]) \cap \mathbb{Z}^2$. When $A = A(x)$ and $B = B(x)$ depend on an auxiliary parameter x, we will employ the abbreviation

$$\mathcal{F}(x) := \mathcal{F}(A(x), B(x)).$$

Fouvry and Murty [9] (in case $r = 0$) and later David and Pappalardi [7] (in case $r \neq 0$) proved Conjecture 1 on average: for any $\varepsilon > 0$, if $\min\{A(x), B(x)\} \geq x^{1+\varepsilon}$ then

$$\frac{1}{|\mathcal{F}(x)|} \sum_{E \in \mathcal{F}(x)} \pi_{E,r}(x) \sim C_r \cdot \frac{\sqrt{x}}{\log x}, \quad \text{as } x \to \infty, \tag{1}$$

where C_r is a specific constant. Baier [2] has recently shortened the length of the average of David and Pappalardi, replacing "$x^{1+\varepsilon}$" with "$x^{3/4+\varepsilon}$" (which is what Fouvry and Murty had obtained in the $r = 0$ case).

Balog et al. [3] have proved a similar average theorem for Conjecture 2: for any $\varepsilon > 0$, if $\min\{A(x), B(x)\} \geq x^{1+\varepsilon}$ then

$$\frac{1}{|\mathcal{F}(x)|} \sum_{E \in \mathcal{F}(x)} \pi_{E,\mathrm{prime}}(x) \sim C_{\mathrm{prime}} \cdot \frac{x}{(\log x)^2}, \quad \text{as } x \to \infty, \tag{2}$$

where C_{prime} is a specific constant.

Finally, Banks and Shparlinski [4] have proved Conjecture 3 unconditionally on average: for any $\varepsilon > 0$, if $x^{1/2+\varepsilon} \leq A(x), B(x) \leq x^{1-\varepsilon}$ then

$$\frac{1}{|\mathcal{F}(x)|} \sum_{E \in \mathcal{F}(x)} \pi_{E,\mathrm{cyclic}}(x) \sim C_{\mathrm{cyclic}} \cdot \frac{x}{\log x}, \quad \text{as } x \to \infty, \tag{3}$$

where C_{cyclic} is a specific constant.

In this paper we will prove that each of these average results is consistent with the corresponding conjectured result on the level of the constants. We will make the notation uniform.

Notation 4 *Throughout the rest of this paper, let C_E denote any one of the constants $C_{E,r}$, $C_{E,prime}$, or $C_{E,cyclic}$, and let C denote the corresponding average constant C_r, C_{prime}, or C_{cyclic}. In the case where E has CM and $r = 0$, we formally extend the definition of the Lang–Trotter constant by setting*

$$C_{E,0} := C_0 = \frac{\pi}{3}.$$

Our first theorem is conditional upon an affirmative answer to the following question of Serre [16, p. 299]. In its statement, $\mathbb{Q}(E[p])$ denotes the pth division field of E, i.e. the field obtained by adjoining to $\mathbb{Q}$ the x and y-coordinates of the p-torsion points of a given Weierstrass model of E.

Question 5 *Does there exist an absolute constant c so that, for any prime $p \geq c$ and any elliptic curve E over $\mathbb{Q}$ without complex multiplication, one has*

$$\mathit{Gal}\,(\mathbb{Q}(E[p])/\mathbb{Q}) \simeq GL_2(\mathbb{Z}/p\mathbb{Z})?$$

We remark that, in [17, p. 199], Serre asks in particular whether one can take $c = 41$. In the present paper, we prove the following theorem.

Theorem 6 *Assume that Question 5 has an affirmative answer. Then there exists an exponent $\gamma > 0$ such that, for any positive integer k, we have*

$$\frac{1}{|\mathcal{F}(A,B)|}\sum_{E\in\mathcal{F}(A,B)} |C_E - C|^k \ll_k \max\left\{ \left(\frac{\log B\cdot(\log A)^7}{B}\right)^{\frac{k}{k+1}}, \frac{\log^\gamma(\min\{A,B\})}{\sqrt{\min\{A,B\}}} \right\}.$$

In the case of the Lang–Trotter conjecture, the implied constant also depends on r.

Note that, since there is no prime-counting function involved in our theorem, there is no need to regard A and B as depending on an auxiliary variable x.

Remark 7 As we will see in Proposition 15 below, there is always a positive proportion of elliptic curves E over $\mathbb{Q}$ for which $C_E \neq C$. As studied in [10,20], for number fields $K \neq \mathbb{Q}$ with $K \cap \mathbb{Q}^{\text{cyc}} = \mathbb{Q}$, for "almost all" elliptic curves E over K one has $C_E = C$ in these problems. Thus, this is a setting in which the situation over $\mathbb{Q}$ is more delicate than that over number fields.

The following corollary of Theorem 6 partially addresses a question in Sect. 2 of [7].

Corollary 8 *Suppose that $A(x)$ and $B(x)$ both tend to infinity with x. Then, provided that Question 5 has an affirmative answer and that*

$$\lim_{x\to\infty} \frac{\log B(x)\cdot(\log A(x))^7}{B(x)} = 0,$$

we have

$$\frac{1}{|\mathcal{F}(x)|}\sum_{E\in\mathcal{F}(x)} C_E \longrightarrow C \quad \mathit{as} \quad x\to\infty.$$

Taking $k = 2$, we also note the following corollary to Theorem 4 of [2] (see also Theorem 1.4 of [7]), which bounds the mean-square error in the Lang–Trotter conjecture.

Corollary 9 *Let $\varepsilon > 0$ and $c > 0$ be given and suppose that Question 5 has an affirmative answer. Then, provided that $x^{1+\varepsilon} < \min\{A(x), B(x)\}$, $x^{3+\varepsilon} < A(x)\cdot B(x) < \exp(\exp(\sqrt{x}/(\log x)^c))$, and that*

$$\left(\frac{\log B(x)\cdot(\log A(x))^7}{B(x)}\right)^{2/3} \ll \frac{1}{(\log x)^{c-2}},$$

one has

$$\frac{1}{|\mathcal{F}(x)|}\sum_{E\in\mathcal{F}(x)}\left|\pi_{E,r}(x)-C_{E,r}\frac{\sqrt{x}}{\log x}\right|^2 \ll \frac{x}{(\log x)^c}.$$

(The difference between this corollary and Theorem 4 of [2] is that we have replaced the average constant C_r with $C_{E,r}$).

Unconditionally, we prove a statement about averages over Serre curves (we will review the notion of a Serre curve in Sect. 3).

Theorem 10 *For any positive integer k, one has*

$$\frac{1}{|\mathcal{F}(A,B))|}\sum_{\substack{E\in\mathcal{F}(A,B)\\ E\ \textit{is a Serre curve}}}|C_E-C|^k \ll_k \frac{1}{A}+\left(\frac{\log B\cdot(\log A)^7}{B}\right)^{\frac{k}{k+1}}.$$

Because of the fact that

$$\frac{1}{|\mathcal{F}(A,B)|}\sum_{\substack{E\in\mathcal{F}(A,B)\\ E\text{ is a Serre curve}}}1 \longrightarrow 1$$

as $\min\{A,B\}\to\infty$ (c.f. [12]), Theorem 10 provides evidence that Theorem 6 should hold unconditionally.

2 The constants

In this section we will describe precisely the constants occuring in the conjectures under consideration, as well as the corresponding average constants. Their description involves the division fields of E, whose notation we now fix. Recall that we are always implicitly viewing E as a Weierstrass equation over $\mathbb{Q}$.

Notation 11 *For each positive integer n, denote by* $\mathbb{Q}(E[n])$ *the nth division field of E, obtained by adjoining to* $\mathbb{Q}$ *the x and y-coordinates of the n-torsion of E, and by*

$$G_n(E) := Gal\,(\mathbb{Q}(E[n])/\mathbb{Q})$$

the associated Galois group. Since $E[n]$ *is a free* $\mathbb{Z}/n\mathbb{Z}$*-module of rank* 2*, we may (by fixing a* $\mathbb{Z}/n\mathbb{Z}$*-basis) view* $G_n(E)$ *as a subgroup of* $GL_2(\mathbb{Z}/n\mathbb{Z})$*.*

We will distinguish between the case where E has complex multiplication (CM) and the case where E does not (non-CM). Since almost all elliptic curves are non-CM [8], our only interest in the CM case is to obtain upper bounds for C_E.

2.1 The non-CM case and the average constants

In the non-CM case, the constant C_E has the form

$$C_E = f(m_E, G_{m_E}(E)) \cdot \prod_{\ell \nmid m_E} f(\ell, GL_2(\mathbb{Z}/\ell\mathbb{Z})),$$

where $f(n, G)$ is some function of the level n and the subgroup $G \leq GL_2(\mathbb{Z}/n\mathbb{Z})$, and where m_E is a positive integer depending on the torsion representation attached to E. We begin by describing m_E. Another way to phrase Notation 11 is to say that there is a group homomorphism

$$\varphi_{E,n} : G_{\mathbb{Q}} \to GL_2(\mathbb{Z}/n\mathbb{Z}),$$

defined by letting the absolute Galois group $G_{\mathbb{Q}} := \text{Gal}\,(\overline{\mathbb{Q}}/\mathbb{Q})$ act on the n-torsion points of E, and we are denoting the image of $\varphi_{E,n}$ by $G_n(E)$. Taking the inverse limit of the $\varphi_{E,n}$ over positive integers n (with bases chosen compatibly), one obtains a continuous group homomorphism

$$\varphi_E : G_{\mathbb{Q}} \to GL_2(\hat{\mathbb{Z}}).$$

(Here $\hat{\mathbb{Z}} := \varprojlim \mathbb{Z}/n\mathbb{Z} = \prod_p \mathbb{Z}_p$.) Serre [16] proved the following "open image" theorem.

Theorem 12 *Suppose that E is an elliptic curve over $\mathbb{Q}$ which has no complex multiplication. Then, with the notation as above, we have*

$$[GL_2(\hat{\mathbb{Z}}) : \varphi_E(G_{\mathbb{Q}})] < \infty.$$

In other words, there exists a positive integer level m so that, if

$$\pi : GL_2(\hat{\mathbb{Z}}) \to GL_2(\mathbb{Z}/m\mathbb{Z})$$

is the natural projection, one has

$$\varphi_E(G_{\mathbb{Q}}) = \pi^{-1}(G_m(E)). \tag{4}$$

For a non-CM curve E over $\mathbb{Q}$, let us denote by m_E the smallest positive integer m such that the above equation holds. In particular, m_E has the property that, for m_1 dividing m_E and m_2 coprime to m_E one has

$$G_{m_1 m_2}(E) \simeq G_{m_1}(E) \times GL_2(\mathbb{Z}/m_2\mathbb{Z}). \tag{5}$$

In order to write the Lang–Trotter constant $C_{E,r}$, we follow the notation in [14]: for $G \subseteq GL_2(\mathbb{Z}/n\mathbb{Z})$ any subgroup, let

$$G_r := \{g \in G : \text{tr}\, g \equiv r \mod n\}.$$

Then the specific constant of Conjecture 1 predicted by Lang and Trotter is

$$\begin{aligned} C_{E,r} &= \frac{2}{\pi} \cdot \frac{m_E |G_{m_E}(E)_r|}{|G_{m_E}(E)|} \cdot \prod_{\ell \nmid m_E} \frac{\ell |GL_2(\mathbb{Z}/\ell\mathbb{Z})_r|}{|GL_2(\mathbb{Z}/\ell\mathbb{Z})|} \\ &= \frac{2}{\pi} \cdot \frac{m_E |G_{m_E}(E)_r|}{|G_{m_E}(E)|} \cdot \prod_{\substack{\ell | r \\ \ell \nmid m_E}} \left(1 + \frac{1}{\ell^2 - 1}\right) \cdot \prod_{\substack{\ell \nmid r \\ \ell \nmid m_E}} \left(1 - \frac{1}{(\ell - 1)(\ell^2 - 1)}\right), \end{aligned}$$

where m_E is as in (4). On the other hand, the average constant in (1) is

$$C_r = \frac{2}{\pi} \cdot \prod_{\ell | r} \left(1 + \frac{1}{\ell^2 - 1}\right) \cdot \prod_{\ell \nmid r} \left(1 - \frac{1}{(\ell - 1)(\ell^2 - 1)}\right).$$

To write the Koblitz constant [13] (as refined by Zywina in [19]), we define, for any positive integer n, the subset

$$\Phi_n := \{g \in GL_2(\mathbb{Z}/n\mathbb{Z}) : \det(1 - g) \in (\mathbb{Z}/n\mathbb{Z})^*\}. \tag{6}$$

Then the constant of Conjecture 2 predicted by Koblitz is

$$\begin{aligned} C_{E,\text{prime}} &= \frac{|G_{m_E}(E) \cap \Phi_{m_E}| / |G_{m_E}(E)|}{\prod_{\ell | m_E} (1 - 1/\ell)} \cdot \prod_{\ell \nmid m_E} \frac{|GL_2(\mathbb{Z}/\ell\mathbb{Z}) \cap \Phi_\ell| / |GL_2(\mathbb{Z}/\ell\mathbb{Z})|}{(1 - 1/\ell)} \\ &= \frac{|G_{m_E}(E) \cap \Phi_{m_E}| / |G_{m_E}(E)|}{\prod_{\ell | m_E} (1 - 1/\ell)} \cdot \prod_{\ell \nmid m_E} \left(1 - \frac{\ell^2 - \ell - 1}{(\ell - 1)^3(\ell + 1)}\right). \end{aligned}$$

In this case the average constant in (2) is given by

$$C_{\text{prime}} = \prod_{\ell} \left(1 - \frac{\ell^2 - \ell - 1}{(\ell - 1)^3(\ell + 1)}\right).$$

Finally, the cyclicity constant in Conjecture 3 is given by

$$C_{E,\text{cyclic}} = \sum_{n \geq 1} \frac{\mu(n)}{[\mathbb{Q}(E[n]) : \mathbb{Q}]},$$

where

$$\mu(n) := \begin{cases} (-1)^{|\{p \text{ prime } : p | n\}|} & \text{if } n \text{ is square-free} \\ 0 & \text{otherwise} \end{cases}$$

is the Möbius function. In the non-CM case, this may be expressed as

$$C_{E,\text{cyclic}} = \left(\sum_{n \mid m_E} \frac{\mu(n)}{|G_n(E)|} \right) \cdot \prod_{\ell \nmid m_E} \left(1 - \frac{1}{\ell(\ell-1)^2(\ell+1)} \right),$$

because of (5) and the fact that any square-free integer n may be decomposed as $n = n_1 \cdot n_2$, where $n_1 \mid m_E$ and $(n_2, m_E) = 1$. The average constant in (3) is

$$C_{\text{cyclic}} = \prod_{\ell} \left(1 - \frac{1}{\ell(\ell-1)^2(\ell+1)} \right).$$

Note that if any non-CM elliptic curve E were to satisfy $m_E = 1$ (i.e. if φ_E were surjective), then we would have $C_E = C$. However, as observed by Serre, *no* elliptic curve over $\mathbb{Q}$ has $m_E = 1$. The main difficulty in proving Theorem 6 is tracking the variation of m_E with E. To prove the theorem, we will focus on a density one subset of curves E for which m_E is essentially equal to the square-free part of the discriminant of E. These curves are called *Serre curves* and will be discussed in detail in Sect. 3.

2.2 The CM case

The Galois representation on the torsion of a CM elliptic curve may be viewed as a "one-dimensional" analogue of the non-CM situation. Suppose E is an elliptic curve over $\mathbb{Q}$ with CM by an order $\mathcal{O}$ in an imaginary quadratic field K. In this case, $\varphi_E(G_{\mathbb{Q}})$ is *not* an open subgroup of $GL_2(\hat{\mathbb{Z}})$. In fact, it has an index two subgroup which is abelian. Indeed, the torsion of E

$$E_{\text{tors}} := \bigcup_{n \geq 1} E[n]$$

is a one-dimensional $\hat{\mathcal{O}}$-module ($\hat{\mathcal{O}} := \varprojlim \mathcal{O}/n\mathcal{O}$) on which $G_K := \text{Gal}\,(\overline{\mathbb{Q}}/K)$ acts, preserving the $\hat{\mathcal{O}}$-action. Thus, the image of φ_E restricted to G_K maps into $(\hat{\mathcal{O}})^*$:

$$\varphi_E : G_K \longrightarrow (\hat{\mathcal{O}})^* = GL_1(\hat{\mathcal{O}}).$$

The following theorem (see [16, Sect. 4.5] and the references therein) is the classical CM analogue of Theorem 12.

Theorem 13 *Suppose E is an elliptic curve over $\mathbb{Q}$ with CM by an imaginary quadratic order $\mathcal{O}$. Then*

$$[(\hat{\mathcal{O}})^* : \varphi_E(G_K)] < \infty.$$

In other words, viewing each $\mathrm{Gal}\,(K(E[n])/K)$ as a subgroup of $(\mathcal{O}/n\mathcal{O})^*$, Theorem 13 states that there is a positive integer m with the property that, for each positive integer n, we have

$$\mathrm{Gal}\,(K(E[n])/K) \simeq \pi^{-1}(\mathrm{Gal}\,(K(E[\gcd(n, m)])/K)), \tag{7}$$

where $\pi : (\mathcal{O}/n\mathcal{O})^* \to (\mathcal{O}/\gcd(n, m)\mathcal{O})^*$ is the canonical projection. The condition (7) continues to hold if one replaces the integer m by any multiple. For a CM curve E over $\mathbb{Q}$, let us denote by m_E the smallest positive integer m such that (7) holds and (for notational convenience) for which

$$4 \cdot \left(\prod_{\ell \text{ ramifies in } \mathcal{O}} \ell \right) \text{ divides } m.$$

Note in particular that, for ℓ not dividing m_E, one has

$$\mathrm{Gal}\,(K(E[\ell])/K) \simeq (\mathcal{O}/\ell\mathcal{O})^*.$$

The constant occurring in Conjecture 1 in the CM case is

$$C_{E,r} = \frac{1}{2\pi} \cdot \frac{m_E|\mathrm{Gal}\,(K(E[m_E])/K)_r|}{|\mathrm{Gal}\,(K(E[m_E])/K)|} \cdot \prod_{\ell \nmid m_E} \frac{\ell|(\mathcal{O}/\ell\mathcal{O})_r^*|}{|(\mathcal{O}/\ell\mathcal{O}^*)|}.$$

Explicitly, we have that

$$C_{E,r} = \frac{m_E|\mathrm{Gal}\,(K(E[m_E])/K)_r|}{2\pi|\mathrm{Gal}\,(K(E[m_E])/K)|} \cdot \prod_{\substack{\ell \nmid m_E \\ \ell | r}} \frac{\ell}{\ell - \chi_{\mathcal{O}}(\ell)} \cdot \prod_{\substack{\ell \nmid m_E \\ \ell \nmid r}} \frac{\ell^2 - (1 + \chi_{\mathcal{O}}(\ell))\,\ell}{(\ell - 1)\,(\ell - \chi_{\mathcal{O}}(\ell))}, \tag{8}$$

where $\chi_{\mathcal{O}}(\ell)$ is the character determining the splitting of ℓ in the order $\mathcal{O}$, namely

$$\chi_{\mathcal{O}}(\ell) := \begin{cases} 1 & \text{if } \ell \text{ splits in } \mathcal{O} \\ -1 & \text{if } \ell \text{ is inert in } \mathcal{O}. \end{cases}$$

(Since E is defined over $\mathbb{Q}$, the class number of $\mathcal{O}$ must be one, and so "splitting of ℓ in $\mathcal{O}$" makes sense).

To describe the Koblitz constant in the CM case, first notice that, by fixing a $\mathbb{Z}/n\mathbb{Z}$-basis of $\mathcal{O}/n\mathcal{O}$, we may view $GL_1(\mathcal{O}/n\mathcal{O})$ as a subgroup of $GL_2(\mathbb{Z}/n\mathbb{Z})$, and doing so, we have

$$\Phi_n \cap (\mathcal{O}/n\mathcal{O})^* = \{g \in (\mathcal{O}/n\mathcal{O})^* : N(1 - g) \in (\mathbb{Z}/n\mathbb{Z})^*\},$$

where Φ_n is defined by (6) and $N : (\mathcal{O}/n\mathcal{O})^* \to (\mathbb{Z}/n\mathbb{Z})^*$ is the norm map. Then we have

$$C_{E,\text{prime}} := \frac{|\text{Gal}\,(K(E[m_E])/K) \cap \Phi_{m_E}|/|\text{Gal}\,(K(E[m_E])/K)|}{\prod_{\ell | m_E}(1-1/\ell)} \times \prod_{\ell \nmid m_E} \frac{|(\mathcal{O}/\ell\mathcal{O})^* \cap \Phi_\ell|/|(\mathcal{O}/\ell\mathcal{O})^*|}{1-1/\ell}.$$

Note that

$$C_{E,\text{prime}} = \frac{|\text{Gal}\,(K(E[m_E])/K) \cap \Phi_{m_E}|/|\text{Gal}\,(K(E[m_E])/K)|}{\prod_{\ell | m_E}(1-1/\ell)} \times \prod_{\ell \nmid m_E} \left(1 - \chi_{\mathcal{O}}(\ell)\frac{\ell^2-\ell-1}{(\ell-\chi_{\mathcal{O}}(\ell))(\ell-1)^2}\right). \tag{9}$$

Finally, we recall that the cyclicity constant is the same as in the non-CM case:

$$C_{E,\text{cyclic}} = \sum_{n \geq 1} \frac{\mu(n)}{[\mathbb{Q}(E[n]) : \mathbb{Q}]}.$$

2.3 Overview of the proof of Theorem 6

The proof of Theorem 6 will proceed as follows. We will decompose the sum

$$\sum_{E \in \mathcal{F}(A,B)} |C_E - C|^k = \sum_{\substack{E \in \mathcal{F}(A,B) \\ E \text{ is a Serre curve}}} |C_E - C|^k + \sum_{\substack{E \in \mathcal{F}(A,B) \\ E \text{ is not a Serre curve}}} |C_E - C|^k.$$

In Sect. 4, we will prove Theorem 10, which says that

$$\frac{1}{|\mathcal{F}(A,B)|} \sum_{\substack{E \in \mathcal{F}(A,B) \\ E \text{ is a Serre curve}}} |C_E - C|^k \ll_k \frac{1}{A} + \left(\frac{\log B \cdot (\log A)^7}{B}\right)^{k/(k+1)}.$$

In Sect. 5 we will show that, assuming an affirmative answer to Question 5, one has

$$\frac{1}{|\mathcal{F}(A,B)|} \sum_{\substack{E \in \mathcal{F}(A,B) \\ E \text{ is not a Serre curve}}} |C_E - C|^k \ll_k \frac{\log^\gamma(\min\{A,B\})}{\sqrt{\min\{A,B\}}},$$

concluding the proof of Theorem 6.

3 Serre curves

Serre [16, Sect. 5.5] observed that although the torsion representation φ_E has finite index in $GL_2(\hat{\mathbb{Z}})$, it is never surjective when the base field is $\mathbb{Q}$. For each elliptic curve E over $\mathbb{Q}$, there is an index two subgroup $H_E \subseteq GL_2(\hat{\mathbb{Z}})$ with $\varphi_E(G_{\mathbb{Q}}) \subseteq H_E$. (We will presently describe this subgoup).

Definition 14 An elliptic curve E over $\mathbb{Q}$ is a *Serre curve* if $\varphi_E(G_{\mathbb{Q}}) = H_E$.

Equivalently, an elliptic curve over $\mathbb{Q}$ is a Serre curve if $\varphi_E(G_{\mathbb{Q}})$ has index 2 in $GL_2(\hat{\mathbb{Z}})$. Thus, a Serre curve is an elliptic curve whose torsion representation has image which is "as large as possible."

We now describe the subgroup H_E, which will be the full preimage under the canonical surjection

$$\pi : GL_2(\hat{\mathbb{Z}}) \to GL_2(\mathbb{Z}/M_E\mathbb{Z})$$

of a particular index two subgroup of $GL_2(\mathbb{Z}/M_E\mathbb{Z})$ for a certain level M_E. Suppose that E is given by the Weierstrass equation $y^2 = (x - e_1)(x - e_2)(x - e_3)$. Then, the 2-torsion of E may be given explicitly as

$$E[2] = \{\mathcal{O}, (e_1, 0), (e_2, 0), (e_3, 0)\}.$$

By considering the action of $\mathrm{Aut}(E[2])$ on the 3-element set $E[2] - \{\mathcal{O}\}$, we see that $\mathrm{Aut}(E[2])$ is isomorphic to the symmetric group on 3 letters:

$$GL_2(\mathbb{Z}/2\mathbb{Z}) \simeq \mathrm{Aut}(E[2]) \simeq S_3. \tag{10}$$

Since

$$\Delta_E = [(e_1 - e_2)(e_2 - e_3)(e_1 - e_3)]^2,$$

we see that $\mathbb{Q}(\sqrt{\Delta_E}) \subseteq \mathbb{Q}(E[2])$ and that the action of $\sigma \in G_2(E)$ on $\sqrt{\Delta_E}$ is given by

$$\sigma\left(\sqrt{\Delta_E}\right) = \varepsilon(\sigma)\sqrt{\Delta_E}, \tag{11}$$

where $\varepsilon : GL_2(\mathbb{Z}/2\mathbb{Z}) \to \{\pm 1\}$ is the identification (10) followed by the signature character on S_3. On the other hand, the field $\mathbb{Q}(\sqrt{\Delta_E})$, being an abelian extension of $\mathbb{Q}$, is contained in a cyclotomic extension. Let D_E be the *conductor* of $\mathbb{Q}(\sqrt{\Delta_E})$, i.e. the smallest positive integer for which

$$\mathbb{Q}(\sqrt{\Delta_E}) \subseteq \mathbb{Q}(\zeta_{D_E})$$

(where ζ_{D_E} is a primitive D_Eth root of unity). In fact,

$$D_E = \begin{cases} |\Delta_{sf}(E)| & \text{if } \Delta_{sf}(E) \equiv 1 \mod 4 \\ 4|\Delta_{sf}(E)| & \text{otherwise,} \end{cases}$$

where $\Delta_{sf} = \Delta_{sf}(E)$ denotes the square-free part of the discriminant of E, i.e. the unique square-free integer so that

$$\frac{\Delta_E}{\Delta_{sf}(E)} \in (\mathbb{Q}^\times)^2.$$

Note that $\Delta_{sf}(E)$ only depends on $E/\mathbb{Q}$, and not on the particular Weierstrass model. Next we define

$$M_E = \begin{cases} 2|\Delta_{sf}(E)| & \text{if } \Delta_{sf}(E) \equiv 1 \mod 4 \\ 4|\Delta_{sf}(E)| & \text{otherwise,} \end{cases} \tag{12}$$

which is the least common multiple of 2 and D_E. The field $\mathbb{Q}(E[M_E])$ is the compositum of $\mathbb{Q}(E[2])$ and $\mathbb{Q}(E[D_E])$. Since

$$\mathbb{Q}(\sqrt{\Delta_E}) \subseteq \mathbb{Q}(E[2]) \cap \mathbb{Q}(E[D_E]),$$

the corresponding Galois group $G_{M_E}(E)$ must be a proper subgroup of $GL_2(\mathbb{Z}/M_E\mathbb{Z})$ (this is still true in case $D_E \in \{1, 4, 8\}$, but requires extra thought). In particular, considering the tower of fields

$$\mathbb{Q}(\sqrt{\Delta_E}) \subseteq \mathbb{Q}(\zeta_{D_E}) \subseteq \mathbb{Q}(E[M_E]),$$

we see that, for $\sigma \in G_{M_E}(E) \subseteq GL_2(\mathbb{Z}/M_E\mathbb{Z})$, one has

$$\sigma\left(\sqrt{\Delta_E}\right) = \left(\frac{\Delta_{sf}(E)}{\det\sigma}\right)\sqrt{\Delta_E}.$$

(Here we are using the Kronecker symbol

$$\left(\frac{\Delta_{sf}}{\cdot}\right) := \left(\frac{\Delta_{sf}/|\Delta_{sf}|}{\cdot}\right) \cdot \prod_{p|\Delta_{sf}} \left(\frac{p}{\cdot}\right),$$

where

$$\left(\frac{2}{\cdot}\right) := (-1)^{((\cdot)^2-1)/8}, \quad \left(\frac{\pm 1}{\cdot}\right) = (\pm 1)^{((\cdot)-1)/2}, \quad \text{and} \quad \left(\frac{p}{n}\right) := \prod_{\ell^k \| n} \left(\frac{p}{\ell}\right)^k,$$

$\left(\frac{p}{\ell}\right)$ being the usual Legendre symbol for the odd primes p and ℓ).

This observation, together with (11), shows that, with the notation as defined,

$$G_{M_E}(E) \subseteq \ker\left(\varepsilon(\cdot)\left(\frac{\Delta_{sf}(E)}{\det(\cdot)}\right)\right) \subseteq GL_2(\mathbb{Z}/M_E\mathbb{Z}).$$

In this case we therefore define $\overline{H_E} \subset GL_2(\mathbb{Z}/M_E\mathbb{Z})$ and $H_E \subset GL_2(\hat{\mathbb{Z}})$ by

$$\overline{H_E} := \ker\left(\varepsilon(\cdot)\left(\frac{\Delta_{sf}(E)}{\det(\cdot)}\right)\right) \quad \text{and} \quad H_E := \pi^{-1}(\overline{H_E}).$$

Note that in the degenerate case $\mathbb{Q}(\sqrt{\Delta_E}) = \mathbb{Q}$ (i.e. in case $\Delta_{sf}(E) = 1$) we have $M_E = 2$ and

$$\overline{H_E} = \ker \varepsilon =: A_3 = \text{ the alternating subgroup of } S_3.$$

For any elliptic curve E over $\mathbb{Q}$ we have

$$E \text{ is a Serre curve} \iff m_E = M_E \quad \text{and} \quad G_{M_E}(E) = \overline{H_E}.$$

One shows easily that, for d a proper divisor of M_E and π denoting the natural projection $GL_2(\mathbb{Z}/M_E\mathbb{Z}) \to GL_2(\mathbb{Z}/d\mathbb{Z})$, one has

$$\pi\left(\overline{H_E}\right) = GL_2(\mathbb{Z}/d\mathbb{Z}).$$

Thus in particular, when E is a Serre curve and $d \mid M_E$, one has

$$G_d(E) = \begin{cases} \overline{H_E} & \text{if } d = M_E \\ GL_2(\mathbb{Z}/d\mathbb{Z}) & \text{otherwise.} \end{cases} \tag{13}$$

4 The average over Serre curves

We will now prove Theorem 10, which says that

$$\frac{1}{|\mathcal{F}(A,B)|} \sum_{\substack{E\in\mathcal{F}(A,B) \\ E \text{ is a Serre curve}}} |C_E - C|^k \ll_k \frac{1}{A} + \left(\frac{\log B \cdot (\log A)^7}{B}\right)^{k/(k+1)}.$$

In Sect. 4.1, we compute the constants C_E when E is a Serre curve, showing that each summand in the left-hand side of the above equation is $O\left(\frac{1}{|\Delta_{sf}(E)|^k}\right)$. In Sect. 4.2, we finish the estimate.

4.1 The constants associated to Serre curves

In this section, we will explicitly compute the constants $C_{E,r}$, $C_{E,\text{prime}}$ and $C_{E,\text{cyclic}}$ for E a Serre curve. For the Lang–Trotter constant $C_{E,r}$, we must fix some notation. First define the exponent $k_2 \in \{1, 2, 3\}$ and the odd integer W by the decomposition

$$M_E =: 2^{k_2} \cdot W,$$

where M_E is as in (12). Explicitly, we have

$$W := \frac{\Delta_{sf}}{(\Delta_{sf}, 2)} \quad \text{and} \quad k_2 := \begin{cases} 1 & \text{if } \Delta_{sf} \equiv 1 \mod 4 \\ 2 & \text{if } \Delta_{sf} \equiv 3 \mod 4 \\ 3 & \text{if } \Delta_{sf} \equiv 2 \mod 4. \end{cases}$$

We also recall the notation $\omega(n) := |\{p \text{ prime } : p \mid n\}|$ and

$$\mu(n) := \begin{cases} (-1)^{\omega(n)} & \text{if } n \text{ is square-free} \\ 0 & \text{otherwise.} \end{cases}$$

When 2^{k_2-1} divides r, we further define the symbol $\delta(\Delta_{sf}, r) \in \{\pm 1\}$ by

$$\delta(\Delta_{sf}, r) := (-1)^{\omega\left(\frac{W}{(W,r)}\right) + \frac{W+1}{2} + \frac{r}{2^{k_2-1}}} \cdot \chi_4\left(-\frac{\Delta_{sf}}{2}\right),$$

where we are defining $\chi_4 : \mathbb{Q} \to \{\pm 1\}$ by

$$\chi_4(x) := \begin{cases} -1 & \text{if } x \in \mathbb{Z} \text{ and } x \equiv -1 \mod 4 \\ 1 & \text{otherwise.} \end{cases}$$

Note that when x is an odd integer, our $\chi_4(x)$ coincides with the usual character of conductor 4.

Proposition 15 *Suppose that E is an elliptic curve over $\mathbb{Q}$ which is a Serre curve. Then*

$$C_{E,r} = \begin{cases} C_r \left(1 + \delta(\Delta_{sf}, r) \cdot \dfrac{M_E \cdot 2^{k_2-1} \cdot \varphi((W,r))}{|GL_2(\mathbb{Z}/M_E\mathbb{Z})_r|}\right) & \textit{if } 2^{k_2-1} \mid r \\ C_r & \textit{otherwise,} \end{cases} \tag{14}$$

$$C_{E,prime} = \begin{cases} C_{prime} \left(1 + \prod_{p \mid \Delta_{sf}} \dfrac{1}{p^3 - 2p^2 - p + 3}\right) & \textit{if } \Delta_{sf} \equiv 1 \mod 4 \\ C_{prime} & \textit{otherwise,} \end{cases} \tag{15}$$

and

$$C_{E,cyclic} = \begin{cases} C_{cyclic}\left(1 + \dfrac{\mu(M_E)}{\prod_{p\mid M_E}(|GL_2(\mathbb{Z}/p\mathbb{Z})|-1)}\right) & \textit{if } \Delta_{sf} \equiv 1 \mod 4 \\ C_{cyclic} & \textit{otherwise}. \end{cases} \tag{16}$$

The proof of Proposition 15 will require the use of some technical lemmas. We now describe the set-up of the first of these lemmas. Let M be any positive integer and recall the isomorphism of the Chinese remainder theorem:

$$GL_2(\mathbb{Z}/M\mathbb{Z}) \simeq \prod_{p^{k_p}\|M} GL_2(\mathbb{Z}/p^{k_p}\mathbb{Z}), \quad x \mapsto (x_{p^{k_p}}).$$

Suppose that $X_M \subseteq GL_2(\mathbb{Z}/M\mathbb{Z})$ is any subset which, under the above isomorphism, satisfies

$$X_M \simeq \prod_{p^{k_p}\|M} X_{p^{k_p}}, \tag{17}$$

where $X_{p^{k_p}}$ denotes the projection of X_M onto the p^{k_p}th factor. Suppose further that, for each prime p dividing M we have a group homomorphism

$$\psi_{p^{k_p}} : GL_2(\mathbb{Z}/p^{k_p}\mathbb{Z}) \longrightarrow \{\pm 1\},$$

and write

$$\psi_M : GL_2(\mathbb{Z}/M\mathbb{Z}) \longrightarrow \{\pm 1\}, \quad \psi_M(x) := \prod_{p^{k_p}\|M} \psi_{p^{k_p}}(x_{p^{k_p}}).$$

Lemma 16 *With notation as just outlined, we have*

$$|\psi_M^{-1}(\pm 1) \cap X_M| = \frac{1}{2}\left(|X_M| \pm \prod_{p^{k_p}\|M}\left(|\psi_{p^{k_p}}^{-1}(1) \cap X_{p^{k_p}}| - |\psi_{p^{k_p}}^{-1}(-1) \cap X_{p^{k_p}}|\right)\right).$$

Proof Define the set $\mathcal{S}$ by

$$\mathcal{S} := \left\{(s_p)_{p\mid M} \in \mathbb{R}^{\omega(M)} : \forall p \mid M, s_p \in \{\pm 1\}\right\}.$$

(Thus, $|\mathcal{S}| = 2^{\omega(M)}$.) We begin by noting that

$$\begin{aligned}\left|\psi_M^{-1}(\pm 1) \cap X_M\right| &= \sum_{\substack{(s_p)\in\mathcal{S} \\ \prod s_p = \pm 1}} \prod_{p^{k_p} \| M} \left|\psi_{p^{k_p}}^{-1}(s_p) \cap X_{p^{k_p}}\right| \\ &= \sum_{\substack{(s_p)\in\mathcal{S} \\ \prod s_p = \pm 1}} \prod_{p^{k_p} \| M} \left(F_1(p^{k_p}) + s_p F_{-1}(p^{k_p})\right),\end{aligned}$$

where

$$F_1(p^{k_p}) := \frac{1}{2}|X_{p^{k_p}}| \quad \text{and} \quad F_{-1}(p^{k_p}) := \frac{1}{2}\left(|\psi_{p^{k_p}}^{-1}(1) \cap X_{p^{k_p}}| - |\psi_{p^{k_p}}^{-1}(-1) \cap X_{p^{k_p}}|\right).$$

Expanding the product and reversing summation, we obtain

$$\left|\psi_M^{-1}(\pm 1) \cap X_M\right| = \sum_{(t_p)\in\mathcal{S}} \prod_{p^{k_p} \| M} F_{t_p}(p^{k_p}) \left(\sum_{\substack{(s_p)\in\mathcal{S} \\ \prod s_p = \pm 1}} \left(\prod_{\substack{p \mid M \\ t_p = -1}} s_p \right)\right), \tag{18}$$

Now we show that, for all tuples (t_p) except $(t_p) \in \{(1, 1, \dots, 1), (-1, -1, \dots, -1)\}$, the innermost sum is equal to zero. For suppose that

$$\{p : p \mid M, t_p = 1\} \neq \emptyset \neq \{p : p \mid M, t_p = -1\},$$

and fix a prime p_1 with $t_{p_1} = 1$ and a prime p_2 with $t_{p_2} = -1$. For a tuple (s_p), define its dual $(\hat{s}_p)$ by

$$\hat{s}_{p_1} = -s_{p_1}, \quad \hat{s}_{p_2} = -s_{p_2}, \quad \text{and} \quad \hat{s}_p = s_p \quad (p \notin \{p_1, p_2\}).$$

Noting that

$$\prod_{p \mid M} s_p = \prod_{p \mid M} \hat{s}_p \quad \text{and} \quad \prod_{\substack{p \mid M \\ t_p = -1}} s_p + \prod_{\substack{p \mid M \\ t_p = -1}} \hat{s}_p = 0,$$

we see that, except when $(t_p) \in \{(1, 1, \dots, 1), (-1, -1, \dots, -1)\}$, the innermost sum in (18) vanishes. Thus,

$$\left|\psi_M^{-1}(\pm 1) \cap X_M\right| = 2^{\omega(M)-1}\left(\prod_{p^{k_p} \| M} F_1(p^{k_p}) \pm \prod_{p^{k_p} \| M} F_{-1}(p^{k_p})\right).$$

By (17), this proves Lemma 16. □

To prove Proposition 15, we will take $M = M_E$ and

$$\psi_{p^{k_p}}(\sigma) := \begin{cases} \left(\frac{\det(\sigma)}{p}\right) & \text{if } p \text{ is odd} \\ \varepsilon(\sigma) & \text{if } p^{k_p} = 2 \text{ and } \Delta_{sf} \equiv 1 \mod 4 \\ \chi_4(\det\sigma)\varepsilon(\sigma) & \text{if } p^{k_p} = 4 \text{ and } \Delta_{sf} \equiv 3 \mod 4 \\ \chi_8(\det\sigma)\varepsilon(\sigma) & \text{if } p^{k_p} = 8 \text{ and } \Delta_{sf} \equiv 2 \mod 8 \\ \chi_4(\det\sigma)\chi_8(\det\sigma)\varepsilon(\sigma) & \text{if } p^{k_p} = 8 \text{ and } \Delta_{sf} \equiv 6 \mod 8, \end{cases} \tag{19}$$

where the characters χ_4 and χ_8 are defined for odd integers x by

$$\chi_4(x) := \begin{cases} 1 & \text{if } x \equiv 1 \mod 4 \\ -1 & \text{if } x \equiv -1 \mod 4 \end{cases}$$

and

$$\chi_8(x) := \begin{cases} 1 & \text{if } x \equiv \pm 1 \mod 8 \\ -1 & \text{if } x \equiv \pm 3 \mod 8. \end{cases}$$

Note that $\varepsilon(\cdot) \cdot \left(\frac{\Delta_{sf}}{\det(\cdot)}\right) = \prod_{p^{k_p} \| M_E} \psi_{p^{k_p}}(\cdot)$, and so we have

$$\overline{H_E} = \psi_{M_E}^{-1}(1).$$

Proof of (14) If E is any non-CM elliptic curve then

$$\frac{C_{E,r}}{C_r} = \frac{m_E |G_{m_E}(E)_r|}{|G_{m_E}(E)|} \cdot \frac{|GL_2(\mathbb{Z}/m_E\mathbb{Z})|}{m_E |GL_2(\mathbb{Z}/m_E\mathbb{Z})_r|}.$$

Here, we have used the fact (see [14, pp. 34–35]) that

$$\frac{m_E |GL_2(\mathbb{Z}/m_E\mathbb{Z})_r|}{|GL_2(\mathbb{Z}/m_E\mathbb{Z})|} = \prod_{\ell | m_E} \frac{\ell |GL_2(\mathbb{Z}/\ell\mathbb{Z})_r|}{|GL_2(\mathbb{Z}/\ell\mathbb{Z}|}.$$

Thus, when E is a Serre curve, we have

$$\frac{C_{E,r}}{C_r} = \frac{2 \left| \left(\overline{H_E}\right)_r \right|}{|GL_2(\mathbb{Z}/M_E\mathbb{Z})_r|}.$$

To evaluate $|(\overline{H_E})_r|$, we will apply Lemma 16 with $M = M_E$, $\psi_{p^{k_p}}$ as in (19) and

$$X_{M_E} := GL_2(\mathbb{Z}/M_E\mathbb{Z})_r = \{g \in GL_2(\mathbb{Z}/M_E\mathbb{Z}) : \operatorname{tr} g \equiv r \mod M_E\}.$$

Thus, by Lemma 16, we have

$$\begin{aligned}\left|(\overline{H_E})_r\right| &= \left|\psi_{M_E}^{-1}(1) \cap GL_2(\mathbb{Z}/M_E\mathbb{Z})_r\right| \\ &= \frac{1}{2}\left(|GL_2(\mathbb{Z}/M_E\mathbb{Z})_r| + \prod_{p^{k_p} \| M} \left(|\psi_{p^{k_p}}^{-1}(1)_r| - |\psi_{p^{k_p}}^{-1}(-1)_r|\right)\right),\end{aligned}$$

from which it follows that

$$\frac{C_{E,r}}{C_r} = 1 + \frac{\prod_{p^{k_p} \| M_E} \left(|\psi_{p^{k_p}}^{-1}(1)_r| - |\psi_{p^{k_p}}^{-1}(-1)_r|\right)}{|GL_2(\mathbb{Z}/M_E\mathbb{Z})_r|}. \tag{20}$$

Note that for any odd prime p dividing M_E we have $k_p = 1$.

Lemma 17 *For odd primes p and ψ_p as in (19), one has*

$$|\psi_p(1)_r| - |\psi_p(-1)_r| = \begin{cases} \left(\frac{-1}{p}\right) p(p-1) & \text{if } p \mid r \\ -\left(\frac{-1}{p}\right) p & \text{if } p \nmid r. \end{cases}$$

Proof of Lemma 17 See the table in [14, p. 45]. The relationship between their notation and ours is

$$E(q)_r = \psi_q(1)_r \quad \text{and} \quad O(q)_r = \psi_q(-1)_r.$$

Lemma 17 follows immediately. □

For $p = 2$, we use the following lemma, whose proof is a straightforward calculation which we omit.

Lemma 18 *If $\Delta_{sf} \equiv 1 \mod 4$ then*

$$|\psi_2^{-1}(1)_0| = 1,\ |\psi_2^{-1}(-1)_0| = 3,\ |\psi_2^{-1}(1)_1| = 2, \ \textit{and}\ |\psi_2^{-1}(-1)_1| = 0.$$

If $\Delta_{sf} \equiv 3 \mod 4$ then

$$|\psi_4^{-1}(1)_0| = |\psi_4^{-1}(-1)_2| = 12,\ |\psi_4^{-1}(-1)_0| = |\psi_4^{-1}(1)_2| = 20,$$

and for any odd r modulo 4,

$$|\psi_4^{-1}(\pm 1)_r| = 8.$$

If $\Delta_{sf} \equiv 2 \mod 4$ then

$$|\psi_8^{-1}(\pm 1)_r| = \begin{cases} 16 \cdot 8 & \textit{if } r \equiv 2 \mod 4 \\ 16 \cdot 4 & \textit{if } r \textit{ is odd.} \end{cases},$$

while

$$|\psi_8^{-1}(1)_0| = |\psi_8^{-1}(-1)_4| = \begin{cases} 16 \cdot 9 & \textit{if } \Delta_{sf} \equiv 2 \mod 8 \\ 16 \cdot 7 & \textit{if } \Delta_{sf} \equiv 6 \mod 8 \end{cases}$$

and

$$|\psi_8^{-1}(-1)_0| = |\psi_8^{-1}(1)_4| = \begin{cases} 16 \cdot 7 & \textit{if } \Delta_{sf} \equiv 2 \mod 8 \\ 16 \cdot 9 & \textit{if } \Delta_{sf} \equiv 6 \mod 8. \end{cases}$$

Corollary 19 *For $\psi_{2^{k_2}}$ as in (19), we have*

$$|\psi_{2^{k_2}}(1)_r| - |\psi_{2^{k_2}}(-1)_r| = \begin{cases} -(-1)^{r/2^{k_2-1}} \chi_4(-\Delta_{sf}/2) \cdot 2^{2k_2-1} & \textit{if } 2^{k_2-1} \mid r \\ 0 & \textit{otherwise,} \end{cases}$$

where here we use the convention that $\chi_4(x) = 1$ if x is not an integer.

Inserting the results of Corollary 19 and Lemma 17 into (20) (and remembering that W is square-free), we find that $C_{E,r}/C_r = 1$ unless 2^{k_2-1} divides r, in which case

$$\frac{C_{E,r}}{C_r} = 1 + \frac{(-1)^{r/2^{k_2-1}+1} \chi_4(-\Delta_{sf}/2) \cdot 2^{2k_2-1} \cdot \left(\frac{-1}{W}\right) W \cdot (-1)^{\omega(W/(W,r))} \varphi((W,r))}{|GL_2(\mathbb{Z}/M_E\mathbb{Z})_r|}.$$

This finishes the proof of (14). □

Having proved (14), we now proceed to

Proof of (15) This computation may also be found in [19]. For any non-CM elliptic curve E, we have

$$\frac{C_{E,\text{prime}}}{C_{\text{prime}}} = \frac{|G_{m_E}(E) \cap \Phi_{m_E}|}{|G_{m_E}(E)|} \cdot \prod_{\ell \mid m_E} \left(\frac{|GL_2(\mathbb{Z}/\ell\mathbb{Z})|}{|\Phi_\ell|} \right).$$

If E is a Serre curve, then we have

$$\frac{C_{E,\text{prime}}}{C_{\text{prime}}} = \frac{2|\psi_{M_E}^{-1}(1) \cap \Phi_{M_E}|}{|\Phi_{M_E}|}. \tag{21}$$

Applying Lemma 16, we find that

$$|\psi_{M_E}^{-1}(1) \cap \Phi_{M_E}| = \frac{1}{2} \left(|\Phi_{M_E}| + \prod_{p^{k_p} \| M_E} \left(|\psi_{p^{k_p}}^{-1}(1) \cap \Phi_{p^{k_p}}| - |\psi_{p^{k_p}}^{-1}(-1) \cap \Phi_{p^{k_p}}| \right) \right).$$

Lemma 20 *For p odd, one has*

$$|\Phi_p| = p(p^3 - 2p^2 - p + 3)$$

and

$$|\psi_p^{-1}(1) \cap \Phi_p| - |\psi_p^{-1}(-1) \cap \Phi_p| = p.$$

Proof of Lemma 20 By considering matrices in $GL_2(\mathbb{Z}/p\mathbb{Z})$ which are conjugate to

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 \\ 0 & x \end{pmatrix} \quad (x \neq 1),$$

we see that

$$\left|\psi_p^{-1}(\pm 1) \cap (GL_2(\mathbb{Z}/p\mathbb{Z}) - \Phi_p)\right| = \frac{1}{2}\left(p^3 - (2 \pm 1)p\right).$$

Thus we have

$$|\psi_p^{-1}(\pm 1) \cap \Phi_p| = \frac{1}{2} \cdot p\left(p^3 - 2p^2 - p + 3 \pm 1\right),$$

upon which Lemma 20 follows immediately. □

The following lemma is a straightforward calculation.

Lemma 21 *One has*

$$|\psi_{2^{k_2}}^{-1}(1) \cap \Phi_{2^{k_2}}| - |\psi_{2^{k_2}}^{-1}(-1) \cap \Phi_{2^{k_2}}| = \begin{cases} 2 & \text{if } k_2 = 1 \\ 0 & \text{if } k_2 \in \{2, 3\}. \end{cases}$$

Inserting the results of Lemmas 20 and 21 into (21), we finish the proof of (15). □

Proof of (16) We have

$$\frac{C_{E,\text{cyclic}}}{C_{\text{cyclic}}} = \frac{\sum_{n|m_E} \frac{\mu(n)}{|G_n(E)|}}{\prod_{\ell|m_E}\left(1 - \frac{1}{|GL_2(\mathbb{Z}/\ell\mathbb{Z})|}\right)}.$$

If E is a Serre curve then $m_E = M_E$. Note that M_E is square-free if and only if $\Delta_{sf}(E) \equiv 1 \mod 4$. Thus, if E is a Serre curve, we deduce from (13) that

$$\sum_{n|m_E} \frac{\mu(n)}{|G_n(E)|} = \begin{cases} \prod_{\ell|m_E} \left(1 - \frac{1}{|GL_2(\mathbb{Z}/\ell\mathbb{Z})|}\right) + \frac{\mu(m_E)}{|GL_2(\mathbb{Z}/m_E\mathbb{Z})|} & \text{if } \Delta_{sf} \equiv 1 \mod 4 \\ \prod_{\ell|m_E} \left(1 - \frac{1}{|GL_2(\mathbb{Z}/\ell\mathbb{Z})|}\right) & \text{otherwise.} \end{cases}$$

This proves (16). □

We have now proved (14)–(16), finishing the proof of the Proposition 15.

4.2 Averaging the Serre curve constants

Considering Proposition 15, we see that when E is a Serre curve, C_E has the form

$$C_E = C\left(1 + g(\Delta_{sf}(E))\right) \quad \text{where} \quad g(\Delta_{sf}(E)) \ll \frac{1}{|\Delta_{sf}(E)|}.$$

Since the discriminant of the curve $Y^2 = X^3+aX+b$ is $-16(4a^3+27b^2)$, Theorem 10 will follow from

$$\frac{1}{4AB} \sum_{\substack{|a|\le A \\ |b|\le B \\ 4a^3+27b^2\neq 0}} \frac{1}{|(4a^3+27b^2)_{sf}|^k} \ll \frac{1}{A} + \left(\frac{\log B \cdot (\log A)^7}{B}\right)^{k/(k+1)}. \quad (22)$$

Let Z be a positive real number to be chosen later. Since the left hand side is bounded by

$$\frac{1}{4AB} \sum_{\substack{|a|\le A \\ |b|\le B \\ 4a^3+27b^2\neq 0 \\ |(4a^3+27b^2)_{sf}|\le Z}} 1 + \frac{1}{4AB} \sum_{\substack{|a|\le A \\ |b|\le B \\ |(4a^3+27b^2)_{sf}|>Z}} \frac{1}{Z^k},$$

we are led to the following lemma.

Lemma 22 *For A, B and $Z \geq 2$, we have*

$$\sum_{\substack{|a|\le A \\ |b|\le B \\ 4a^3+27b^2\neq 0 \\ |(4a^3+27b^2)_{sf}|\le Z}} 1 \ll B + \log B \cdot A \cdot (\log A)^7 \cdot Z, \quad (23)$$

with an absolute implied constant.

Proof The proof boils down to counting ideals of bounded norm in various quadratic fields. I would like to thank R. Daileda for helpful discussions regarding this viewpoint. We wish to count the number of integer pairs $(a, b) \in [-A, A] \times [-B, B]$ which satisfy the equation

$$4a^3 + 27b^2 = dy^2,$$

where $y > 0$ and d are integers with $d \neq 0$ square-free and $|d| \leq Z$. Re-writing this equation as

$$x^2 - 3dy^2 = 12(-a)^3,$$

where $x = 9b$, we see that the left hand side of (23) is bounded by

$$\sum_{\substack{1 \leq |d| \leq Z \\ d \text{ square-free}}} \sum_{|a| \leq A} \left| \{(x, y) \in \mathbb{Z} \times \mathbb{Z}_{>0} : (x + y\sqrt{3d})(x - y\sqrt{3d}) = 12a^3, \ |x| \leq 9B\} \right|. \tag{24}$$

Regarding $\alpha = x + y\sqrt{3d}$ as an integer in $\mathbb{Q}(\sqrt{3d})$, we are led to counting factorizations of $12a^3$ in $\mathcal{O}_{\mathbb{Q}(\sqrt{3d})}$, ring of integers of $\mathbb{Q}(\sqrt{3d})$.

We begin by dealing with the degenerate case $d = 3$ (where $\mathbb{Q}(\sqrt{3d}) = \mathbb{Q}$). In this case, the inner sum of (24) is bounded by

$$6B + \sum_{1 \leq |a| \leq A} \left| \{(x, y) \in \mathbb{Z} \times \mathbb{Z}_{\geq 0} : (x + y)(x - y) = 12a^3, \ |x| \leq 9B\} \right|, \tag{25}$$

where the $6B$ term corresponds to $a = 0$. For a given $a \neq 0$ and assuming $y \geq 0$, the factorization $12a^3 = (x + y)(x - y)$ uniquely determines the pair (x, y), up to the sign of x. Therefore (25) is bounded by

$$6B + \sum_{1 \leq |a| \leq A} 2\tau(12|a|^3) \ll B + \sum_{1 \leq a \leq A} \tau(a)^3 \ll B + A(\log A)^7,$$

where the last inequality follows from [11, equation (1.80)].

Returning to (24) and writing D for the square-free part of $3d$, we see that it remains to bound

$$\sum_{\substack{2 \leq |D| \leq 3Z \\ D \text{ square-free}}} \sum_{1 \leq |a| \leq A} \left| \{x + y\sqrt{D} =: \alpha \in \mathcal{O}_{\mathbb{Q}(\sqrt{D})} : \alpha \cdot \overline{\alpha} = 12a^3, \ |\alpha + \overline{\alpha}| \leq 18B\} \right|, \tag{26}$$

where $\overline{\alpha}$ denotes the Galois conjugate of α. We will presently transform this into counting ideals rather than elements, but in the real quadratic case we must worry about the presence of an infinite unit group. Fix a principal integral ideal $\beta \cdot \mathcal{O}_{\mathbb{Q}(\sqrt{D})}$.

We will now show that

$$\left|\left\{\alpha \in \mathcal{O}_{\mathbb{Q}(\sqrt{D})} : \alpha \cdot \mathcal{O}_{\mathbb{Q}(\sqrt{D})} = \beta \cdot \mathcal{O}_{\mathbb{Q}(\sqrt{D})} \text{ and } |\alpha + \overline{\alpha}| \leq 18B\right\}\right| \ll \log B, \tag{27}$$

with an absolute constant. To see this, first note that any $\alpha \in \mathcal{O}_{\mathbb{Q}(\sqrt{D})}$ we have

$$\alpha \cdot \mathcal{O}_{\mathbb{Q}(\sqrt{D})} = \beta \cdot \mathcal{O}_{\mathbb{Q}(\sqrt{D})} \iff \alpha = \beta \cdot \eta,$$

where η is a unit in $\mathcal{O}_{\mathbb{Q}(\sqrt{D})}$. Thus, if $D < 0$ we see that the left hand side of (27) is at most 6, the maximal cardinality of the unit group of any imaginary quadratic field.

We will now prove (27) when $D > 0$. In this case, any unit has the form $\eta = \pm\varepsilon_D^n$, where $n \in \mathbb{Z}$ and ε_D is the fundamental unit. First note that we may replace β in (27) with $\pm\varepsilon_D^n \cdot \beta$ and not affect the left hand side. Thus replacing β, we may assume that

$$\varepsilon_D^{-2}|\beta| < |\overline{\beta}| < |\beta|$$

and that $|\beta| \geq 1$, where $|\beta|$ denotes the absolute value of β as a real number. We will now prove (27) by showing that

$$\left|\left\{n \in \mathbb{Z} : |\beta \cdot \varepsilon_D^n + \overline{\beta} \cdot \overline{\varepsilon}_D^n| \leq 18B\right\}\right| \ll \log B. \tag{28}$$

By the reverse triangle inequality and noting that $\varepsilon = \varepsilon_D > 1$, we see that

$$\begin{aligned} \sum_{\substack{n \in \mathbb{Z} \\ \left|\beta\varepsilon^n + \overline{\beta}\overline{\varepsilon}^n\right| \leq 18B}} 1 \leq \sum_{\substack{n \in \mathbb{Z} \\ \left||\beta|\varepsilon^n - |\overline{\beta}|\varepsilon^{-n}\right| \leq 18B}} 1 &\leq \sum_{\substack{n \geq 0 \\ \varepsilon^n - \varepsilon^{-n} \leq \frac{18B}{|\beta|}}} 1 + \sum_{\substack{-n \leq -1 \\ \varepsilon^{n-2} - \varepsilon^{-n} \leq \frac{18B}{|\beta|}}} 1 \\ &\leq \sum_{\substack{n \geq 0 \\ \varepsilon^n \leq 18B+1}} 1 + \sum_{\substack{n \geq 1 \\ \varepsilon^{n-2} \leq 18B+1}} 1. \end{aligned}$$

The inequality (28) then follows from the fact that, uniformly in D, one has $\varepsilon_D > 2$.

Thus, (26) is bounded by a constant times

$$\log B \cdot \sum_{\substack{2 < |D| \leq 3Z \\ D \text{ square-free}}} \sum_{1 \leq a \leq A} \eta_D^{\text{princ}}(12a^3) \ \leq \ \log B \cdot \sum_{\substack{2 < |D| \leq 3Z \\ D \text{ square-free}}} \sum_{1 \leq a \leq A} \eta_D(12a^3),$$

where $\eta_D(m)$ (resp. $\eta_D^{\text{princ}}(m)$) is the number of integral ideals (resp. the number of *principal* ideals) in the ring $\mathcal{O}_{\mathbb{Q}(\sqrt{D})}$ of norm equal to m. We will now show that

$$\eta_D(m) \leq \tau(m) := \sum_{d|m} 1. \tag{29}$$

Since both $\eta_D(m)$ and $\tau(m)$ are multiplicative in m, we only need to consider m of the form p^β. One sees that

$$\eta_D(p^\beta) = \begin{cases} 1 & \text{if } p \text{ ramifies in } \mathbb{Q}(\sqrt{D}) \\ 1 & \text{if } p \text{ is inert in } \mathbb{Q}(\sqrt{D}) \text{ and } \beta \text{ is even} \\ 0 & \text{if } p \text{ is inert in } \mathbb{Q}(\sqrt{D}) \text{ and } \beta \text{ is odd} \\ \beta+1 & \text{if } p \text{ is split in } \mathbb{Q}(\sqrt{D}). \end{cases}$$

For example in the last case,

$$\left\{ \text{integral ideals } I \subseteq \mathcal{O}_{\mathbb{Q}(\sqrt{D})},\ N(I) = p^\beta \right\} = \left\{ \mathfrak{P}^i \overline{\mathfrak{P}}^{\beta-i} : i = 0, 1, \dots, \beta \right\},$$

where $\mathfrak{P}$ is a prime ideal above p. From this, (29) is immediate, since $\tau(p^\beta) = \beta + 1$ in all cases. Thus, we see that

$$\sum_{1 \le a \le A} \eta_D(12a^3) \ll \sum_{1 \le a \le A} \tau(a^3) \le \sum_{1 \le a \le A} \tau(a)^3 \ll A \log(A)^7,$$

the last inequality following as before from [11, equation (1.80)]. Lemma 22 follows at once. □

Finally, using $Z = (B/(\log B \cdot (\log A)^7))^{1/(k+1)}$, (22) follows, and thus so does Theorem 10.

5 The average over non-Serre curves

We finally turn to bounding the kth moment over the non-Serre curves. It is here that we assume an affirmative answer to Question 5, which implies that certain bounds on the constants C_E are uniform in E, as detailed in the following lemma.

Lemma 23 *For any elliptic curve E over $\mathbb{Q}$, we have*

$$C_{E,cyclic} \le 1.$$

If E is a CM elliptic curve over $\mathbb{Q}$, then

$$C_{E,r} \ll \log\log(|r| + 3) \quad \textit{and} \quad C_{E,prime} \ll 1. \tag{30}$$

If E is a non-CM elliptic curve over $\mathbb{Q}$, then, assuming an affirmative answer to Question 5, we have

$$C_{E,r} \ll 1 \quad \textit{and} \quad C_{E,prime} \ll 1. \tag{31}$$

Proof The bound on the cyclicity constant $C_{E,\text{cyclic}}$ comes from its heuristic interpretation as a density.

To prove (31), write $m_E = m_1 \cdot m_2$, where

$$p \mid m_1 \iff p \in \{2, 3, 5\} \text{ or } G_p(E) \subsetneq GL_2(\mathbb{Z}/p\mathbb{Z}).$$

It follows from [6, Theorem 1 of Appendix] that $G_{m_2}(E) = GL_2(\mathbb{Z}/m_2\mathbb{Z})$. Thus, we see that

$$\pi\left(G_{m_E}(E)_r\right) \subseteq GL_2(\mathbb{Z}/m_2\mathbb{Z})_r,$$

where $\pi : G_{m_E}(E) \twoheadrightarrow GL_2(\mathbb{Z}/m_2\mathbb{Z})$ is the projection map. From this it follows that

$$\frac{m_E\left|G_{m_E}(E)_r\right|}{\left|G_{m_E}(E)\right|} \leq m_1 \cdot \frac{m_2\,|GL_2(\mathbb{Z}/m_2\mathbb{Z})_r|}{|GL_2(\mathbb{Z}/m_2\mathbb{Z})|} = m_1 \cdot \prod_{p \mid m_2} \frac{p\,|GL_2(\mathbb{Z}/p\mathbb{Z})_r|}{|GL_2(\mathbb{Z}/p\mathbb{Z})|},$$

and therefore

$$C_{E,r} \leq \frac{2}{\pi} \cdot m_1 \cdot \prod_p \left(1 + \frac{1}{p^2 - 1}\right) \ll m_1.$$

An affirmative answer to Question 5 implies that $m_1 \ll 1$ (a bound on the exponents of the primes dividing m_1 follows from [1, Theorem 1.2]). The proof that $C_{E,\text{prime}} \ll 1$ in the non-CM case is entirely analogous, so we omit it. Noting that the analogue of Question 5 for CM elliptic curves E does have an affirmative answer, we see that (30) follows from (8) and (9). (This uses the fact that, since E is defined over $\mathbb{Q}$, there are only 13 possible endomorphism rings of E.) This finishes the proof of Lemma 23. □

We are now ready to bound the contribution to the kth moment from the non-Serre curves.

Proposition 24 *Let k be a positive integer. In the cases of the Lang–Trotter conjecture and the Koblitz conjecture, assume that Question 5 has an affirmative answer. Then there exists a constant γ so that*

$$\frac{1}{|\mathcal{F}(A,B)|} \sum_{\substack{E \in \mathcal{F}(A,B) \\ E \textit{ is not a Serre curve}}} |C_E - C|^k \ll_k \frac{\log^\gamma(\min\{A,B\})}{\sqrt{\min\{A,B\}}}.$$

In the case of the Lang–Trotter conjecture, the implied constant also depends on r. (Note that the above bound is unconditional in the case of the cyclicity conjecture).

Proof By Lemma 23, we have

$$\frac{1}{|\mathcal{F}(A,B)|} \sum_{\substack{E \in \mathcal{F}(A,B) \\ E \text{ is not a Serre curve}}} |C_E - C|^k \ll_k \frac{1}{|\mathcal{F}(A,B)|} \sum_{\substack{E \in \mathcal{F}(A,B) \\ E \text{ is not a Serre curve}}} 1$$

We finally use [12, Theorem 4], which in our situation may be stated as follows:

Theorem 25 *There is a $\gamma > 0$ so that*

$$\sum_{\substack{E \in \mathcal{F}(A,B) \\ E \text{ is not a Serre curve}}} 1 \ll \frac{|\mathcal{F}(A,B)| \log^{\gamma}(\min\{A,B\})}{\sqrt{\min\{A,B\}}},$$

with an absolute implied constant.

This finishes the proof Proposition 24, and thus also the proof of Theorem 6. □

Acknowledgments I wish to thank W. Duke for bringing this problem to my attention and also C. David and A. Granville for comments on an earlier version of this paper. I would also like to thank the referee for a careful reading of the manuscript and several helpful comments.

References

1. Arai, K.: On uniform lower bound of the Galois images associated to elliptic curves, preprint (Available at http://arxiv.org/abs/math/0703686)
2. Baier, S.: The Lang–Trotter conjecture on average. J. Ramanujan Math. Soc. **22**, 299–314 (2007)
3. Balog, A., Cojocaru, A.C., David, C.: Average twin prime conjecture for elliptic curves (2009, preprint)
4. Banks, W., Shparlinski, I.: Sato-Tate, cyclicity, and divisibility statistics on average for elliptic curves of small height. Isreal J. Math. (2009, in press)
5. Cojocaru, A.C.: Cyclicity of CM elliptic curves modulo p. Trans. Am. Math. Soc. **355**, 2651–2662 (2003)
6. Cojocaru, A.C.: On the surjectivity of the Galois representations associated to non-CM elliptic curves. Can. Math. Bull. **48**(1), 16–31 (2005)
7. David, C., Papalardi, F.: Average frobenius distributions of elliptic curves. Int. Math. Res. Not. **4**, 165–183 (1999)
8. Duke, W.D.: Elliptic curves with no exceptional primes. C. R. Math. Acad. Sci. Paris Sér. I **325**, 813–818 (1997)
9. Fouvry, E., Murty, M.R.: On the distribution of supersingular primes. Canad. J. Math. **48**, 81–104 (1996)
10. Greicius, A.: Elliptic curves with surjective global Galois representation. Ph.D. dissertation, UC Berkeley (2007)
11. Iwaniec, H., Kowalski, E.: Analytic number theory. AMS Colloquium Publications, vol. 53. American Mathematical Society, Providence (2004)
12. Jones, N.: Almost all elliptic curves are Serre curves. Trans. Am. Math. Soc. (2009, in press)
13. Koblitz, N.: Primality of the number of points of an elliptic curve over a finite field. Pacific J. Math. **131**(1), 157–165 (1988)
14. Lang, S., Trotter, H.: Frobenius distributions in GL_2-extensions. Lecture Notes in Math., vol. 504. Springer, Berlin (1976)
15. Murty, M.R.: On Artin's conjecture. J. Number Theory **16**(2), 147–168 (1983)
16. Serre, J.-P.: Propriétés galoisiennes des points d'ordre fini des courbes elliptiques. Invent. Math. **15**, 259–331 (1972)
17. Serre, J.-P.: Quelques applications du théorème de densité de Chebotarev. Publ. Math. I.H.E.S. **54**, 123–201 (1981)
18. Serre, J.-P.: Résumé des cours de 1977–1978. Annuaire du Collège de France 1978, 67–70, in Collected Papers, vol. III, pp. 465–468. Springer, Heidelberg (1986)
19. Zywina, D.: A refinement of Koblitz's Conjecture (2009, preprint)
20. Zywina, D.: Elliptic curves with maximal Galois action on their torsion points (2009, preprint)