

Square-free discriminants of Frobenius rings

Chantal David and Jorge Jiménez Urroz

August 13, 2008

Abstract

Let E be an elliptic curve over \mathbb{Q} . It is well known that the ring of endomorphisms of E_p , the reduction of E modulo a prime p of ordinary reduction, is an order of the quadratic imaginary field $\mathbb{Q}(\pi_p)$ generated by the Frobenius element π_p . When the curve has complex multiplication (CM), this is always a fixed field as the prime varies. However, when the curve has no CM, very little is known, not only about the order, but about the fields that might appear as algebra of endomorphisms varying the prime. The ring of endomorphisms is obviously related with the arithmetic of $a_p^2 - 4p$, the discriminant of the characteristic polynomial of the Frobenius element. In this paper, we are interested in the function $\pi_{E,r,h}^{\text{sf}}(x)$ counting the number of primes p up to x such that $a_p^2 - 4p$ is square-free and in the congruence class r modulo h . We give in this paper the precise asymptotic for $\pi_{E,r,h}^{\text{sf}}(x)$ when averaging over elliptic curves defined over the rationals, and we discuss the relation of this result with the Lang-Trotter conjecture, and with some other problems related to the curve modulo p .

1 Introduction and statement of results

Let E be an elliptic curve over \mathbb{Q} with conductor N_E . For each prime p of good reduction (i.e. $p \nmid N_E$), E reduces to an elliptic curve over the finite field \mathbb{F}_p . The Frobenius endomorphism $(x, y) \mapsto (x^p, y^p)$ of E/\mathbb{F}_p has characteristic polynomial

$$x^2 - a_p x + p = (x - \pi_p)(x - \bar{\pi}_p)$$

where $|a_p| \leq 2\sqrt{p}$ by the Hasse bound. Let $\text{End}(E/\mathbb{F}_p)$ be the ring of endomorphism of E/\mathbb{F}_p . If $p > 3$ is an ordinary prime for E (or equivalently, $a_p \neq 0$), then

$$\text{End}(E/\mathbb{F}_p) \otimes \mathbb{Q} = \mathbb{Q}(\pi_p)$$

is completely determined by the Frobenius ring $\mathbb{Z}[\pi_p]$. The ring of endomorphism $\text{End}(E/\mathbb{F}_p)$ is more subtle, and the Frobenius ring $\mathbb{Z}[\pi_p]$ can be a proper subset of $\text{End}(E/\mathbb{F}_p)$. In fact, it follows from Deuring's Theorem [De] that any order R such that

$$\mathbb{Z}[\pi_p] \subseteq R \subseteq \mathbb{Q}(\pi_p)$$

can occur as the ring of endomorphism of some curve E over \mathbb{F}_p such that the Frobenius of E has characteristic polynomial $x^2 - a_p x + p$. It is then a natural question to ask whether

$\mathbb{Z}[\pi_p] \simeq \text{End}(E/\mathbb{F}_p)$, or whether $\mathbb{Z}[\pi_p]$ is the maximal order of $\mathbb{Q}(\pi_p)$ (the second question is of course a refinement of the first one), and when this happens for a fixed curve E when varying the prime p .

For a fixed E/\mathbb{F}_p , let Δ_p be the discriminant of the order $\text{End}(E/\mathbb{F}_p)$. Then, the “distance” between the endomorphism ring and the Frobenius ring is encoded by the unique positive integer b_p defined by the relation $a_p^2 - 4p = b_p^2 \Delta_p$. The b_p ’s were studied by Duke and Toth in [DuTo], and by Cojocaru and Duke [CoDu] who showed that under the GRH

$$\#\{p \leq x : p \nmid N_E \text{ and } b_p = 1\} \sim C_{CD}(E) \frac{x}{\log x},$$

for an explicit non-zero constant $C_{CD}(E)$. The condition $b_p = 1$ is equivalent to the triviality of the Tate-Shafarevich group of E over its function field $\mathbb{F}_p(E)$.

In general, properties of the Frobenius ring can be deduced from a study of the arithmetic of $a_p^2 - 4p$ and, in particular, in studying the square divisors of $a_p^2 - 4p$. In this paper, we study the following question: given an elliptic curve E over \mathbb{Q} , for which primes p is $a_p^2 - 4p$ square-free? For those primes, the Frobenius ring $\mathbb{Z}[\pi_p]$ is the maximal order in $\mathbb{Q}(\pi_p)$. This also provides a refinement of the question of Cojocaru and Duke. Since the result of Cojocaru and Duke is itself a refinement of Serre’s result about the cyclicity of the group $E(\mathbb{F}_p)$, we have the string of implications

$$a_p^2 - 4p \text{ squarefree} \implies b_p = 1 \implies E(\mathbb{F}_p) \text{ cyclic.}$$

Can we then show that there is a positive proportion of $a_p^2 - 4p$ which are square-free, refining Cojocaru and Duke’s and Serre’s result? Maybe surprisingly, we believe that this is a much more difficult question. If E has complex multiplication (CM), $a_p^2 - 4p$ square-free is equivalent to p lying in some quadratic progression. For example, let E be the CM elliptic curve $y^2 = x^3 - x$ with complex multiplication by $\mathbb{Z}[i]$. Let p be an ordinary prime, which is a prime congruent to 1 modulo 4. Since E has rational 2-torsion, a_p is even, and then 4 divides $a_p^2 - 4p$, and the natural question to consider is the square-freeness of $(a_p^2 - 4p)/4$. Since E has complex multiplication by $\mathbb{Z}[i]$, $a_p^2 - 4p = -4f^2$ for some $f \in \mathbb{Z}$, and $(a_p^2 - 4p)/4$ is square-free if and only if $f = 1$ if and only if $p = (a_p/2)^2 + 1$.

Another reason to investigate square-freeness in the sequence $a_p^2 - 4p$ is that it might shed light on the following conjecture.

Conjecture 1 (Lang-Trotter conjecture [LaTr]) *Let K be an imaginary quadratic number field, and E an elliptic curve over \mathbb{Q} without complex multiplication. Let*

$$\Pi_{E,K}(x) = \#\{p \leq x : p \nmid N_E \text{ and } \mathbb{Q}(\pi_p) = K\}.$$

Then as $x \rightarrow \infty$

$$\Pi_{E,K}(x) \sim C_{LT}(E, K) \frac{\sqrt{x}}{\log x},$$

for some explicit constant $C_{LT}(E, K)$ depending on E and K .

Upper bounds for $\Pi_{E,K}(x)$ were investigated by Serre, by Cojocaru, Fouvry and Murty [CoFoMu], and by Cojocaru and David [CoDa]. But there are no known lower bounds for $\Pi_{E,K}(x)$, in particular there are no known examples of (non-CM) elliptic curves such that

$\mathbb{Q}(\pi_p) = K$ for infinitely many primes p . In fact, we know much less than that, as there are no known examples of (non-CM) elliptic curves such that D_p , the discriminant of $\mathbb{Q}(\pi_p)$, lies in a fixed congruence class for infinitely many primes p . Here again, one could give a lower bound for the number of such primes by counting the number of primes p such that $a_p^2 - 4p$ is square-free and in the prescribed congruence class, since in this case $a_p^2 - 4p = D_p$.

Let h be a positive odd integer, and let r be any integer such that the greatest common divisor (r, h) is square-free. Let

$$\Pi_{E,r,h}^{\text{sf}}(x) = \#\{2 < p \leq x, \text{ prime} : a_p^2 - 4p \in \Delta(r, h)\},$$

where $\Delta(r, h)$ is the set of square-free integers n such that $n \equiv r \pmod{h}$. We remark that if (r, h) is not square-free, then $\Pi_{E,r,h}^{\text{sf}}(x) = 0$. As for the restriction to h odd, it simplifies technical aspects of the proof in various places, but it is not an essential restriction. Unfortunately, we cannot give an asymptotic (or a lower bound) for $\Pi_{E,r,h}^{\text{sf}}(x)$, but we can prove that the correct asymptotic holds on average over elliptic curves.

In all the following, $E(a, b)$ denotes the elliptic curve given by the equation $y^2 = x^3 + ax + b$, with $4a^3 + 27b^2 \neq 0$.

Theorem 2 *Let h be a positive odd integer, and let r be any integer such that (r, h) is square-free. Fix any $\varepsilon > 0$. Let A, B be such that $AB > x \log^8 x$, $A, B > x^\varepsilon$. Let $\mathcal{C}(A, B)$ be the set of all elliptic curves $E(a, b)$ with integer coefficients a, b such that $|a| \leq A$ and $b \leq B$. Then, as $x \rightarrow \infty$,*

$$\frac{1}{|\mathcal{C}(A, B)|} \sum_{E(a,b) \in \mathcal{C}(A,B)} \Pi_{E(a,b),r,h}^{\text{sf}}(x) = \mathfrak{C} \frac{x}{\log x} + O\left(\frac{x}{\log^2 x}\right),$$

where \mathfrak{C} is the positive constant

$$\mathfrak{C} = \frac{1}{3h} \prod_{\substack{\ell|h \\ \ell \nmid r}} \frac{\ell-1}{\ell} \prod_{\substack{\ell|h \\ \ell \nmid r}} \frac{\ell(\ell-1 - (\frac{r}{\ell}))}{(\ell-1)(\ell - (\frac{r}{\ell}))} \prod_{\ell|h} \frac{\ell^4 - 2\ell^2 - \ell + 1}{\ell^2(\ell^2 - 1)}, \quad (1)$$

where all products are taken over odd primes ℓ with the specified conditions.

The result above is similar in spirit, and by the techniques used to prove it, to other average results as [BaCoDa] for the Koblitz's conjecture and [FoMu] and [DaPa] for the Lang-Trotter conjecture. All those results rely on the fact that the average over elliptic curves in $\mathcal{C}(A, B)$ can be rewritten as an average over elliptic curves over \mathbb{F}_p by interchanging the summations, and this average can be reduced to an average of class numbers by means of Deuring's Theorem. Hence, Theorem 2 is equivalent to

Theorem 3 *Let h be a positive odd integer, and let r be any integer such that (r, h) is square-free. Let*

$$\Pi^{\text{sf}}(p) = \#\{E \text{ over } \mathbb{F}_p : a_p^2 - 4p \in \Delta(r, h)\}.$$

Then, as $x \rightarrow \infty$,

$$\sum_{p \leq x} \Pi^{\text{sf}}(p) = \frac{\mathfrak{C}}{3} \frac{x^3}{\log x} + O\left(\frac{x^3}{\log^2 x}\right),$$

where \mathfrak{C} is the constant of Theorem 2.

There is another sequence related to elliptic curves which was investigated for square-freeness, namely the sequence $p + 1 - a_p = |E(\mathbb{F}_p)|$. Again, there should be a positive proportion of square-free $p + 1 - a_p$, but this is still an open question, and there are no known examples of non-CM elliptic curves over \mathbb{Q} with infinitely many square-free $p + 1 - a_p$. The case of CM curves is very different. Let E be an elliptic curve with CM by an order \mathcal{O} in a quadratic imaginary field K , and let p be a prime of ordinary reduction. Then, $p + 1 - a_p$ is square-free if and only if $\pi_p - 1 = \mathfrak{p}_1 \dots \mathfrak{p}_s$ where the \mathfrak{p}_i are distinct split primes in \mathcal{O} , and the number of such $\pi \in \mathcal{O}$ can be counted. When E has CM by the maximal order of a quadratic imaginary field K , the asymptotic for the number of primes p such that $p + 1 - a_p$ is square-free was obtained by Cojocaru [Co]. The number of square-free $p + 1 - a_p$ was also investigated in average over elliptic curves \mathbb{F}_p as p tends to infinity by Gekeler [Ge]. The techniques that he uses are completely different from ours, and rely on the work of Howe [Ho] on the moduli spaces of elliptic curves over \mathbb{F}_p with a given group structure. Gekeler then proved the analogue of our Theorem 3 for the sequence $p + 1 - a_p$ (his results are stated in a different way, but the two forms are easily seen to be equivalent). As a corollary of his result [Ge, Proposition 4.2 (i)], one obtains, by going from the average over \mathbb{F}_p to the average over \mathbb{Q} as indicated in Section 2, that the number of primes $p \leq x$ such that $p + 1 - a_p$ is square-free follows the predicted asymptotic on average over all elliptic curves. It is unclear if the techniques used by Gekeler could be used to prove Theorem 2, as $a_p^2 - 4p$ is not easily related to the group structure of $E(\mathbb{F}_p)$.

The structure of our paper is as follows. In section 2, we reduce the average over elliptic curves over \mathbb{Q} to an average over elliptic curves over \mathbb{F}_p . There are several ways to do that, and we use multiplicative characters as in [BaSh] and [BaCoDa] to obtain the short average of Theorem 2. Section 3 deals with the main error term by applying the Theorem of Barban, Davenport and Halberstam. Section 4 deals with the main term of the asymptotic, and it is basically a computation of the constant (1) of Theorem 2. Finally, we explain in Section 5 what is the conjectural asymptotic for $\Pi_{E,r,h}(x)$. The conjectural constant is very similar to our average constant \mathfrak{C} , and Theorem 2 then gives additional evidence for the conjecture.

2 Reduction to a short average over finite fields

As we have mentioned in the introduction, the proof of Theorem 2 relies in reducing the average of $\Pi_{E,r,h}^{\text{sf}}(x)$ to an average of class numbers by means of Deuring's Theorem. This could be done in a straightforward manner by just writing

$$\frac{1}{|\mathcal{C}(A, B)|} \sum_{E \in \mathcal{C}(A, B)} \Pi_{E,r,h}^{\text{sf}}(x) = \sum_{p \leq x} \sum_{(s,t) \in \mathcal{S}} \frac{1}{|\mathcal{C}(A, B)|} \sum_{\substack{|a| \leq A, |b| \leq B \\ a \equiv s \pmod{p} \\ b \equiv t \pmod{p}}} 1, \quad (2)$$

where

$$\mathcal{S} = \mathcal{S}(p, r, h) = \{(s_1, s_2) \in \mathbb{F}_p^2 : a_p(E(s_1, s_2))^2 - 4p \in \Delta(r, h)\}.$$

Observe that the number of terms in the middle sum is $\Pi^{\text{sf}}(p)$. The innermost sum is simply

$$\left(\frac{2A}{p} + O(1)\right) \left(\frac{2B}{p} + O(1)\right) \sim \frac{4AB}{p^2}$$

when A, B are big enough with respect to x . Hence, we get the asymptotic

$$\frac{1}{|\mathcal{C}(A, B)|} \sum_{E \in \mathcal{C}(A, B)} \Pi_{E, r, h}^{\text{sf}}(x) \sim \sum_{p \leq x} \frac{\Pi^{\text{sf}}(p)}{p^2}.$$

However, this approach results a poor average, as we need to take $AB > x^{2+\epsilon}$ to get the asymptotic above.

Following [BaSh] and [BaCoDa], one can obtain a substantial improvement from a better use of the uniform distribution of equivalent elliptic curves. Indeed, since two elliptic curves $E(s_1, s_2)$ and $E(t_1, t_2)$ are isomorphic over \mathbb{F}_p if and only if $t_1 = s_1 u^4$ and $t_2 = s_2 u^6$ for some $u \in \mathbb{F}_p^*$, there are $(p-1)/\#\text{Aut}(E(s_1, s_2))$ elliptic curves over \mathbb{F}_p which are \mathbb{F}_p -isomorphic to a given elliptic curve $E(s_1, s_2)$. Hence, we can write

$$\begin{aligned} & \frac{1}{|\mathcal{C}(A, B)|} \sum_{E \in \mathcal{C}(A, B)} \Pi_{E, r, h}^{\text{sf}}(x) \\ &= \frac{1}{|\mathcal{C}(A, B)|} \sum_{p \leq x} \sum_{(s, t) \in \mathcal{S}} \frac{\#\text{Aut}(E(s_1, s_2))}{p-1} \sum_{\substack{|a| \leq A, |b| \leq B, \exists 1 \leq u < p: \\ a \equiv s_1 u^4 \pmod{p}, \\ b \equiv s_2 u^6 \pmod{p}}} 1, \end{aligned} \quad (3)$$

where for each fixed $s_1, s_2 \in \mathbb{F}_p$, the innermost sum is over all integers $|a| \leq A$, $|b| \leq B$ such that there exists $u \in \mathbb{F}_p^*$ with $a \equiv s_1 u^4 \pmod{p}$ and $b \equiv s_2 u^6 \pmod{p}$. We now want to approximate the innermost sum by its main term. This is the content of Lemma 4 of [BaCoDa], and we will use it to prove the short average of Theorem 2. We include it here for reading convenience.

Lemma 4 [BaCoDa, Lemma 4] *For any positive integer k , we have as $x \rightarrow \infty$*

$$\begin{aligned} & \sum_{p \leq x} \frac{1}{p} \sum_{(s_1, s_2) \in (\mathbb{F}_p^*)^2} \left| \sum_{\substack{|a| \leq A, |b| \leq B, \exists 1 \leq u < p: \\ a \equiv s_1 u^4 \pmod{p}, \\ b \equiv s_2 u^6 \pmod{p}}} 1 - \frac{2AB}{p} \right| \\ & \ll ABx^{1-\frac{1}{2k}} \log^{\frac{k}{2}-1} x + (A\sqrt{B} + B\sqrt{A})x^{1+\frac{1}{2k}} \log^{\frac{k}{2}-1} x + \sqrt{AB}x^{3/2} \log^2 x. \end{aligned}$$

Now, since

$$|\mathcal{C}(A, B)| = 4AB + O(A+B),$$

and $\#\text{Aut}(E(s, t)) = 2$ except when $p|a$ or $p|b$, we can rewrite (3) as

$$\frac{1}{AB} \sum_{p \leq x} \sum_{(s, t) \in \mathcal{S}} \frac{1}{2(p-1)} \sum_{\substack{|a| \leq A, |b| \leq B, \exists 1 \leq u < p: \\ a \equiv s_1 u^4 \pmod{p}, \\ b \equiv s_2 u^6 \pmod{p}}} 1 + O\left(\log \log x + \frac{A+B}{AB} \frac{x}{\log x}\right) \quad (4)$$

using the trivial bound

$$\frac{1}{|\mathcal{C}(A, B)|} \sum_{p \leq x} \sum_{\substack{|a| \leq A \\ |b| \leq B \\ ab \equiv 0 \pmod{p}}} 1 \ll \log \log x + \frac{A+B}{AB} \frac{x}{\log x}.$$

Then, using Lemma 4 to replace the first term in (4), and observing that the error term in (4) is smaller than the error term of Lemma 4, we have

$$\frac{1}{|\mathcal{C}(A, B)|} \sum_{E \in \mathcal{C}(A, B)} \Pi_{E, r, h}^{\text{sf}}(x) = \sum_{p \leq x} \frac{\Pi^{\text{sf}}(p)}{p(p-1)} \quad (5)$$

$$+ O\left(x^{1-\frac{1}{2k}} \log^{\frac{k}{2}-1} x + (A^{-1/2} + B^{-1/2})x^{1+\frac{1}{2k}} \log^{\frac{k}{2}-1} x + (AB)^{-1/2} x^{3/2} \log^2 x\right).$$

By choosing A, B such that $A, B > x^\varepsilon$ and $AB > x \log^8 x$, and k large enough to have that $\varepsilon k > 1$, we have that

$$\frac{1}{|\mathcal{C}(A, B)|} \sum_{E \in \mathcal{C}(A, B)} \Pi_{E, r, h}^{\text{sf}}(x) = \sum_{p \leq x} \frac{\Pi^{\text{sf}}(p)}{p(p-1)} + O\left(\frac{x}{\log^2 x}\right). \quad (6)$$

The heart of the proof of Theorem 2 then consists in an estimate for the main term in (6), and we will show that

$$\sum_{p \leq x} \frac{\Pi^{\text{sf}}(p)}{p(p-1)} = \mathfrak{e} \frac{x}{\log x} + O\left(\frac{x}{\log^2 x}\right), \quad (7)$$

which will complete the proof of Theorem 2. With (7) and partial summation, we also get that

$$\sum_{p \leq x} \Pi^{\text{sf}}(p) = \frac{\mathfrak{e}}{3} \frac{x^3}{\log x} + O\left(\frac{x^3}{\log^2 x}\right), \quad (8)$$

and the proof of Theorem 3 also follows from (7). We first apply Deuring's Theorem to rewrite the sum in (6) as a sum of class numbers.

Theorem 5 (*Deuring's Theorem [De]*)

Let $p > 3$ be a prime and let t be an integer such that $t^2 - 4p < 0$. Let $\mathcal{E}_r(p)$ be the set of \mathbb{F}_p -isomorphism classes of elliptic curves defined over \mathbb{F}_p with $a_p = t$. Then,

$$\sum_{E \in \mathcal{E}_r(p)} \frac{1}{\#\text{Aut}(E)} = H(t^2 - 4p),$$

where $\text{Aut}(E)$ is the automorphism group of E and for any $D < 0$, $H(D)$ is the Kronecker class number

$$H(D) := \sum_{\substack{f^2 | D \\ \frac{D}{f^2} \equiv 0, 1 \pmod{4}}} \frac{h(D/f^2)}{w(D/f^2)}$$

defined in terms of the class number $h(D/f^2)$ and number of units $w(D/f^2)$ of $\mathbb{Q}(\sqrt{D/f^2})$. Then, for any fixed $-2\sqrt{p} \leq t \leq 2\sqrt{p}$, there are exactly $(p-1)H(t^2 - 4p)$ elliptic curves defined over \mathbb{F}_p with $a_p = t$.

Using the previous theorem, and noting that for square-free D we have

$$H(D) = \frac{h(D)}{w(D)},$$

we can write

$$\sum_{p \leq x} \frac{\Pi^{\text{sf}}(p)}{p(p-1)} = \sum_{\substack{p \leq x, |t| \leq 2\sqrt{p} \\ t^2 - 4p \in \Delta(r, h)}} \frac{h(t^2 - 4p)}{w(t^2 - 4p)p} = 2 \sum_{p \leq x} \sum_{\substack{\text{odd} \\ 1 \leq t \leq 2\sqrt{p} \\ t^2 - 4p \in \Delta(r, h)}} \frac{h(t^2 - 4p)}{w(t^2 - 4p)p} \quad (9)$$

since $t^2 - 4p$ is not square-free when t is even.

3 An average of class numbers

As usual, we will characterize squarefree values by using the Mobius function. In order to do so, we first let $\beta > 0$ a parameter to be chosen later, and let

$$K = [\log^{2\beta} x], \quad Y = \frac{x}{K+1}, \quad R = \log^\beta x. \quad (10)$$

Then

$$\sum_{p \leq x} \frac{\Pi^{\text{sf}}(p)}{p(p-1)} = T_1 + T_2 + O(Y/\log x), \quad (11)$$

where

$$T_i = \sum_{\substack{Y < p \leq x, |t| \leq 2\sqrt{p} \\ t^2 - 4p \in \Delta(r, h)}} \frac{h(t^2 - 4p)}{w(t^2 - 4p)p} \sum_{\substack{d^2 | t^2 - 4p \\ d \in I_i}} \mu(d),$$

for $I_1 = [1, R]$, $I_2 = (R, \sqrt{x}]$, and the error term is to take account of the primes $p < Y$. The main term of (9) will come from T_1 . In order to bound T_2 we observe that, by the well known bound $h(d) \ll \sqrt{d} \log d$, we find the trivial upper bound

$$\begin{aligned} T_2 &\ll \log x \sum_{R < d < \sqrt{x}} \sum_{|t| \leq 2\sqrt{x}} \sum_{\substack{p \leq x \\ 4p \equiv t^2 \pmod{d^2}}} \frac{1}{\sqrt{p}} \\ &= \log x \sum_{R < d < \sqrt{x}} \sum_{|t| \leq 2\sqrt{x}} \sum_{l=0}^{x/d^2-1} \sum_{\substack{ld^2 < p \leq (l+1)d^2 \\ 4p \equiv t^2 \pmod{d^2}}} \frac{1}{\sqrt{p}} \\ &\ll \sqrt{x} \log x \sum_{R < d < \sqrt{x}} \frac{1}{d} \sum_{l=0}^{x/d^2-1} \frac{1}{\sqrt{l}} \ll \frac{x \log x}{R}. \end{aligned} \quad (12)$$

Now, to estimate T_1 , we first split the sum over the primes up to x into K sums over the intervals $[kY, kY + Y]$ of length Y , and write T_1 as

$$T_1 = 2 \sum_{1 \leq k \leq K} \sum_{\substack{kY < p \leq kY + Y \\ t^2 - 4p \in \Delta(r, h)}} \sum_{\substack{\text{odd} \\ 1 \leq t \leq 2\sqrt{p}}} \frac{h(t^2 - 4p)}{w(t^2 - 4p)p} \sum_{\substack{d^2 | t^2 - 4p \\ d \leq R}} \mu(d) \quad (13)$$

On each interval $kY < p \leq kY + Y$, we approximate the sum over $\{1 \leq t \leq 2\sqrt{p}\}$ by $\{1 \leq t \leq 2\sqrt{kY}\}$ with an error term of $RY \log x$. Replacing (12) and (13) in (11), we have

$$\begin{aligned} \sum_{p \leq x} \frac{\Pi^{\text{sf}}(p)}{p(p-1)} &= 2 \sum_{1 \leq k \leq K} \sum_{1 \leq t \leq 2\sqrt{kY}}^{\text{odd}} \sum_{d \leq R}^{\text{odd}} \mu(d) \sum_{\substack{kY < p \leq kY+Y \\ d^2 | t^2 - 4p \\ t^2 - 4p \equiv r \pmod{h}}} \frac{h(t^2 - 4p)}{\omega(t^2 - 4p)p} \\ &\quad + O(x/\log^{\beta-1} x) \\ &= \frac{1}{\pi} \sum_{1 \leq k \leq K} \sum_{1 \leq t \leq 2\sqrt{kY}}^{\text{odd}} \sum_{d \leq R}^{\text{odd}} \mu(d) \sum_{\substack{kY < p \leq kY+Y \\ d^2 | t^2 - 4p \\ t^2 - 4p \equiv r \pmod{h}}} \frac{\sqrt{4p - t^2} L(1, \chi_{t^2 - 4p})}{p} \\ &\quad + O(x/\log^{\beta-1} x), \end{aligned}$$

using the class number formula. Let

$$U = x^{1/2} R^2. \tag{14}$$

Since

$$\begin{aligned} L(1, \chi_{t^2 - 4p}) &= \sum_{n \geq 1} \frac{\chi_{t^2 - 4p}(n)}{n} = \sum_{n \geq 1}^{\text{odd}} \frac{\chi_{t^2 - 4p}(n)}{n} + \frac{1}{2} \sum_{n \geq 1} \frac{\chi_{t^2 - 4p}(2n)}{n} \\ &= \left(1 + \frac{\chi_{t^2 - 4p}(2)}{2}\right) \sum_{n \geq 1}^{\text{odd}} \frac{\chi_{t^2 - 4p}(n)}{n} + \frac{1}{4} \sum_{n \geq 1} \frac{\chi_{t^2 - 4p}(n)}{n}, \end{aligned}$$

using Polya-Vinogradov inequality to bound the tail of the L-function, we obtain

$$\begin{aligned} \sum_{p \leq x} \frac{\Pi^{\text{sf}}(p)}{p(p-1)} &= \frac{2}{3\pi} \sum_{1 \leq k \leq K} \sum_{\substack{n \leq U \\ 1 \leq t \leq 2\sqrt{kY} \\ (t^2 - r, h) = 1}}^{\text{odd}} \frac{1}{n} \sum_{d \leq R}^{\text{odd}} \mu(d) \sum_{\substack{kY < p \leq kY+Y \\ d^2 | t^2 - 4p \\ t^2 - 4p \equiv r \pmod{h}}} \frac{\sqrt{4p - t^2}}{p} \chi_{t^2 - 4p}(n) \\ &\quad + O(x/\log^{\beta-1} x), \end{aligned} \tag{15}$$

since $\chi_{t^2 - 4p}(2) = -1$ for any t, p in the previous conditions. Using quadratic reciprocity, we can write χ as a character mod n , and we use this to rewrite the main term of (15) as

$$M = \frac{2}{3\pi} \sum_{1 \leq k \leq K} \sum_{\substack{n \leq U \\ 1 \leq t \leq 2\sqrt{kY} \\ (t^2 - r, h) = 1}}^{\text{odd}} \frac{1}{n} \sum_{\substack{\alpha \pmod{n} \\ (t^2 - \alpha, n) = 1 \\ \alpha \equiv r \pmod{(n, h)}}} \left(\frac{\alpha}{n}\right) \sum_{\substack{d \leq R \\ (d, nt) = 1 \\ r \equiv 0 \pmod{(d^2, h)}}}^{\text{odd}} \mu(d) S(k, Y, n, t, \alpha, d)$$

where

$$S(k, Y, n, t, \alpha, d) = \sum_{\substack{kY \leq p \leq kY+Y \\ p \equiv \nu(t, \alpha, r) \pmod{[nd^2, h]}}} \frac{\sqrt{4p - t^2}}{p},$$

and $\nu(t, \alpha, r)$ is the invertible residue modulo $[nd^2, h]$ solving the congruences

$$\begin{aligned} 4p &\equiv t^2 - r \pmod{h}, \text{ for } (t^2 - r, h) = 1, \\ 4p &\equiv t^2 - \alpha \pmod{n}, \text{ for } (t^2 - \alpha, n) = 1, \\ 4p &\equiv t^2 \pmod{d^2} \text{ for } (t, d) = 1, \end{aligned}$$

under the conditions $(n, d) = 1$, $r \equiv \alpha \pmod{(n, h)}$ and $r \equiv 0 \pmod{(d^2, h)}$ insuring that the congruences are compatible.

In order to compute the asymptotic of M , we first change the weights and consider

$$M' = \frac{2}{3\pi} \sum_{1 \leq k \leq K} \sum_{\substack{\text{odd} \\ n \leq U \\ 1 \leq t \leq 2\sqrt{kY} \\ (t^2 - r, h) = 1}} \frac{1}{n} \sum_{\substack{\alpha \pmod{n} \\ (t^2 - \alpha, n) = 1 \\ \alpha \equiv r \pmod{(n, h)}}} \left(\frac{\alpha}{n}\right) \sum_{\substack{\text{odd} \\ d \leq R \\ (d, nt) = 1 \\ r \equiv 0 \pmod{(d^2, h)}}} \mu(d) S'(k, Y, n, t, \alpha, d),$$

where

$$S'(k, Y, n, t, \alpha, d) = \sum_{\substack{kY \leq p \leq kY + Y \\ p \equiv \nu(t, \alpha, r) \pmod{[nd^2, h]}}} \frac{\sqrt{4kY - t^2} \log p}{kY \log kY}.$$

Lemma 6 *Let M and M' be the two sums defined above where $Y = x/(K + 1)$ and $K = \lfloor \log^{2\beta} x \rfloor$ for some $\beta \geq 1$. Then as $x \rightarrow \infty$*

$$M - M' \ll \frac{x}{\log^{\beta-1} x}.$$

Proof: It is enough to note that

$$\begin{aligned} S(k, Y, n, t, \alpha, d) - S'(k, Y, n, t, \alpha, d) &= \sum_{\substack{kY \leq p \leq kY + Y \\ p \equiv \nu(t, \alpha, r) \pmod{[nd^2, h]}}} \left(\frac{\sqrt{4p - t^2}}{p} - \frac{\sqrt{4kY - t^2}}{kY} \right) \\ &+ \sum_{\substack{kY \leq p \leq kY + Y \\ p \equiv \nu(t, \alpha, r) \pmod{[nd^2, h]}}} \left(\frac{\sqrt{4kY - t^2}}{kY} - \frac{\sqrt{4kY - t^2} \log p}{kY \log kY} \right) \ll \frac{\sqrt{Y}}{knd^2}. \end{aligned}$$

Summing over k, n, t, α, d , with trivial bounds for the remaining terms, gives the result. ■

We then have to evaluate the sum M' above. We first define some notation. For any positive integers a, q such that $(a, q) = 1$, let

$$\begin{aligned} \psi(X, Y; a, q) &= \sum_{\substack{X < p < X + Y \\ p \equiv a \pmod{q}}} \log p \\ E(X, Y; a, q) &= \psi(X, Y; a, q) - \frac{Y}{\varphi(q)}, \end{aligned}$$

where, as usual, φ denotes the Euler function.

Subtracting and adding $\frac{Y}{\varphi([nd^2, h])}$ to the sum M' above, we write it as $M' = S_1 + S_2$ where

$$S_1 = \frac{2Y}{3\pi} \sum_{1 \leq k \leq K} \sum_{\substack{\text{odd} \\ n \leq U \\ 1 \leq t \leq 2\sqrt{kY} \\ (t^2 - r, h) = 1}} \frac{\sqrt{4kY - t^2}}{kY \log kY} \frac{1}{n} \sum_{\substack{\alpha \pmod{n} \\ (t^2 - \alpha, n) = 1 \\ \alpha \equiv r \pmod{(n, h)}}} \left(\frac{\alpha}{n}\right) \sum_{\substack{d \leq R \\ (d, nt) = 1 \\ r \equiv 0 \pmod{(d^2, h)}}} \frac{\mu(d)}{\varphi([nd^2, h])} \quad (16)$$

and

$$\begin{aligned} S_2 &\ll \sum_{1 \leq k \leq K} \frac{1}{\sqrt{kY}} \sum_{\substack{\text{odd} \\ 1 \leq t \leq 2\sqrt{kY} \\ (t^2 - r, h) = 1}} \sum_{\substack{n \leq U \\ d \leq R \\ (d, nt) = 1 \\ r \equiv 0 \pmod{(d^2, h)}}} \frac{1}{n} \sum_{\substack{\alpha \pmod{n} \\ (t^2 - \alpha, n) = 1 \\ \alpha \equiv r \pmod{(n, h)}}} |E(kY, Y; \nu(t, \alpha, r), [nd^2, h])| \\ &\ll \sum_{1 \leq k \leq K} \sum_{q \leq R^2 U h} \frac{R^3}{q} \sum_{\substack{a \pmod{q} \\ (a, q) = 1}} |E(kY, Y; a, q)| \end{aligned} \quad (17)$$

since each $q \leq R^2 U h$ can be written as $q = [nd^2, h]$ in at most R ways, and for each fixed integer t , and for any $\alpha \not\equiv \alpha' \pmod{n}$, $\nu(t, \alpha, r) \not\equiv \nu(t, \alpha', r)$ modulo $q = [nd^2, h]$. We now use:

Theorem 7 (*Barban-Davenport-Halberstam*) *Let X, Y be positive integers such that $X + Y \leq x$. Then, for any $N > 0$, there exist $M > 0$ such that*

$$\sum_{q \leq Q} \sum_{\substack{a \pmod{q} \\ (a, q) = 1}} |E(X, Y; a, q)|^2 \ll \frac{x^2}{\log^N x}$$

whenever $Q \leq x \log^{-M} x$.

Using Cauchy-Schartz and Theorem 7 on (17), we get that for any $A > 0$

$$\begin{aligned} S_2 &\ll R^3 \sum_{1 \leq k \leq K} \left(\sum_{\substack{q \leq R^2 U h \\ a \pmod{q} \\ (a, q) = 1}} |E(kY, Y; a, q)|^2 \right)^{1/2} \left(\sum_{\substack{q \leq R^2 U h \\ a \pmod{q} \\ (a, q) = 1}} \frac{1}{q^2} \right)^{1/2} \\ &\ll x \log^{-A} x \end{aligned} \quad (18)$$

since

$$R^2 U h \leq \frac{x}{\log^M x}$$

for any $M > 0$, and x sufficiently large.

We now evaluate S_1 . Consider the sum

$$S(T) = \sum_{\substack{1 \leq t \leq T \\ (t^2 - r, h) = 1}} \sum_{n \leq U} \frac{1}{n} \sum_{\substack{\alpha \pmod{n} \\ (t^2 - \alpha, n) = 1 \\ \alpha \equiv r \pmod{(n, h)}}} \left(\frac{\alpha}{n}\right) \sum_{\substack{d \leq R \\ (d, nt) = 1 \\ r \equiv 0 \pmod{(d^2, h)}}} \frac{\mu(d)}{\varphi([nd^2, h])}. \quad (19)$$

Theorem 8 *Let R, U as before. Then, as $T, x \rightarrow \infty$ we have for any $\varepsilon > 0$*

$$S(T) = \frac{3}{2}\mathfrak{C}T + O\left(\frac{T \log R}{R}\right) + O\left(\frac{1}{T^{1-\varepsilon}}\right).$$

where \mathfrak{C} is the constant given in (1).

Let us assume for a moment the previous theorem, which will be proved in the next section, and let us show how the evaluation of S_1 in (16) follows from the asymptotic of $S(T)$. This will complete the proof of Theorem 2.

We define $F(t)$ such that

$$S(T) = \sum_{\substack{\text{odd} \\ 1 \leq t \leq T \\ (t^2-r, h)=1}} F(t),$$

where $S(T)$ is the sum defined in (19), and consider

$$S_1(X) = \sum_{\substack{\text{odd} \\ 1 \leq t \leq 2\sqrt{X} \\ (t^2-r, h)=1}} \frac{\sqrt{4X-t^2}}{X \log X} F(t).$$

Using Theorem 8 and partial summation we get

$$S_1(X) = \frac{\frac{3}{2}\mathfrak{C}}{X \log X} \int_0^{2\sqrt{X}} t^2 (4X-t^2)^{-1/2} dt = \frac{\frac{3}{2}\mathfrak{C}\pi}{\log X} + O(1/R). \quad (20)$$

Replacing (20) into (16), we get

$$\begin{aligned} S_1 &= \frac{2Y}{3\pi} \sum_{1 \leq k \leq K} S_1(kY) = \mathfrak{C}Y \sum_{1 \leq k \leq K} \frac{1}{\log kY} + O(x/R) \\ &= \mathfrak{C} \frac{x}{\log x} + O\left(\frac{x}{\log^2 x}\right). \end{aligned} \quad (21)$$

Using (21), and collecting the error terms from (18), Lemma 6 and (15) with $\beta \geq 3$, this shows that

$$\sum_{p \leq x} \frac{\Pi^{\text{sf}}(p)}{p(p-1)} = \mathfrak{C} \frac{x}{\log x} + O\left(\frac{x}{\log^2 x}\right),$$

and it completes the proof of Theorem 2.

4 Proof of Theorem 8

Let

$$c_t(n) = \sum_{\substack{\alpha \bmod n \\ (t^2-\alpha, n)=1 \\ \alpha \equiv r \bmod (n, h)}} \left(\frac{\alpha}{n}\right)$$

be the sum over residues modulo n that appears in (19).

Lemma 9 Let r, h, t be integers such that $(t^2 - r, h) = 1$, and let p be an odd prime. Then, $c_t(n)$ is a multiplicative function of n , with value at prime powers given by

1. If $p \mid h$, then $c_t(p^\ell)p^{-\ell} = \frac{1}{(p^\ell, h)} \left(\frac{r}{p^\ell} \right)$.

2. If $p \nmid h$, then

$$c_t(p^\ell)p^{-\ell} = \begin{cases} 0 & \text{if } p \mid t, \ell \text{ odd;} \\ -1/p & \text{if } p \nmid t, \ell \text{ odd;} \\ 1 - 1/p & \text{if } p \mid t, \ell \text{ even;} \\ 1 - 2/p & \text{if } p \nmid t, \ell \text{ even.} \end{cases}$$

Proof: This is a straightforward case by case computation, and we omit it. ■

Now, for each fixed odd t in (19), such that $(t^2 - r, h) = 1$, we evaluate the sum

$$\sum_{n \leq U} \frac{c_t(n)}{n} \sum_{\substack{d \leq R \\ (d, nt)=1 \\ r \equiv 0 \pmod{(d^2, h)}}} \frac{\mu(d)}{\varphi([nd^2, h])}.$$

For any $(d, n) = 1$, we have that

$$\varphi([nd^2, h]) = \frac{\varphi(n)\varphi(d^2)\varphi(h)}{\varphi((n, h))\varphi((d^2, h))} > \varphi(n)\varphi(d^2). \quad (22)$$

We also observe that

$$\sum_{\substack{d > R \\ (d, nt)=1 \\ r \equiv 0 \pmod{(d^2, h)}}} \frac{1}{\varphi(d^2)} \ll \sum_{d > R} \frac{\log \log d}{d^2} < \frac{\log R}{R},$$

and

$$\sum_{n > U} \frac{c_t(n)}{n\varphi(n)} \ll \sum_{n > U} \frac{\log \log n}{n \prod_{p^{2\gamma+1} \mid n} p} \ll \sum_{n > U^{1/2}} \frac{2^{\omega(n)} \log \log n}{n^2} < \frac{1}{U^{1/2-\varepsilon}}$$

for any $\varepsilon > 0$. Then,

$$\sum_{n \leq U} \frac{c_t(n)}{n} \sum_{\substack{d \leq R \\ (d, nt)=1 \\ r \equiv 0 \pmod{(d^2, h)}}} \frac{\mu(d)}{\varphi([nd^2, h])} = C_t + O\left(\frac{\log R}{R}\right), \quad (23)$$

where

$$C_t = \sum_{n \geq 1} \frac{c_t(n)}{n} \sum_{\substack{d \geq 1 \\ (d, nt)=1 \\ r \equiv 0 \pmod{(d^2, h)}}} \frac{\mu(d)}{\varphi([nd^2, h])} = \sum_{\substack{d \geq 1 \\ (d, t)=1 \\ (d^2, h) \mid r}} \mu(d) \sum_{\substack{n \geq 1 \\ (n, d)=1}} \frac{c_t(n)}{n\varphi([nd^2, h])}. \quad (24)$$

We now use the multiplicativity of the functions involved to write the constant C_t as an Euler product.

Lemma 10 *Let r, h, t be integers such that $(t^2 - r, h) = 1$, and (h, r) is squarefree. Then,*

$$C_t = P(r, h) \prod_{\substack{p|t \\ p \nmid h}} \left(1 + \frac{2p^2 + p - 1}{p^4 - p^3 - 2p^2 - p + 1} \right),$$

where

$$P(r, h) = \frac{1}{\varphi(h)} \prod_{p|h} \frac{p}{p - \left(\frac{r}{p}\right)} \prod_{\substack{p||h \\ p|r}} \left(1 - \frac{1}{p} \right) \prod_{p \nmid h} \frac{p^4 - p^3 - 2p^2 - p + 1}{p(p-1)(p^2-1)},$$

and all the products run over odd prime numbers.

Proof: We first observe that

$$\frac{p^4 - p^3 - 2p^2 - p + 1}{p(p-1)(p^2-1)} = 1 - \frac{p^2 + 2p - 1}{p(p-1)(p^2-1)}, \quad (25)$$

so each of the products of C_t is convergent. Although intricate, the proof consists of a combination of straightforward computations using Lemma 9. We always assume that every prime appearing in the products below is odd. We first note that

$$\sum_{\substack{\text{odd} \\ d \geq 1 \\ (d, nt) = 1 \\ r \equiv 0 \pmod{(d^2, h)}}} \frac{\mu(d)\varphi((d^2, h))}{\varphi(d^2)} = \prod_{\substack{p \nmid nt \\ p||h \\ p|r}} \left(1 - \frac{1}{p} \right) \prod_{p \nmid h} \left(1 - \frac{1}{p(p-1)} \right)$$

Hence, using (22), we get

$$C_t = F(h, t) \sum_{n \geq 1}^{\text{odd}} \frac{c_t(n)\varphi((n, h))}{n\varphi(n)} \prod_{\substack{p|(r, n) \\ p||h \\ p \nmid t}} \left(1 - \frac{1}{p} \right)^{-1} \prod_{\substack{p|n \\ p \nmid h}} \left(1 - \frac{1}{p(p-1)} \right)^{-1},$$

for

$$F(h, t) = \frac{1}{\varphi(h)} \prod_{p \nmid ht} \left(1 - \frac{1}{p(p-1)} \right) \prod_{\substack{p||h \\ p|r \\ p \nmid t}} \left(1 - \frac{1}{p} \right).$$

Again, writing the sum over n as an Euler product, we get that

$$C_t = F(h, t) \prod_p \left(1 + \delta(p) \sum_{j \geq 1} \frac{c_t(p^j)\varphi((p^j, h))}{p^j \varphi(p^j)} \right), \quad (26)$$

where

$$\delta(p) = \begin{cases} \left(1 - \frac{1}{p(p-1)} \right)^{-1} & \text{when } p \nmid ht; \\ \left(1 - \frac{1}{p} \right)^{-1} & \text{when } p \parallel h, p \mid r, p \nmid t; \\ 1 & \text{otherwise.} \end{cases}$$

Finally, we use Lemma 9 to note that when $p|h$, $c_t(p^j) \frac{\varphi((p^j, h))}{\varphi(p^j)} = \left(\frac{r}{p}\right)^j$, independently of t . For primes $p \nmid h$, we have that $\varphi((p^j, h)) = 1$. Hence, by splitting the inner sum in (26) into odd and even terms, we get

$$\sum_{j \geq 1} \frac{c_t(p^j)}{p^j} \frac{\varphi((p^j, h))}{\varphi(p^j)} = \frac{p}{(p-1)(p^2-1)} \left(p \frac{c_t(p)}{p} + \frac{c_t(p^2)}{p^2} \right).$$

The result now follows by considering the different cases in Lemma 9. ■

We now proceed to prove Theorem 8. In the following, all primes p are odd and all products are restricted to odd primes. Let us call G the multiplicative function with value at a prime p given by $G(p) = \frac{2p^2 + p - 1}{p^4 - p^3 - 2p^2 - p + 1}$. Using (23) and Lemma 10, we have that

$$\begin{aligned} S(T) &= P(r, h) \sum_{\substack{t \leq T \\ (t^2 - r, h) = 1}}^{\text{odd}} \sum_{\substack{d|t \\ (d, h) = 1}} \mu^2(d) G(d) + O\left(\frac{T \log R}{R}\right) \\ &= P(r, h) \sum_{\substack{d \leq T \\ (d, h) = 1}} \mu^2(d) G(d) \sum_{\substack{t \leq T/d \\ (d^2 t^2 - r, h) = 1}}^{\text{odd}} 1 + O\left(\frac{T \log R}{R}\right). \end{aligned} \quad (27)$$

To compute the inner sum, we write

$$\sum_{\substack{t \leq T/d \\ (d^2 t^2 - r, h) = 1}}^{\text{odd}} 1 = \sum_{t \leq T/d}^{\text{odd}} \sum_{k|(d^2 t^2 - r, h)} \mu(k) = \sum_{k|h} \mu(k) \sum_{\substack{t \leq T/d \\ t^2 \equiv r d^{-2} \pmod{k} \\ t \equiv 1 \pmod{2}}} 1. \quad (28)$$

Because of the Moebius function, we can suppose that k is square-free. In that case, by the Chinese Remainder Theorem, the number of solutions modulo $2k$ to the congruences $t^2 \equiv r d^{-2} \pmod{k}$ and $t \equiv 1 \pmod{2}$ is

$$\prod_{p|k} \left(1 + \left(\frac{r}{p}\right) \right).$$

Using that in (28), we have that

$$\sum_{\substack{t \leq T/d \\ (d^2 t^2 - r, h) = 1}}^{\text{odd}} 1 = \frac{T}{2d} \left(\prod_{p|h} \left(1 + \frac{1 + \left(\frac{r}{p}\right)}{p} \right) \right) + O(\tau(h) 2^{\omega(h)}). \quad (29)$$

Now we observe that

$$\sum_{\substack{d \leq T \\ (d, h) = 1}} \frac{\mu^2(d) G(d)}{d} = \prod_{p \nmid h} \left(1 + \frac{G(p)}{p} \right) + O\left(\frac{1}{T^{2-\varepsilon}}\right). \quad (30)$$

Using (29) and (30) in (27) completes the proof of Theorem 8 with

$$\begin{aligned} \mathfrak{c} &= \frac{P(r, h)}{3} \prod_{p|h} \left(1 - \frac{1 + \left(\frac{r}{p}\right)}{p} \right) \prod_{p \nmid h} \left(1 + \frac{G(p)}{p} \right) \\ &= \frac{1}{2h} \prod_{\substack{p|h \\ p|r}} \frac{p-1}{p} \prod_{\substack{p|h \\ p \nmid r}} \frac{p(p-1 - \left(\frac{r}{p}\right))}{(p-1)(p - \left(\frac{r}{p}\right))} \prod_{p \nmid h} \frac{p^4 - 2p^2 - p + 1}{p^2(p^2 - 1)}. \end{aligned}$$

■

5 Conjectures and Constants

We write in this section a precise conjecture for the asymptotic behavior of

$$\Pi_{E, r, h}^{\text{sf}}(x) = \#\{2 < p \leq x, \text{ prime} : a_p^2 - 4p \in \Delta(r, h)\}.$$

Using the Moebius function to detect square-free numbers, we write

$$\begin{aligned} \Pi_{E, r, h}^{\text{sf}}(x) &= \sum_{\substack{p \leq x \\ a_p^2 - 4p \equiv r \pmod{h}}} \sum_{d^2 | a_p^2 - 4p} \mu(d) \\ &= \sum_{d \leq 2\sqrt{x}} \mu(d) \pi_{d^2}(x; r, h) \end{aligned} \tag{31}$$

where

$$\pi_{d^2}(x; r, h) = \sum_{\substack{p \leq x \\ d^2 | a_p^2 - 4p \\ a_p^2 - 4p \equiv r \pmod{h}}} 1.$$

To evaluate $\pi_{d^2}(x; r, h)$, we use the Chebotarev Density Theorem on the field extension $\mathbb{Q}(E[m])/ \mathbb{Q}$ obtained by adjoining the m -torsion points of E to \mathbb{Q} where $m = [d^2, h]$ is the least common multiple of d^2 and h . Since $E[m] \simeq \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$ as an abstract group, by considering the Galois action on a basis for the m -torsion, we get a Galois representation

$$\rho_m : \text{Gal}(\mathbb{Q}(E[m])/ \mathbb{Q}) \rightarrow \text{GL}_2(\mathbb{Z}/m\mathbb{Z})$$

which sends the Frobenius element σ_p of $\text{Gal}(\mathbb{Q}(E[m])/ \mathbb{Q})$ to a conjugacy class of matrices such that, for each g in the conjugacy class, we have

$$\begin{aligned} \text{tr } g &\equiv a_p \pmod{m}, \\ \det g &\equiv p \pmod{m}. \end{aligned}$$

Then, by applying the Chebotarev Density Theorem to $\text{Gal}(\mathbb{Q}(E[m])/ \mathbb{Q})$, we can evaluate $\pi_{d^2}(x; r, h)$ as desired. Let us set some notation. Let $G_E([d^2, h]) \subseteq \text{GL}_2(\mathbb{Z}/[d^2, h]\mathbb{Z})$ be the image of $\rho_{[d^2, h]}$ and let

$$C_E(d, h, r) = \{g \in G_E([d^2, h]) : \text{tr}^2 g - 4 \det g \equiv 0 \pmod{d^2}, \text{tr}^2 g - 4 \det g \equiv r \pmod{h}\}.$$

Using the Chebotarev Density Theorem under the Generalized Riemann Hypothesis (see for example [Se81, p. 133]), it follows that

$$\begin{aligned}\pi_{d^2}(x; r, h) &= \#\{p \leq x : \sigma_p \in C_E(d, h, r)\} \\ &= \frac{|C_E(d, h, r)|}{|G_E([d^2, h])|} \pi(x) + O\left(|C_E(d, h, r)| x^{1/2} \log(dx)\right)\end{aligned}\quad (32)$$

where the error term depends on E and h .

In order to use (32) in (31) to estimate $\Pi_{E,r,h}^{\text{sf}}(x)$, we need to control the error term which should be $o(\pi(x))$. This seems to be out of reach with the present techniques, given the size of the error term in (32). There are some known techniques to improve the error term of (32) by replacing $|C_E(d, h, r)|$ by $|C_E(d, h, r)|^{1/2}$ [MuMuSa], but even such an improved error term will not allow to evaluate (31). We remark that in the result of Cojocaru and Duke [CoDu], the Chebotarev Density Theorem is used to count primes splitting completely in some extensions depending on E , where the error term is then much smaller.

We now concentrate on the conjectural main term of $\Pi_{E,r,h}^{\text{sf}}(x)$,

$$\pi(x) \sum_{d=1}^{\infty} \mu(d) \frac{|C_E(d, r, h)|}{|G_E([d^2, h])|},$$

obtained by using the main term of (32) in (31) and the convergence of the sum, proved below.

Using properties of the Galois representations $\rho_{[d^2, h]}$ associated to E we will rewrite the constant as an Euler product. For any curve without complex multiplication (CM), Serre proved in [Se72] that the image of the Galois representation associated to all the torsion points of E is open in $\text{GL}_2(\hat{\mathbb{Z}})$. It follows that for each E without CM, there is an integer M_E such that for any n with $(n, M_E) = 1$, we have

$$G_E(n) = \text{GL}_2(\mathbb{Z}/n\mathbb{Z}), \quad (33)$$

and for any n, n' with $(n, M_E) = (n', M_E) = 1$, we have

$$G_E(nn') = G_E(n) \times G_E(n'). \quad (34)$$

We remark that the integer M_E is always even. Using (34) and the Chinese Remainder Theorem and noting that the sum is restricted to squarefree numbers d , we write

$$\begin{aligned}\sum_d \mu(d) \frac{|C_E(d, r, h)|}{|G_E([d^2, h])|} &= \sum_{k|[h, M_E]} \sum_{(d, [h, M_E])=k} \mu(d) \frac{|C_E(d, r, h)|}{|G_E([d^2, h])|} \\ &= \sum_{k|[h, M_E]} \mu(k) \frac{|C_E(k, r, h)|}{|G_E([k^2, h])|} \sum_{(j, M_E)=(j, h)=1} \mu(j) \frac{|C_E(j^2)|}{|G_E(j^2)|}\end{aligned}\quad (35)$$

where

$$C_E(j^2) = \{g \in G_E(j^2) : \text{tr}^2 g - 4 \det g \equiv 0 \pmod{j^2}\}.$$

Since $(j, M_E) = 1$, $G_E(j^2) = \text{GL}_2(\mathbb{Z}/j^2\mathbb{Z})$ by (33), and we get that

$$\sum_{(j, M_E)=(j, h)=1} \mu(j) \frac{|C_E(j^2)|}{|G_E(j^2)|} = \prod_{\ell \nmid h M_E} \frac{\ell^4 - 2\ell^2 - \ell + 1}{\ell^2(\ell^2 - 1)}. \quad (36)$$

We now look at the first sum of (35). We write $h = h_1 h_2$ where $p \mid h_1 \Rightarrow p \mid M_E$ and $p \mid h_2 \Rightarrow p \nmid M_E$. Then, $[h, M_E] = [h_1, M_E] h_2$. Similarly, for any $k \mid [h, M_E]$, we write $k = k_1 k_2$ where $p \mid k_1 \Rightarrow p \mid M_E$ and $p \nmid k_2 \Rightarrow p \nmid M_E$. Using (34) and the Chinese Remainder Theorem, we get that

$$\sum_{k \mid [h, M_E]} \mu(k) \frac{|C_E(k, r, h)|}{|G_E([k^2, h])|} = \sum_{k_2 \mid h_2} \mu(k_2) \frac{|C_E(k_2, r, h_2)|}{|G_E([k_2^2, h_2])|} \sum_{k_1 \mid [h_1, M_E]} \mu(k_1) \frac{|C_E(k_1, r, h_1)|}{|G_E([k_1^2, h_1])|}.$$

Let $h_2 = \prod_{\ell} \ell^{\alpha(\ell)}$, $\beta(\ell) = \max(\alpha(\ell), 2)$ and

$$C_E(\ell^{\alpha(\ell)}, r) = \left\{ g \in G_E(\ell^{\alpha(\ell)}) : \text{tr } g^2 - 4 \det g \equiv r \pmod{\ell^{\alpha(\ell)}} \right\}.$$

Then,

$$\frac{|C_E(k_2, r, h_2)|}{|G_E([k_2^2, h_2])|} = \prod_{\ell \mid h_2} \frac{|C_E(\ell^{\alpha(\ell)}, r)|}{|G_E(\ell^{\alpha(\ell)})|} \prod_{\ell \mid k_2} \frac{|C_E(\ell, r, \ell^{\alpha(\ell)})|}{|C_E(\ell^{\alpha(\ell)}, r)|} \frac{|G_E(\ell^{\alpha(\ell)})|}{|G_E(\ell^{\beta(\ell)})|},$$

and

$$\begin{aligned} \sum_{k_2 \mid h_2} \mu(k_2) \frac{|C_E(k_2, r, h_2)|}{|G_E([k_2^2, h_2])|} &= \prod_{\ell \mid h_2} \frac{|C_E(\ell^{\alpha(\ell)}, r)|}{|G_E(\ell^{\alpha(\ell)})|} - \frac{|C_E(\ell, r, \ell^{\alpha(\ell)})|}{|G_E(\ell^{\beta(\ell)})|} \\ &= \prod_{\ell \mid h_2} \frac{|D_E(\ell, r, \ell^{\alpha(\ell)})|}{|G_E(\ell^{\beta(\ell)})|} \end{aligned} \quad (37)$$

where

$$D_E(\ell, r, \ell^{\alpha(\ell)}) = \left\{ g \in G(\ell^{\beta(\ell)}) : \text{tr}(g)^2 - 4 \det(g) \not\equiv 0 \pmod{\ell^2} \text{ and } \text{tr}(g)^2 - 4 \det(g) \equiv r \pmod{\ell^{\alpha(\ell)}} \right\}.$$

It is a lengthy, but straightforward computation, to show that for any odd prime ℓ and any integer $\alpha \geq 1$,

$$\frac{|D_E(\ell, r, \ell^{\alpha})|}{|G_E(\ell^{\max(\alpha, 2)})|} = \begin{cases} \frac{\ell(\ell - 1 - (\frac{r}{\ell}))}{(\ell - 1)(\ell - (\frac{r}{\ell}))\ell^{\alpha}} & \text{when } \ell \nmid r; \\ \frac{\ell - 1}{\ell^{\alpha+1}} & \text{when } \ell \mid r \text{ and } \alpha = 1; \\ \frac{1}{\ell^{\alpha}} & \text{when } \ell \parallel r \text{ and } \alpha \geq 2; \\ 0 & \text{when } \ell^2 \mid r \text{ and } \alpha \geq 2. \end{cases}$$

The last case ($\ell^2 \mid r$ and $\alpha \geq 2$) happens only when $\ell^2 \mid (r, h)$, which is impossible when (r, h) is square-free. If (r, h) is not square-free, it is clear that $\Pi_{E,r,h}^{\text{sf}}(x)$ is empty.

Replacing (36) and (37) in (35), we get that

$$\begin{aligned} \sum_d \mu(d) \frac{|C_E(d, r, h)|}{|G_E([d^2, h])|} &= \sum_{k_1 \mid [h_1, M_E]} \mu(k_1) \frac{|C_E(k_1, r, h_1)|}{|G_E([k_1^2, h_1])|} \times \\ &\frac{1}{h} \prod_{\substack{\ell \mid M_E \\ \ell \mid h \\ \ell \mid r}} \frac{\ell - 1}{\ell} \prod_{\substack{\ell \mid M_E \\ \ell \mid h \\ \ell \mid r}} \frac{\ell(\ell - 1 - (\frac{r}{\ell}))}{(\ell - 1)(\ell - (\frac{r}{\ell}))} \prod_{\substack{\ell \mid M_E \\ \ell \nmid h}} \frac{\ell^4 - 2\ell^2 - \ell + 1}{\ell^2(\ell^2 - 1)}. \end{aligned} \quad (38)$$

Conjecture 11 *Let E be an elliptic curve over \mathbb{Q} without complex multiplication, and let M_E be an integer satisfying (33) and (34). Let h be a positive odd integer, and r an integer such that (h, r) is square-free. Write $h = h_1 h_2$ where $p \mid h_1 \Rightarrow p \mid M_E$ and $p \mid h_2 \Rightarrow p \nmid M_E$. Then as $x \rightarrow \infty$*

$$\Pi_{E(a,b),r,h}^{\text{sf}}(x) \sim C_{SF}(E, r, h) \frac{x}{\log x},$$

where $C_{SF}(E, r, h)$ is the constant given by (38).

Then, Theorem 2 gives evidence for the asymptotic of Conjecture 11, and also for the constant $C_{SF}(E, r, h)$ appearing on the asymptotic. We remark that the average constant \mathfrak{C} of Theorem 2 corresponds to a curve with $M_E = 1$, even if there are no such curves over \mathbb{Q} , as 2 always divides M_E . This also explains the factor $1/3$ which appears in the average constant \mathfrak{C} since

$$\frac{|C_E(4)|}{|\text{GL}_2(\mathbb{Z}/4\mathbb{Z})|} = \frac{1}{3}.$$

A similar heuristic leads to the following conjecture for the number of primes p smaller than x such that $p + 1 - a_p(E)$ is square-free.

Conjecture 12 *Let E be an elliptic curve over \mathbb{Q} without complex multiplication, and let M_E be an integer satisfying (33) and (34). Let $\Pi_E(x)$ be the number of primes of good reduction such that $p + 1 - a_p(E)$ is square-free. Then as $x \rightarrow \infty$*

$$\Pi_E(x) \sim C'_{SF}(E) \frac{x}{\log x},$$

where $C'_{SF}(E)$ is the constant

$$C'_{SF}(E) = \sum_{k \mid M_E} \mu(k) \frac{|C'_E(k^2)|}{|G_E(k^2)|} \prod_{\ell \mid M_E} 1 - \frac{\ell^3 - \ell - 1}{(\ell^2 - 1)\ell^2(\ell - 1)},$$

where $G_E(k^2) \subseteq \text{GL}_2(\mathbb{Z}/k^2\mathbb{Z})$ is the image of ρ_{k^2} and

$$C'_E(k^2) = \{g \in G_E(k^2) : \det(g) + 1 - \text{tr}(g) \equiv 0 \pmod{k^2}\}.$$

The work of Gekeler mentioned in the introduction then gives evidence for Conjecture 12, and in particular for the constant $C'_{SF}(E)$ appearing in the conjecture. We stress again that the techniques used by Gekeler and our techniques are completely different, but they

both uncover the conjectural constants computed from Galois representations in that section (and of course, neither the work of Gekeler or the work of this present paper relies on Galois representations, and counting elements of Galois groups).

Acknowledgements: Both authors would like to thank Nathan Jones for useful discussions during the preparation of this paper. This work was done during the visit of the second author at the Centre de Recherches Mathématiques (CRM), and he thanks the CRM for providing an excellent working environment.

References

- [BaCoDa] A. Balog, A. Cojocaru and C. David, Average twin prime conjecture for elliptic curves, submitted, 2008.
- [BaSh] B. Banks and I. Shparlinski, Sato–Tate, cyclicity, and divisibility statistics for elliptic curves of small height, *Israel J. Math.*, to appear.
- [Co] A.C. Cojocaru, Squarefree orders for CM elliptic curves modulo p , *Math. Annalen*, to appear.
- [CoDa] C. David and A.C. Cojocaru, Frobenius fields for Elliptic Curves, *Amer. J. Math.*, to appear.
- [CoDu] A.C. Cojocaru and W. Duke, Reductions of an elliptic curve and their Tate–Shafarevich groups, *Math. Ann.* **329**, 2004, 513–534.
- [CoFoMu] A.C. Cojocaru, E. Fouvry and M.R. Murty, The square sieve and the Lang–Trotter conjecture, *Canad. J. Math.* **57**, 2005, 1155–1177.
- [DaPa] C. David and F. Pappalardi, Average Frobenius distributions of elliptic curves, *Internat. Math. Res. Notices* **4**, 1999, 165–183.
- [De] M. Deuring, Die Typen der Multiplikatorenringe elliptischer Funktionenkörper, *Hamb. Abh.*, 1941, 197–272.
- [DuTo] W. Duke and Á. Tóth, The splitting of primes in division fields of elliptic curves, *Experiment. Math.* **11**, 2003, 555–565.
- [FoMu] E. Fouvry and R. Murty, On the distribution of supersingular primes, *Canad. J. Math.* **48**, 1996, 81–104.
- [FrIw] J. B. Friedlander and H. Iwaniec, The divisor problem for arithmetic progressions, *Acta Arith.* **45**, 1985, 273–277.
- [Ge] E.-U. Gekeler, Statistics about elliptic curves over finite prime fields, *Manuscripta Math.*, to appear.

- [Ho] E. W. Howe, On the group orders of elliptic curves over finite fields, *Comp. Math.* **85**, 1993, 229-247.
- [Ko] N. Koblitz, Primality of the number of points on an elliptic curve over a finite field, *Pacific J. Math.* **131**, 1988 (no. 1), 157–165.
- [LaTr] S. Lang, H. Trotter, Frobenius distributions in GL_2 -extensions, Lecture Notes in Mathematics **504**, Springer Verlag, 1976.
- [MuMuSa] M. R. Murty, V. K. Murty and N. Saradha, Modular Forms and the Chebotarev Density Theorem, *Amer. J. Math.* **110**, 1988, 253–281.
- [Se72] J.-P. Serre, Propriétés galoisiennes des points d’ordre fini des courbes elliptiques, *Invent. Math.* **15**, 1972, 259–331.
- [Se81] J.-P. Serre, Quelques applications du théorème de densité de Chebotarev, *Publications Mathématiques I.H.E.S.*, no. 54, 1981, 123–201.

[Chantal David] Department of Mathematics and Statistics, Concordia University, 1455 de Maisonneuve West, Montréal, QC, Canada. *Email:* c david@mathstat.concordia.ca

[Jorge Jimenez Urroz] Universitat Politècnica de Catalunya, Campus Nord, edifici C3, C. Jordi Girona, 1-3, 08034 Barcelona, Spain. *Email:* jjimenez@ma4.upc.edu