

On group structures realized by elliptic curves over arbitrary finite fields

WILLIAM D. BANKS

Department of Mathematics, University of Missouri
Columbia, MO 65211 USA
`bankswd@missouri.edu`

FRANCESCO PAPPALARDI

Dipartimento di Matematica, Università Roma Tre
Roma, I-00146, Italy
`pappa@mat.uniroma3.it`

IGOR E. SHPARLINSKI

Department of Computing, Macquarie University
Sydney, NSW 2109, Australia
`igor@ics.mq.edu.au`

November 2, 2018

Abstract

We study the collection of group structures that can be realized as a group of rational points on an elliptic curve over a finite field (such groups are well known to be of rank at most two). We also study various subsets of this collection which correspond to curves over prime fields or to curves with a prescribed torsion. Some of our results are rigorous and are based on recent advances in analytic number theory, some are conditional under certain widely believed conjectures, and others are purely heuristic in nature.

1 Introduction

Let \mathbb{F}_q denote the finite field with q elements. It is well known that the group $E(\mathbb{F}_q)$ of points on an elliptic curve E defined over \mathbb{F}_q has rank at most two, and therefore,

$$E(\mathbb{F}_q) \cong \mathbb{Z}_n \times \mathbb{Z}_{kn} \quad (1)$$

for some natural numbers n and k , where \mathbb{Z}_m denotes the ring of congruence classes modulo m for each natural number m ; see [7, 13, 21, 23]. On the other hand, little is known about the structure of the set of groups $\mathbb{Z}_n \times \mathbb{Z}_{kn}$ that can be realized as the group of points on an elliptic curve defined over a finite field. For this reason, we introduce and investigate the set

$$\mathcal{S}_{\Pi} = \{(n, k) \in \mathbb{N}^2 : \exists \text{ prime power } q \text{ and } E/\mathbb{F}_q \text{ with } E(\mathbb{F}_q) \cong \mathbb{Z}_n \times \mathbb{Z}_{kn}\}.$$

We are also interested in groups $\mathbb{Z}_n \times \mathbb{Z}_{kn}$ with a realization (1) in which $q = p$ is a prime number, hence we study the subset $\mathcal{S}_{\pi} \subset \mathcal{S}_{\Pi}$ defined by

$$\mathcal{S}_{\pi} = \{(n, k) \in \mathbb{N}^2 : \exists \text{ prime } p \text{ and } E/\mathbb{F}_p \text{ with } E(\mathbb{F}_p) \cong \mathbb{Z}_n \times \mathbb{Z}_{kn}\}.$$

Although one can expect \mathcal{S}_{π} and \mathcal{S}_{Π} to be reasonably “dense” in \mathbb{N}^2 , the complementary sets also appear to be rather large. For example, here is the list of pairs $(n, k) \notin \mathcal{S}_{\Pi}$ with $n, k \leq 25$:

$$\begin{aligned} & (11, 1), (11, 14), (13, 6), (13, 25), (15, 4), \\ & (19, 7), (19, 10), (19, 14), (19, 15), (19, 18), \\ & (21, 18), (23, 1), (23, 5), (23, 8), (23, 19), (25, 5), (25, 14). \end{aligned} \quad (2)$$

To investigate the distribution in \mathbb{N}^2 of the elements of \mathcal{S}_{π} and of \mathcal{S}_{Π} , for natural numbers N and K we introduce the sets

$$\begin{aligned} \mathcal{S}_{\pi}(N, K) &= \{(n, k) \in \mathcal{S}_{\pi} : n \leq N, k \leq K\}, \\ \mathcal{S}_{\Pi}(N, K) &= \{(n, k) \in \mathcal{S}_{\Pi} : n \leq N, k \leq K\}. \end{aligned}$$

These sets are the main objects of study in this note.

For natural numbers n and k , we also put

$$\mathcal{P}(n, k) = \{\text{primes } p : \exists E/\mathbb{F}_p \text{ for which } E(\mathbb{F}_p) \cong \mathbb{Z}_n \times \mathbb{Z}_{kn}\}.$$

The set $\mathcal{P}(n, k)$ parametrizes the set of finite fields of prime cardinality over which $\mathbb{Z}_n \times \mathbb{Z}_{kn}$ can be realized as the group of points on an elliptic curve. For natural numbers N and K we study the double sum

$$\mathcal{N}_{\mathcal{P}}(N, K) = \sum_{n \leq N} \sum_{k \leq K} \#\mathcal{P}(n, k),$$

for which we obtain an asymptotic formula in certain ranges.

Finally, for natural numbers m, k we introduce and compare the sets

$$\begin{aligned} \mathcal{N}_{m,k} &= \{n \in \mathbb{N} : \exists p \text{ prime and } E/\mathbb{F}_{p^m} \text{ with } E(\mathbb{F}_{p^m}) \cong \mathbb{Z}_n \times \mathbb{Z}_{kn}\}, \\ \tilde{\mathcal{N}}_{m,k} &= \{n \in \mathbb{N} : \exists p \text{ prime, } \ell \in \mathbb{Z} \text{ with } p^m = kn^2 + \ell n + 1, |\ell| \leq 2\sqrt{k}\}. \end{aligned}$$

We remark that the distribution of group structures generated by elliptic curves over a *fixed* finite field \mathbb{F}_q has been studied in [12].

2 Notational conventions

Throughout the paper, the letter p always denotes a prime number, and q always denotes a prime power. As usual, we use $\pi(x)$ to denote the number of $p \leq x$. For coprime integers a and $m \geq 1$, we put

$$\begin{aligned} \pi(x; m, a) &= \#\{p \leq x : p \equiv a \pmod{m}\}, \\ \Pi(x; m, a) &= \#\{q \leq x : q \equiv a \pmod{m}\}. \end{aligned}$$

We also set

$$\psi(x; m, a) = \sum_{\substack{n \leq x \\ n \equiv a \pmod{m}}} \Lambda(n),$$

where $\Lambda(n)$ is the von Mangoldt function.

For any set $\mathcal{A} \subseteq \mathbb{N}$ and real $x > 0$, we denote $\mathcal{A}(x) = \{a \in \mathcal{A} : a \leq x\}$.

For functions F and $G > 0$ the notations $F = O(G)$, $F \ll G$, and $G \gg F$ are all equivalent to the assertion that the inequality $|F| \leq cG$ holds with some constant $c > 0$. In what follows, all constants implied by the symbols O , \ll , and \gg may depend (where obvious) on the small real parameter ε but are absolute otherwise; we write O_ρ , \ll_ρ , and \gg_ρ to indicate that the implied constant depends on a given parameter ρ .

3 Preliminaries

Lemma 1. *If q is a prime power, and E is an elliptic curve defined over \mathbb{F}_q such that $E(\mathbb{F}_q) \cong \mathbb{Z}_n \times \mathbb{Z}_{kn}$, then $q = kn^2 + \ell n + 1$ for some integer ℓ that satisfies $|\ell| \leq 2\sqrt{k}$.*

Proof. By the Hasse bound, we can write $kn^2 = q + 1 - a$ for some integer a that satisfies the bound $a^2 \leq 4q$. Using the Weil pairing one also sees that $q \equiv 1 \pmod{n}$, hence $a = \ell n + 2$ for some integer ℓ , and we have $q = kn^2 + \ell n + 1$. Since

$$\ell^2 n^2 + 4\ell n + 4 = (\ell n + 2)^2 = a^2 \leq 4q = 4kn^2 + 4\ell n + 4,$$

it follows that $|\ell| \leq 2\sqrt{k}$ as required. \square

The following result of Waterhouse [23] (see also [22, Theorems 4.3]) is a characterization of the natural numbers N that can be realized as the cardinality of the group of \mathbb{F}_q -rational points on an elliptic curve E defined over \mathbb{F}_q .

Lemma 2. *Let $q = p^m$ be a prime power, and suppose that $N = q + 1 - a$ for some integer a . Then, there is an elliptic curve E defined over \mathbb{F}_q such that $\#E(\mathbb{F}_q) = N$ if and only if $|a| \leq 2\sqrt{q}$ and one of the following conditions is met:*

- (i) $\gcd(a, p) = 1$;
- (ii) m even and $a = \pm 2\sqrt{q}$;
- (iii) m is even, $p \not\equiv 1 \pmod{3}$, and $a = \pm\sqrt{q}$;
- (iv) m is odd, $p = 2$ or 3 , and $a = \pm p^{(m+1)/2}$;
- (v) m is even, $p \not\equiv 1 \pmod{4}$, and $a = 0$;
- (vi) m is odd and $a = 0$.

For every admissible cardinality N , the following result of Rück [13] (see also [22, Theorems 4.4]) describes the group structures that are possible for $E(\mathbb{F}_q)$ given that $\#E(\mathbb{F}_q) = N$; see also [7, 21].

Lemma 3. *Let $q = p^m$ be a prime power, and suppose that N is an integer such that $\#E(\mathbb{F}_q) = N$ for some elliptic curve E defined over \mathbb{F}_q . Write $N = p^e n_1 n_2$ with $p \nmid n_1 n_2$ and $n_1 \mid n_2$ (possibly $n_1 = 1$). Then, there is an elliptic curve E over \mathbb{F}_q for which*

$$E(\mathbb{F}_q) \cong \mathbb{Z}_{p^e} \times \mathbb{Z}_{n_1} \times \mathbb{Z}_{n_2}$$

if and only if

- (i) $n_1 = n_2$ in case (ii) of Lemma 2;
- (ii) $n_1 \mid q - 1$ in all other cases of Lemma 2.

Combining Lemmas 2 and 3, we get:

Corollary 4. *If p is prime and $N \in \mathbb{N}$ with $|p + 1 - N| \leq 2\sqrt{p}$, then there is an elliptic curve E defined over \mathbb{F}_p with $\#E(\mathbb{F}_p) = N$. In this case, if we write $N = n_1 n_3$ with $p \nmid n_1$ and $n_1 \mid n_3$ (possibly $n_1 = 1$), then $n_1 \mid p - 1$ and $E(\mathbb{F}_p) \cong \mathbb{Z}_{n_1} \times \mathbb{Z}_{n_3}$.*

Lemma 5. *A prime p lies in $\mathcal{P}(n, k)$ if and only if $p = kn^2 + \ell n + 1$ for some integer ℓ such that $|\ell| \leq 2\sqrt{k}$.*

Proof. By definition, if p lies in $\mathcal{P}(n, k)$ then there is an elliptic curve E/\mathbb{F}_p such that $E(\mathbb{F}_p) \cong \mathbb{Z}_n \times \mathbb{Z}_{kn}$. According to Lemma 1, $p = kn^2 + \ell n + 1$ with some integer ℓ such that $|\ell| \leq 2\sqrt{k}$.

Conversely, suppose that $p = kn^2 + \ell n + 1$ and $|\ell| \leq 2\sqrt{k}$. Taking $N = kn^2$ we have

$$|p + 1 - N|^2 = (\ell n + 2)^2 = \ell^2 n^2 + 4\ell n + 4 \leq 4kn^2 + 4\ell n + 4 = 4p,$$

hence $|p + 1 - N| \leq 2\sqrt{p}$. Applying Corollary 4 with $n_1 = n$ and $n_3 = kn$, we see that there is an elliptic curve E/\mathbb{F}_p such that $E(\mathbb{F}_p) \cong \mathbb{Z}_n \times \mathbb{Z}_{kn}$, and thus $p \in \mathcal{P}(n, k)$. \square

Next, we relate $\mathcal{N}_{\mathcal{P}}(N, K)$ to the distribution of primes in short arithmetic progressions.

Lemma 6. *For all $N, K \in \mathbb{N}$ we have*

$$\mathcal{N}_{\mathcal{P}}(N, K) = \sum_{\substack{n \leq N \\ |\ell| \leq 2\sqrt{K}}} \left(\pi(Kn^2 + \ell n + 1; n^2, \ell n + 1) - \pi\left(\frac{1}{4}\ell^2 n^2 + \ell n + 1; n^2, \ell n + 1\right) \right).$$

Proof. Fix $n \leq N$, and let $\mathcal{T}_1(n)$ be the collection of pairs (ℓ, p) such that

$$(i) \quad |\ell| \leq 2\sqrt{K};$$

$$(ii) \quad p \text{ is a prime congruent to } \ell n + 1 \pmod{n^2};$$

$$(iii) \quad \frac{1}{4}\ell^2 n^2 + \ell n + 1 \leq p \leq K n^2 + \ell n + 1.$$

Since $\frac{1}{4}\ell^2 n^2 + \ell n + 1 = (\frac{1}{2}\ell n + 1)^2$ cannot be prime, it is easy to see that

$$\#\mathcal{T}_1(n) = \sum_{|\ell| \leq 2\sqrt{K}} \left(\pi(K n^2 + \ell n + 1; n^2, \ell n + 1) - \pi\left(\frac{1}{4}\ell^2 n^2 + \ell n + 1; n^2, \ell n + 1\right) \right).$$

Let $\mathcal{T}_2(n)$ be the collection of pairs (k, p) such that

$$(iv) \quad k \leq K;$$

$$(v) \quad p \text{ is prime and } p = k n^2 + \ell n + 1 \text{ for some integer } \ell \text{ such that } |\ell| \leq 2\sqrt{k}.$$

By Lemma 5, condition (v) is equivalent to the assertion that $p \in \mathcal{P}(n, k)$, hence

$$\#\mathcal{T}_2(n) = \sum_{k \leq K} \#\mathcal{P}(n, k).$$

Since

$$\sum_{n \leq N} \#\mathcal{T}_1(n) = \sum_{\substack{n \leq N \\ |\ell| \leq 2\sqrt{K}}} \left(\pi(K n^2 + \ell n + 1; n^2, \ell n + 1) - \pi\left(\frac{1}{4}\ell^2 n^2 + \ell n; n^2, \ell n + 1\right) \right)$$

and

$$\sum_{n \leq N} \#\mathcal{T}_2(n) = \sum_{n \leq N} \sum_{k \leq K} \#\mathcal{P}(n, k) = \mathcal{N}_{\mathcal{P}}(N, K),$$

to prove the lemma it suffices to show that $\#\mathcal{T}_1(n) = \#\mathcal{T}_2(n)$ for each $n \leq N$.

First, let $(\ell, p) \in \mathcal{T}_1(n)$. By (ii) we can write $p = k n^2 + \ell n + 1$ for some integer k . Substituting into (iii) we have

$$\frac{1}{4}\ell^2 n^2 + \ell n + 1 \leq k n^2 + \ell n + 1 \leq K n^2 + \ell n + 1,$$

hence $k \leq K$ and $|\ell| \leq 2\sqrt{k}$. This shows that the pair (k, p) lies in $\mathcal{T}_2(n)$. As the map $\mathcal{T}_1(n) \rightarrow \mathcal{T}_2(n)$ given by $(\ell, p) \mapsto (k, p)$ is clearly injective, we have $\#\mathcal{T}_1(n) \leq \#\mathcal{T}_2(n)$.

Next, suppose that $(k, p) \in \mathcal{T}_2(n)$, and let ℓ be as in (v). By (iv) we have $|\ell| \leq 2\sqrt{k} \leq 2\sqrt{K}$, and $p \equiv \ell n + 1 \pmod{n^2}$ by (v). Furthermore, since $\frac{1}{4}\ell^2 \leq k \leq K$ the prime $p = kn^2 + \ell n + 1$ satisfies (iii). This shows that the pair (ℓ, p) lies in $\mathcal{T}_1(n)$. Since the map $\mathcal{T}_2(n) \rightarrow \mathcal{T}_1(n)$ given by $(k, p) \mapsto (\ell, p)$ is injective, we have $\#\mathcal{T}_2(n) \leq \#\mathcal{T}_1(n)$, and the proof is complete. \square

4 Primes in sparse progressions

Below, we use the following result of Baier and Zhao [2], which is a variant of the Bombieri-Vinogradov theorem that deals with primes in arithmetic progressions to square moduli.

Lemma 7. *For fixed $\varepsilon > 0$ and $C > 0$ we have*

$$\sum_{m \leq x^{2/9-\varepsilon}} m \max_{\gcd(a,m)=1} \left| \psi(x; m^2, a) - \frac{x}{\varphi(m^2)} \right| \ll \frac{x}{(\log x)^C},$$

where the implied constant depends only on ε and C .

Via partial summation one obtains the following:

Corollary 8. *For fixed $\varepsilon > 0$ and $C > 0$ we have*

$$\sum_{m \leq x^{2/9-\varepsilon}} m \max_{\gcd(a,m)=1} \left| \pi(x; m^2, a) - \frac{\pi(x)}{\varphi(m^2)} \right| \ll \frac{x}{(\log x)^C},$$

where the implied constant depends only on ε and C .

For our applications of Corollary 8 we also need a well known asymptotic formula

$$\sum_{n \leq X} \frac{n}{\varphi(n)} = \frac{315 \zeta(3)}{2\pi^4} X + O(\log X); \quad (3)$$

for more precise results, we refer the reader to [11, 17, 18].

For any sequence of integers $\mathcal{A} = (a_n)_{n=1}^{\infty}$ and any positive real numbers λ and X , we define the sum

$$\mathcal{P}(\mathcal{A}; \lambda, X) = \sum_{n \leq X} \pi(\lambda n^2; n^2, a_n). \quad (4)$$

Lemma 9. Fix $\varepsilon \in (0, 2/5)$. For any sequence of integers $\mathcal{A} = (a_n)_{n=1}^\infty$ such that $\gcd(a_n, n) = 1$ for all n , and for any real numbers λ and X such that $3 \leq X \leq \lambda^{2/5-\varepsilon}$, the estimate

$$\mathcal{P}(\mathcal{A}; \lambda, X) = \frac{315 \zeta(3)}{2\pi^4} \frac{\lambda X}{\log(\lambda X^2)} + O\left(\frac{\lambda X (\log \log X)^2}{(\log X)^2}\right)$$

holds, where the implied constant depends only on ε .

Proof. Let Δ be an arbitrary real number such that $X^{-1} \leq \Delta \leq 1$, and let

$$J = \left\lfloor \frac{2 \log \log X}{\log(1 + \Delta)} \right\rfloor \ll \Delta^{-1} \log \log X.$$

Put

$$X_j = X(1 + \Delta)^{j-J} \quad (j = 0, 1, \dots, J).$$

Note that

$$\frac{X}{(\log X)^2} \leq X_0 \leq \frac{2X}{(\log X)^2},$$

and we have

$$X_j \leq X_{j+1} \leq 2X_j \quad \text{and} \quad \log X_j \gg \log X.$$

Using the trivial bound $\pi(\lambda n^2; n^2, a_n) \leq \lambda$ for all $n \leq X_0$, we derive that

$$\begin{aligned} \mathcal{P}(\mathcal{A}; \lambda, X) &= \sum_{X_0 < n \leq X} \pi(\lambda n^2; n^2, a_n) + O(\lambda X_0) \\ &= \sum_{j=0}^{J-1} S_j + O\left(\frac{\lambda X}{(\log X)^2}\right), \end{aligned} \tag{5}$$

where

$$S_j = \sum_{X_j < n \leq X_{j+1}} \pi(\lambda n^2; n^2, a_n) \quad (j = 0, 1, \dots, J).$$

Since $X_{j+1} - X_j = \Delta X_j$, for every integer $n \in [X_j, X_{j+1}]$ we have

$$n^2 = X_j^2 + O(\Delta X_j^2). \tag{6}$$

For any such n , the number of primes $p \in [\lambda X_j^2, \lambda n^2]$ with $p \equiv a_n \pmod{n^2}$ does not exceed

$$\frac{\lambda n^2 - \lambda X_j^2}{n^2} + 1 \ll \frac{\Delta \lambda X_j^2}{n^2} + 1 \leq \Delta \lambda + 1 \ll \Delta \lambda$$

(since $\Delta\lambda \geq \Delta X \geq 1$). Therefore,

$$S_j = \sum_{X_j < n \leq X_{j+1}} \pi(\lambda X_j^2; n^2, a_n) + O(\Delta^2 \lambda X_j) \quad (j = 0, 1, \dots, J). \quad (7)$$

Furthermore,

$$\begin{aligned} & \left| \sum_{X_j < n \leq X_{j+1}} \pi(\lambda X_j^2; n^2, a_n) - \pi(\lambda X_j^2) \sum_{X_j < n \leq X_{j+1}} \frac{1}{\varphi(n^2)} \right| \\ & \leq \sum_{X_j < n \leq X_{j+1}} \left| \pi(\lambda X_j^2; n^2, a_n) - \frac{\pi(\lambda X_j^2)}{\varphi(n^2)} \right| \\ & \leq \frac{1}{X_j} \sum_{X_j < n \leq X_{j+1}} n \left| \pi(\lambda X_j^2; n^2, a_n) - \frac{\pi(\lambda X_j^2)}{\varphi(n^2)} \right| \\ & \leq \frac{1}{X_j} \sum_{n \leq X_{j+1}} n \max_{\gcd(a, n)=1} \left| \pi(\lambda X_j^2; n^2, a) - \frac{\pi(\lambda X_j^2)}{\varphi(n^2)} \right|. \end{aligned}$$

In view of the hypothesis that $3 \leq X \leq \lambda^{2/5-\varepsilon}$ we can apply Corollary 8 with $C = 4$ to derive the bound

$$\sum_{X_j < n \leq X_{j+1}} \pi(\lambda X_j^2; n^2, a_n) - \pi(\lambda X_j^2) \sum_{X_j < n \leq X_{j+1}} \frac{1}{\varphi(n^2)} \ll \frac{\lambda X_j}{(\log X)^4}. \quad (8)$$

Using (6) again, we write

$$\pi(\lambda X_j^2) \sum_{X_j < n \leq X_{j+1}} \frac{1}{\varphi(n^2)} = \sum_{X_j < n \leq X_{j+1}} \frac{\pi(\lambda n^2) + O(\Delta \lambda X_j^2)}{\varphi(n^2)}.$$

Using the prime number theorem in its simplest form, namely

$$\pi(y) = \frac{y}{\log y} + O\left(\frac{y}{(\log y)^2}\right),$$

(see [20, Chapter II.4, Theorem 1] for a stronger statement) together with the lower bound

$$\varphi(n^2) = n\varphi(n) \gg \frac{n^2}{\log \log(n+2)} \quad (n \in \mathbb{N})$$

(see [20, Chapter I.5, Theorem 4]) and the trivial inequalities

$$\log(\lambda X^2) \geq \log(\lambda n^2) \geq \log(\lambda X_0^2) = \log(\lambda X^2) + O(\log \log X),$$

which hold for any integer $n \in [X_0, X]$, we derive that

$$\begin{aligned} & \pi(\lambda X_j^2) \sum_{X_j < n \leq X_{j+1}} \frac{1}{\varphi(n^2)} \\ &= \lambda \sum_{X_j < n \leq X_{j+1}} \frac{n^2}{\varphi(n^2) \log(\lambda n^2)} + O\left(\frac{\Delta \lambda X_j \log \log X}{(\log X)^2} + \Delta^2 \lambda X_j \log \log X\right) \\ &= \frac{\lambda}{\log(\lambda X^2)} \sum_{X_j < n \leq X_{j+1}} \frac{n}{\varphi(n)} + O\left(\frac{\Delta \lambda X_j (\log \log X)^2}{(\log X)^2} + \Delta^2 \lambda X_j \log \log X\right). \end{aligned}$$

Combining this result with (7) and (8) we see that

$$\begin{aligned} S_j - \frac{\lambda}{\log(\lambda X^2)} \sum_{X_j < n \leq X_{j+1}} \frac{n}{\varphi(n)} \\ \ll \frac{\lambda X_j}{(\log X)^4} + \frac{\Delta \lambda X_j (\log \log X)^2}{(\log X)^2} + \Delta^2 \lambda X_j \log \log X. \end{aligned}$$

We insert this estimate in (5) and deduce that

$$\begin{aligned} \mathcal{P}(\mathcal{A}; \lambda, X) - \frac{\lambda}{\log(\lambda X^2)} \sum_{X_0 < n \leq X} \frac{n}{\varphi(n)} \\ \ll \left(\frac{\lambda}{(\log X)^4} + \frac{\Delta \lambda (\log \log X)^2}{(\log X)^2} + \Delta^2 \lambda \log \log X \right) \sum_{j=0}^{J-1} X_j \\ \ll \left(\frac{\lambda}{(\log X)^4} + \frac{\Delta \lambda (\log \log X)^2}{(\log X)^2} + \Delta^2 \lambda \log \log X \right) \Delta^{-1} X \\ = \frac{\Delta^{-1} \lambda X}{(\log X)^4} + \frac{\lambda X (\log \log X)^2}{(\log X)^2} + \Delta \lambda X \log \log X. \end{aligned}$$

Taking $\Delta = (\log X)^{-2}$ (for which our hypothesis $X^{-1} \leq \Delta \leq 1$ holds for all $X > 1$) and taking into account that (3) implies the estimate

$$\begin{aligned} \sum_{X_0 < n \leq X} \frac{n}{\varphi(n)} &= \frac{315 \zeta(3)}{2\pi^4} (X - X_0) + O(\log X) \\ &= \frac{315 \zeta(3)}{2\pi^4} X + O\left(\frac{X}{(\log X)^2}\right), \end{aligned}$$

we conclude the proof. \square

We are certain that the error term of Lemma 9 can be improved easily, but we have not attempted to do so as we only require the asymptotic behavior of $\mathcal{P}(\mathcal{A}; \lambda, X)$ stated in the next corollary.

Corollary 10. *Fix $\varepsilon \in (0, 2/5)$. For any sequence of integers $\mathcal{A} = (a_n)_{n=1}^{\infty}$ such that $\gcd(a_n, n) = 1$ for all n , and for any real numbers λ and X such that $\lambda^\varepsilon \leq X \leq \lambda^{2/5-\varepsilon}$, the estimate*

$$\mathcal{P}(\mathcal{A}; \lambda, X) = \left(\frac{315 \zeta(3)}{2\pi^4} + o(1) \right) \frac{\lambda X}{\log(\lambda X^2)}$$

holds, where the function implied by $o(1)$ depends only on ε .

5 The sets $\mathcal{S}_\pi(N, K)$ and $\mathcal{S}_\Pi(N, K)$

We begin with the observation that

$$\#\mathcal{S}_\pi(N, K) \geq \sum_{n \leq N} \pi(Kn^2; n^2, 1). \quad (9)$$

Indeed, if $p = kn^2 + 1$ is a prime which does not exceed Kn^2 , then the pair $(n, (p-1)/n^2)$ lies in $\mathcal{S}_\pi(N, K)$. Clearly, Corollary 10 can be applied to the sum on the right hand side of (9) to derive the lower bound

$$\#\mathcal{S}_\pi(N, K) \geq \left(\frac{315 \zeta(3)}{2\pi^4} + o(1) \right) \frac{KN}{\log(KN^2)}$$

provided that $K^\varepsilon \leq N \leq K^{2/5-\varepsilon}$. Moreover, even without the condition $N \geq K^\varepsilon$ we are able to get a lower bound of the same strength.

Theorem 11. *Fix $\varepsilon \in (0, 2/5)$, and suppose that $N \leq K^{2/5-\varepsilon}$. Then, the following bound holds:*

$$\#\mathcal{S}_\pi(N, K) \gg \frac{KN}{\log K}.$$

Proof. Using (9) together with the elementary bound

$$\frac{\psi(x; m, a)}{\log x} \leq \Pi(x; m, a) = \pi(x; m, a) + O(x^{1/2} \log x),$$

we have

$$\begin{aligned}
\#\mathcal{S}_\pi(N, K) &\geq \sum_{N/2 \leq n \leq N} \pi(Kn^2; n^2, 1) \\
&\geq \sum_{N/2 \leq n \leq N} \left(\frac{\psi(Kn^2; n^2, 1)}{\log(Kn^2)} + O(K^{1/2}n \log(Kn^2)) \right) \\
&\gg \frac{1}{\log K} \sum_{N/2 \leq n \leq N} \psi(\tfrac{1}{4}KN^2; n^2, 1) + O(K^{1/2}N^2 \log K) \\
&= \frac{1}{\log K} \sum_{N/2 \leq n \leq N} \frac{KN^2}{4\varphi(n^2)} + E(N, K) + O(K^{1/2}N^2 \log K),
\end{aligned}$$

where

$$\begin{aligned}
|E(N, K)| &\leq \frac{1}{\log K} \sum_{N/2 \leq n \leq N} \left| \psi(\tfrac{1}{4}KN^2; n^2, 1) - \frac{KN^2}{4\varphi(n^2)} \right| \\
&\leq \frac{2}{N \log K} \sum_{N/2 \leq n \leq N} n \left| \psi(\tfrac{1}{4}KN^2; n^2, 1) - \frac{KN^2}{4\varphi(n^2)} \right|.
\end{aligned}$$

Applying Lemma 7 with $x = \frac{1}{4}KN^2$ and $C = 1$ (which is permissible since our assumption $N \leq K^{2/5-\varepsilon}$ implies that $N \leq (\frac{1}{4}KN^2)^{2/9-\delta}$ for a suitable $\delta > 0$ that depends only on ε) we see that

$$E(N, K) \ll \frac{KN}{(\log K)^2},$$

and therefore,

$$\#\mathcal{S}_\pi(N, K) \gg \frac{KN^2}{\log K} \sum_{N/2 \leq n \leq N} \frac{1}{\varphi(n^2)} + O\left(K^{1/2}N^2 \log K + \frac{KN}{(\log K)^2}\right).$$

Since

$$\sum_{N/2 \leq n \leq N} \frac{1}{\varphi(n^2)} \geq \sum_{N/2 \leq n \leq N} \frac{1}{n^2} \gg \frac{1}{N},$$

the result follows. □

Theorem 12. *For any fixed $K \in \mathbb{N}$ we have*

$$\#\mathcal{S}_\pi(N, K) \ll_K \frac{N}{\log N}.$$

Proof. The Selberg sieve provides the following upper bound on the number of primes represented by an irreducible polynomial $F(n) = an^2 + bn + 1$ with integer coefficients (see Halberstam and Richert [6, Theorem 5.3] for a more general statement):

$$\begin{aligned} \#\{n \leq x : F(n) \text{ is prime}\} &\leq 2 \prod_p \left(1 - \frac{\chi_p(b^2 - 4a)}{p-1}\right) \\ &\times \frac{x}{\log x} \left(1 + O_F\left(\frac{\log \log 3x}{\log x}\right)\right), \end{aligned} \quad (10)$$

where χ_p is the quadratic character modulo p , that is, the Dirichlet character afforded by the Legendre symbol. The constant implied by O_F depends on F , and this is the reason that K is fixed in the statement of the theorem.

Trivially, we have

$$\#\mathcal{S}_\pi(N, K) \leq \sum_{k \leq K} \sum_{|\ell| < 2\sqrt{k}} \#\{n \leq N : kn^2 + \ell n + 1 \text{ is prime}\}.$$

Applying (10) with $F(n) = kn^2 + \ell n + 1$, the result is immediate. \square

Corollary 13. *For any fixed $K \in \mathbb{N}$ we have*

$$\#\mathcal{S}_\Pi(N, K) \ll_K \frac{N}{\log N}.$$

Proof. We have

$$\#\mathcal{S}_\Pi(N, K) \leq \#\mathcal{S}_\pi(N, K) + \sum_{j=2}^{\infty} \#\mathcal{S}_\Pi^{(j)}(N, K), \quad (11)$$

where for each $j \geq 2$, we use $\mathcal{S}_\Pi^{(j)}(N, K)$ to denote the set of pairs (n, k) in $\mathcal{S}_\Pi(N, K)$ associated with prime powers of the form $q = p^j$ with p prime. It is easy to see that

$$\#\mathcal{S}_\Pi^{(j)}(N, K) \ll K^{3/2} \pi((KN^2 + 2K^{1/2}N + 1)^{1/j}) \ll \begin{cases} K^2 N / \log N & \text{if } j = 2, \\ K^{11/6} N^{2/3} & \text{if } j \geq 3. \end{cases}$$

Indeed, for fixed k and p there are only $O(K^{1/2})$ possibilities for ℓ . Thus, for fixed p there are $O(K^{3/2})$ possibilities for (n, k) , where the implied constant

is absolute. Furthermore, $\mathcal{S}_{\Pi}^{(j)}(N, K) = \emptyset$ for all but $O(\log(KN))$ choices of j . Thus, from (11) we deduce that

$$\#\mathcal{S}_{\Pi}(N, K) \leq \#\mathcal{S}_{\pi}(N, K) + O_K(N/\log N),$$

and the result follows from Theorem 12. \square

An immediate consequence of Corollary 13 is that there are infinitely many pairs (n, k) that do not lie in \mathcal{S}_{Π} . In fact, if $k \in \mathbb{N}$ is fixed, then we see that $(n, k) \notin \mathcal{S}_{\Pi}$ for almost all $n \in \mathbb{N}$.

The situation is very different when $n \in \mathbb{N}$ is fixed, for in this case we expect that the pair (n, k) lies in the smaller set \mathcal{S}_{π} for all but finitely many $k \in \mathbb{N}$. To prove this, one needs to show that

$$\pi((k^{1/2}n + 1)^2; n, 1) - \pi((k^{1/2}n - 1)^2; n, 1) > 0$$

for all sufficiently large k . Although this problem is intractable at present, the probabilistic model of Cramér (see, for example, [5, 19]) predicts that

$$\pi((k^{1/2}n + 1)^2; n, 1) - \pi((k^{1/2}n - 1)^2; n, 1) \gg_n k^{1/2}/\log k$$

for all large k . Unconditionally, it may be possible to answer the following questions:

- If $n \in \mathbb{N}$ is fixed, is it true that $(n, k) \in \mathcal{S}_{\Pi}$ for almost all $k \in \mathbb{N}$?
- Is it true that for almost all $n \in \mathbb{N}$, there are only finitely many pairs (n, k) that do not lie in \mathcal{S}_{Π} ?

We conclude this section with the following:

Theorem 14. *The set $\mathcal{S}_{\Pi} \setminus \mathcal{S}_{\pi}$ is infinite. In fact, we have*

$$\#\{n \in N : (n, 1) \in \mathcal{S}_{\Pi} \setminus \mathcal{S}_{\pi}\} \geq (2 + o(1)) \frac{N}{\log N} \quad (N \rightarrow \infty).$$

Proof. Using the prime number theorem for arithmetic progressions together with a standard upper bound from sieve theory such as [6, Theorem 5.3], one sees that there are $(2 + o(1))N/\log N$ natural numbers $n \leq N$ such either $n - 1$ or $n + 1$ is prime, but not both, and such that the integers $n^2 + 1$, $n^2 + n + 1$ and $n^2 - n + 1$ are all composite. For any such n , either $(n - 1)^2$ or $(n + 1)^2$ is a prime power, and we have $(n, 1) \in \mathcal{S}_{\Pi}$; however, $n^2 + \ell n + 1$ is clearly composite for $-2 \leq \ell \leq 2$, and thus $(n, 1) \notin \mathcal{S}_{\pi}$. \square

6 The double sum $\mathcal{N}_{\mathcal{P}}(N, K)$

Here, we study the double sum $\mathcal{N}_{\mathcal{P}}(N, K)$ using the formula of Lemma 6. Our main result is the following:

Theorem 15. *Fix $\varepsilon \in (0, 2/5)$, and suppose that $K^\varepsilon \leq N \leq K^{2/5-\varepsilon}$. Then, the estimate*

$$\mathcal{N}_{\mathcal{P}}(N, K) = \left(\frac{210 \zeta(3)}{\pi^4} + o(1) \right) \frac{K^{3/2} N}{\log(K N^2)}$$

holds, where the function implied by $o(1)$ depends only on ε .

Proof. Using the trivial estimate

$$\pi(x + y; k, a) = \pi(x; k, a) + O(y/k + 1),$$

we see from Lemma 6 that $\mathcal{N}_{\mathcal{P}}(N, K)$ is equal to

$$\begin{aligned} & \sum_{\substack{n \leq N \\ |\ell| \leq 2\sqrt{K}}} \left(\pi(Kn^2; n^2, \ell n + 1) - \pi(\tfrac{1}{4}\ell^2 n^2; n^2, \ell n + 1) + O(\ell/n + 1) \right) \\ &= \sum_{\substack{n \leq N \\ |\ell| \leq 2\sqrt{K}}} \left(\pi(Kn^2; n^2, \ell n + 1) - \pi(\tfrac{1}{4}\ell^2 n^2; n^2, \ell n + 1) \right) \\ & \quad + O(K \log N + K^{1/2} N) \\ &= \sum_{|\ell| \leq 2\sqrt{K}} \left(\mathcal{P}(\mathcal{A}_\ell; K, N) - \mathcal{P}(\mathcal{A}_\ell; \tfrac{1}{4}\ell^2, N) \right) + O(K \log N), \end{aligned}$$

where $\mathcal{A}_\ell = (n\ell + 1)_{n=1}^\infty$ for each ℓ , and the sum $\mathcal{P}(\mathcal{A}_\ell; \lambda, X)$ is defined by (4). Note that we have used the bound $K^{1/2} N \ll K \log N$, which follows from our hypothesis that $N \leq K^{2/5-\varepsilon}$.

We now put $L = 2\sqrt{K}/\log K$ and write

$$\mathcal{N}_{\mathcal{P}}(N, K) = S_1 + S_2 + O(K \log N), \tag{12}$$

where

$$\begin{aligned} S_1 &= \sum_{|\ell| \leq L} \left(\mathcal{P}(\mathcal{A}_\ell; K, N) - \mathcal{P}(\mathcal{A}_\ell; \tfrac{1}{4}\ell^2, N) \right), \\ S_2 &= \sum_{L < |\ell| \leq 2\sqrt{K}} \left(\mathcal{P}(\mathcal{A}_\ell; K, N) - \mathcal{P}(\mathcal{A}_\ell; \tfrac{1}{4}\ell^2, N) \right). \end{aligned}$$

For S_1 we use the trivial estimate

$$S_1 \leq \sum_{|\ell| \leq L} \mathcal{P}(\mathcal{A}_\ell; K, N)$$

together with Corollary 10 to derive the bound

$$S_1 \ll \frac{LKN}{\log K} \ll \frac{K^{3/2}N}{(\log K)^2}. \quad (13)$$

For S_2 we apply Corollary 10 to both terms in the summation. Writing $\Theta = 315\zeta(3)/(2\pi^4)$, and taking into account that

$$\log(\ell^2 N^2/4) = (1 + o(1)) \log(KN^2) \quad (L < |\ell| \leq 2\sqrt{K}),$$

we see that

$$\begin{aligned} S_2 &= \sum_{L < |\ell| \leq 2\sqrt{K}} \left((\Theta + o(1)) \frac{KN}{\log(KN^2)} - (\Theta + o(1)) \frac{\ell^2 N}{4 \log(\ell^2 N^2/4)} \right) \\ &= (\Theta + o(1)) \frac{N}{\log(KN^2)} \sum_{L < |\ell| \leq 2\sqrt{K}} (K - \ell^2/4) = \left(\frac{4}{3}\Theta + o(1)\right) \frac{K^{3/2}N}{\log(KN^2)}. \end{aligned}$$

Using this bound and (13) in (12), we finish the proof. \square

7 The sets $\mathcal{N}_{m,k}$ and $\tilde{\mathcal{N}}_{m,k}$

In this section, we study the sets $\mathcal{N}_{m,k}$ and $\tilde{\mathcal{N}}_{m,k}$ introduced in §1. We begin the following:

Lemma 16. *For all $m, k \in \mathbb{N}$ we have $\mathcal{N}_{m,k} \subseteq \tilde{\mathcal{N}}_{m,k}$.*

Proof. For every $n \in \mathcal{N}_{m,k}$ there is a prime p and an elliptic curve E defined over \mathbb{F}_{p^m} such that $E(\mathbb{F}_{p^m}) \cong \mathbb{Z}_n \times \mathbb{Z}_{kn}$. By Lemma 1, $p^m = kn^2 + \ell n + 1$ for some integer ℓ that satisfies $|\ell| \leq 2\sqrt{k}$, that is, $n \in \tilde{\mathcal{N}}_{m,k}$. \square

7.1 Results with fixed values of m

In the case that $m = 1$, the set inclusion of Lemma 16 is an equality.

Theorem 17. *For all $k \in \mathbb{N}$ we have $\mathcal{N}_{1,k} = \tilde{\mathcal{N}}_{1,k}$.*

Proof. In view of Lemma 16 it suffices to show that $\tilde{\mathcal{N}}_{1,k} \subseteq \mathcal{N}_{1,k}$. For every $n \in \tilde{\mathcal{N}}_{1,k}$ there is a prime p such that $p = kn^2 + \ell n + 1$. Put $a = n\ell + 2$, and note that $|a| \leq 2\sqrt{p}$ since

$$a^2 = n^2\ell^2 + 4n\ell + 4 \leq 4(n^2k + n\ell + 1) = 4p.$$

If $\gcd(a, p) = 1$, then by Lemma 2 (i) there is an elliptic curve E/\mathbb{F}_p such that $\#E(\mathbb{F}_p) = p + 1 - a = kn^2$. On the other hand, if $p \mid a$, then the inequality $|a| \leq 2\sqrt{p}$ implies that either $p \leq 3$ and $a = \pm p$, or $a = 0$. Applying Lemma 2 (iv) in the former case and Lemma 2 (iv) in the latter, we again conclude that there is an elliptic curve E/\mathbb{F}_p such that $\#E(\mathbb{F}_p) = kn^2$. In all cases, since $p \equiv 1 \pmod{n}$, Lemma 3 (ii) guarantees that there is an elliptic curve E defined over \mathbb{F}_p such that $E(\mathbb{F}_p) \cong \mathbb{Z}_n \times \mathbb{Z}_{kn}$. Therefore, $n \in \mathcal{N}_{1,k}$. \square

Lemma 18. *For natural numbers n, k the set*

$$\tilde{\mathcal{P}}(n, k) = \{\text{primes } p : p^2 = kn^2 + \ell n + 1 \text{ for some } \ell \in \mathbb{Z} \text{ with } |\ell| \leq 2\sqrt{k}\}$$

contains at most one prime except for the following cases:

$$(i) \quad \tilde{\mathcal{P}}(n, k) = \{2, 3\} \text{ if } n = 1 \text{ and } 4 \leq k \leq 9;$$

$$(ii) \quad \tilde{\mathcal{P}}(n, k) = \{hn \pm 1\} \text{ if } k = h^2 \text{ for some } h \in \mathbb{N}, \text{ and both } hn - 1 \text{ and } hn + 1 \text{ are primes.}$$

Proof. It is easy to see that

$$\tilde{\mathcal{P}}(n, k) = \{\text{primes } p \in [n\sqrt{k} - 1, n\sqrt{k} + 1] : p^2 \equiv 1 \pmod{n}\}. \quad (14)$$

Since the interval $[n\sqrt{k} - 1, n\sqrt{k} + 1]$ has length two, the result follows immediately. \square

When $m = 2$, the inclusion of Lemma 16 can be proper. Fortunately, we are able to classify those natural numbers k for which this happens.

Theorem 19. *For all $k \in \mathbb{N}$ we have $\mathcal{N}_{2,k} = \tilde{\mathcal{N}}_{2,k}$ except for the following disjoint cases:*

- (i) $k = p^2 + 1$ for some prime $p \equiv 1 \pmod{4}$;
- (ii) $k = p^2 \pm p + 1$ for some prime $p \equiv 1 \pmod{3}$;
- (iii) $k = h^2$ for some integer $h > 1$.

In cases (i) and (ii) we have $\tilde{\mathcal{N}}_{2,k} \setminus \mathcal{N}_{2,k} = \{1\}$, and in case (iii) we have

$$\tilde{\mathcal{N}}_{2,k} \setminus \mathcal{N}_{2,k} = \{n \in \mathbb{N} : hn - 1 \text{ or } hn + 1 \text{ is prime}\}. \quad (15)$$

Proof. Let k be fixed, and suppose that $n \in \tilde{\mathcal{N}}_{2,k}$. Let p and ℓ be such that $p^2 = kn^2 + \ell n + 1$, $|\ell| \leq 2\sqrt{k}$, and put $a = \ell n + 2$. Then $|a| \leq 2p$, and using Lemmas 2 and 3 it is easy to see that n lies in $\mathcal{N}_{2,k}$ except possibly in the following cases:

- (1) $a = 0$ and $p \equiv 1 \pmod{4}$;
- (2) $a = \pm p$ and $p \equiv 1 \pmod{3}$;
- (3) $a = \pm 2p$ and k is not of the form p^j for any $j \geq 0$.

In case (1) we have $\ell n = -2$, which implies either that $(n, \ell) = (2, -1)$ and $p^2 = 4k - 1$, which is impossible, or that $(n, \ell) = (1, -2)$ and $p^2 = k - 1$. This shows that $\tilde{\mathcal{N}}_{2,k} \setminus \mathcal{N}_{2,k} \subseteq \{1\}$ and that k satisfies the condition (i). Since $k \geq 26$ and $k \neq h^2$ for any $h > 1$, we have $\tilde{\mathcal{P}}(n, k) = \{p\}$ by Lemma 18. It remains to show that $1 \notin \mathcal{N}_{2,k}$ in this case. Suppose on the contrary that $1 \in \mathcal{N}_{2,k}$. Then there is a prime p_0 and an elliptic curve E defined over $\mathbb{F}_{p_0^2}$ such that $E(\mathbb{F}_{p_0^2}) \cong \mathbb{Z}_1 \times \mathbb{Z}_k$. By Lemma 1 we see that $p_0^2 = k + \ell + 1$ for some integer ℓ such that $|\ell| \leq 2\sqrt{k}$; that is, $p_0 \in \tilde{\mathcal{P}}(n, k)$. Therefore, $p_0 = p$, and $\#E(\mathbb{F}_{p^2}) = k$. But this is impossible by Lemma 2(v) since $p \equiv 1 \pmod{4}$.

In case (2) we have $p = \pm(\ell n + 2) \equiv \pm 2 \pmod{n}$, thus $p^2 \equiv 4 \pmod{n}$. Since $p^2 = kn^2 + \ell n + 1 \equiv 1 \pmod{n}$ as well, it follows that $n \mid 3$. We claim that $n \neq 3$. Indeed, if $n = 3$, then $p^2 = 9k + 3\ell + 1 = 9k \pm p - 1$, and therefore $p^2 \mp p + 1 \equiv 0 \pmod{9}$. But this is impossible as neither $X^2 + X + 1$ nor $X^2 - X + 1$ has a root in \mathbb{Z}_9 . If $n = 1$, then $p^2 = k + \ell + 1 = k \pm p - 1$. This shows that $\tilde{\mathcal{N}}_{2,k} \setminus \mathcal{N}_{2,k} \subseteq \{1\}$ and that k satisfies the condition (ii). The proof that $1 \notin \mathcal{N}_{2,k}$ is similar to that of the preceding case.

In case (3) we have $p^2 = kn^2 \pm 2p - 1$, or $kn^2 = (p \mp 1)^2$; it follows that $n \mid p \mp 1$, and $k = h^2$ with $h = (p \mp 1)/n$. Since $k \neq p^0$, we see that k satisfies the condition (iii). It remains to establish (15).

Fix $h > 1$, and suppose that $n \in \tilde{\mathcal{N}}_{2,h^2}$. Then $\tilde{\mathcal{P}}(n, h^2) \neq \emptyset$, where by (14) we have

$$\tilde{\mathcal{P}}(n, h^2) = \{\text{primes } p \in [hn - 1, hn + 1] : p^2 \equiv 1 \pmod{n}\}.$$

First, suppose $\tilde{\mathcal{P}}(n, h^2)$ contains a prime p in the open interval $(hn - 1, hn + 1)$. Then, using Lemma 18, we deduce that $\tilde{\mathcal{P}}(n, h^2) = \{p\}$, and thus case (3) does not occur for any prime in $\tilde{\mathcal{P}}(n, h^2)$. Also, the cases (1) and (2) cannot occur, for otherwise $k = h^2$ would satisfy (i) or (ii), respectively, rather than (iii). Consequently, $n \in \mathcal{N}_{2,h^2}$ in this case.

Next, suppose $\tilde{\mathcal{P}}(n, h^2)$ does not contain a prime p in the open interval $(hn - 1, hn + 1)$. If $p \in \tilde{\mathcal{P}}(n, h^2)$, then $p = hn \pm 1$ for some choice of the sign, and we have $p^2 + 1 - h^2n^2 = \pm 2hn + 2 = \pm 2p$. If there were an elliptic curve E defined over \mathbb{F}_{p^2} such that $E(\mathbb{F}_{p^2}) \cong \mathbb{Z}_n \times \mathbb{Z}_{h^2n}$, then by Lemma 2 (ii) and Lemma 3 (i) it would follow that $n = h^2n$, which is impossible since $h > 1$. This argument shows that $n \notin \mathcal{N}_{2,h^2}$ in this case. \square

Corollary 20. *Suppose that k is not a perfect square. Then,*

$$\#\mathcal{N}_{2,k}(T) \ll_k \log T.$$

Proof. In view of Lemma 16, it is enough to show that $\#\tilde{\mathcal{N}}_{2,k}(T) \ll_k \log T$.

Suppose that $n \in \tilde{\mathcal{N}}_{2,k}$ with $n \leq T$. Then there is a prime p and an integer ℓ such that $p^2 = kn^2 + \ell n + 1$, $|\ell| \leq 2\sqrt{k}$, and we have

$$\max\{2kn + \ell, 2p\} \ll_k T. \quad (16)$$

Since

$$(2kn + \ell)^2 - k(2p)^2 = \ell^2 - 4k,$$

the pair $(2kn + \ell, 2p)$ is a solution of the Pell equation

$$X^2 - kY^2 = \ell^2 - 4k. \quad (17)$$

Note that $\ell^2 - 4k \neq 0$ since k is not a perfect square. It is well known (and easy to verify) that every solution $(x, y) \in \mathbb{Z}^2$ to an equation such as (17) has the form

$$x + y\sqrt{k} = (x_0 + y_0\sqrt{k})\omega^t \quad (t \in \mathbb{Z}),$$

where (x_0, y_0) is an arbitrary fixed solution, and ω is a fixed unit in $\mathbb{Q}(\sqrt{k})$; therefore,

$$t \ll_k \log \max\{|x|, |y|\}.$$

In view of (16) we have $t \ll_k \log T$ for every solution $(x, y) = (2kn + \ell, 2p)$ to (17), and the result follows. \square

We remark that Theorem 19 implies

$$\#\mathcal{N}_{2,1}(T) = \pi(T-1) + \pi(T+1) - \#\{p \leq T-1 : p+2 \text{ is prime}\} \sim \frac{2T}{\log T}.$$

For $m \geq 3$, the situation is more complicated. For example, it is easy to see that $3 \in \tilde{\mathcal{N}}_{3,237} \setminus \mathcal{N}_{3,237}$. Indeed, since $13^3 = 3^2 \cdot 237 + 3 \cdot 21 + 1$, we have $3 \in \tilde{\mathcal{N}}_{3,237}$. On the other hand, direct computation shows that there is no elliptic curve over any finite field \mathbb{F}_{p^3} whose group of points $E(\mathbb{F}_{p^3})$ is isomorphic to $\mathbb{Z}_3 \times \mathbb{Z}_{3 \cdot 237}$. In fact, the equation $p^3 = 3^2 \cdot 237 + 3\ell + 1$ with $|\ell| < 2\sqrt{237} = 30.79 \dots$ admits only one solution $(p, \ell) = (13, 21)$, and $13^3 + 1 - 9 \cdot 237 = 5 \cdot 13$ is not a value for the parameter a that is permitted by Lemma 2.

7.2 Results with $k = 1$

Here, we focus on the problem of bounding $\#\mathcal{N}_{m,1}(T)$. We begin by quoting three results on Diophantine equations due to Lebesgue [8], to Nagell [10], and to Ljunggren [9], respectively.

Lemma 21. *For any $m \in \mathbb{N}$, the Diophantine equation $y^m = x^2 + 1$ has only the trivial solutions $(0, \pm 1)$.*

Lemma 22. *For any $m \in \mathbb{N}$ that is not a power of three, the Diophantine equations $y^m = x^2 + x + 1$ and $y^m = x^2 - x + 1$ have only trivial solutions from the set $\{(0, \pm 1), (\pm 1, \pm 1)\}$.*

Lemma 23. *The only solutions of the Diophantine equation $y^3 = x^2 + x + 1$ are the following: $\{(0, \pm 1), (-1, \pm 1), (18, 7), (-19, 7)\}$.*

The main result here is the following:

Theorem 24. *If m is even, then*

$$\#\mathcal{N}_{m,1}(T) = (m + o(1)) \frac{T^{2/m}}{\log T} \quad (T \rightarrow \infty).$$

If $m \geq 5$ and m is odd, then $\mathcal{N}_{m,1} = \emptyset$. Also, $\mathcal{N}_{3,1} = \{18, 19\}$, and

$$\#\mathcal{N}_{1,1}(T) \ll \frac{T}{\log T}.$$

Proof. First, suppose that $m = 2r \geq 2$ and $n \in \mathcal{N}_{m,1}$. Then there exists a prime p such that

$$p^{2r} = n^2 + \ell n + 1 \quad \text{for some } \ell \in \{0, \pm 1, \pm 2\}.$$

However, the cases $\ell \in \{0, \pm 1\}$ can be excluded in view of Lemmas 21 and 22. Since the numbers n for which this relation holds with $\ell \in \{\pm 2\}$ are those of the form $n = p^r \pm 1$, by the prime number theorem it follows that

$$\#\{n \leq T : n = p^r \pm 1\} = (2 + o(1)) \frac{T^{1/r}}{\log T^{1/r}} = (m + o(1)) \frac{T^{2/m}}{\log T},$$

and the proof is complete when m is even.

Next suppose that $m = 2r + 1 \geq 5$. Combining Lemmas 21, 22 and 23, one sees that there is no integer n for which any one of the numbers $n^2 + 1$, $n^2 + n + 1$, or $n^2 - n + 1$ is the m -th power of a prime. Since the relation $(n \pm 1)^2 = p^{2r+1}$ is also impossible, it follows $\mathcal{N}_{m,1} = \emptyset$ as stated.

When $m = 3$ we are lead to consider the three Diophantine equations

$$y^3 = x^2 + 1, \quad y^3 = x^2 + x + 1 \quad \text{and} \quad y^3 = x^2 - x + 1.$$

The first equation has no nontrivial solution by Lemma 21, the second only the nontrivial solution $(18, 7)$ by Lemma 23, and the third only the nontrivial solution $(19, 7)$ by Lemma 23. Since $\gcd(7, 20) = \gcd(7, -17) = 1$, using Lemmas 2 and 3 we conclude that $\mathcal{N}_{3,1} = \{18, 19\}$.

As an application of Theorem 17, we deduce that

$$\mathcal{N}_{1,1}(T) = \{n \leq T : n^2 + 1, n^2 + n + 1, \text{ or } n^2 - n + 1 \text{ is prime}\}.$$

Using Brun sieve (see [20, Chapter I.4, Theorem 3]) or the Selberg sieve (see (10) in §5) we see that $\#\mathcal{N}_{1,1}(T) \ll T/\log T$ as required. \square

Remark 1. Recalling the asymptotic version of Schinzel's *Hypothesis H* (see [14]) given by Bateman and Horn [1], it is reasonable to conjecture that

$$\#\mathcal{N}_{1,1}(T) = (C + o(1)) \frac{T}{\log T} \quad (T \rightarrow \infty),$$

where

$$C = \frac{1}{2} \prod_{p \geq 3} \left(1 - \frac{\left(\frac{-1}{p}\right)}{p-1}\right) + \prod_{p \geq 3} \left(1 - \frac{\left(\frac{-3}{p}\right)}{p-1}\right)$$

and $\left(\frac{\cdot}{p}\right)$ is the Legendre symbol modulo p . We note that two distinct polynomials are simultaneously prime for $O(T/(\log T)^2)$ arguments $n \leq T$, so we simply estimate the number of prime values for each of the above polynomials independently.

7.3 Finiteness of $\mathcal{N}_{m,k}$ when $m \geq 3$

In this section, we set

$$\mathcal{K}_k = \bigcup_{m \geq 3} \mathcal{N}_{m,k} \quad \text{and} \quad \mathcal{M}_m = \bigcup_{k \geq 1} \mathcal{N}_{m,k}.$$

We show that there are only finitely many prime powers p^m with $m \geq 3$ for which there is an elliptic curve E defined over \mathbb{F}_{p^m} with $E(\mathbb{F}_{p^m}) \cong \mathbb{Z}_n \times \mathbb{Z}_{kn}$ for some $n \in \mathbb{N}$. In other words, we have:

Theorem 25. *For every $k \geq 2$ the set \mathcal{K}_k is finite.*

Proof. We apply a result of Schinzel and Tijdeman [15] which asserts that if a polynomial f with rational coefficients has at least two distinct zeros, then the equation $y^m = f(x)$, where x and y are integers with $y \neq 0$, implies that $m \leq c(f)$, where $c(f)$ is a computable constant that depends only on f .

For any $n \in \mathcal{K}_k$, there exists a prime p and integers m, ℓ with $m \geq 3$ and $|\ell| \leq 2\sqrt{k}$ such that $p^m = kn^2 + \ell n + 1$.

For values of ℓ with $|\ell| < 2\sqrt{k}$, the polynomial $kX^2 + \ell X + 1$ has distinct roots. Thus we apply a result of Schinzel and Tijdeman [15] which asserts that if a polynomial f with rational coefficients has at least two distinct zeros, then the equation $y^m = f(x)$, where x and y are integers with $y \neq 0$, implies that $m \leq c(f)$, where $c(f)$ is a computable constant that depends only on f , see also [16, Theorem 10.2]. Hence, there are only finitely many possibilities

for the number m . For any fixed pair (m, ℓ) , using a classical result in the theory of Diophantine equations (see [16, Theorem 6.1]), we conclude that there are only finitely many possibilities for the pair (n, p) .

If $\ell = \pm 2\sqrt{k}$, then $k = h^2$ is a perfect square, and we have $p^m = (hn \pm 1)^2$. Thus, m is even, and $h^2 n^2 = p^m + 1 - a$, where $a = \pm 2p^{m/2}$. Applying Lemma 3 (i) it follows that $kn = h^2 n = n$; this contradicts our hypothesis that $k \geq 2$ and shows that the case $\ell = \pm 2\sqrt{k}$ does not occur. \square

Remark 2. All of the underlying ingredients in the proof of Theorem 25 are effective, so one can easily obtain explicit bounds on $\#\mathcal{K}_k$ and $\max\{n \in \mathcal{K}_k\}$. Using the explicit estimates of Bugeaud [3, Theorem 2], it can be shown that $\mathcal{N}_{m,k} = \emptyset$ for any $m > 2^{137} k^{3/2} (\log_2 4k)^6$. Further a result of Bugeaud [4, Theorem 2] on solutions of superelliptic equations imply the bound $\max\{n \in \mathcal{N}_{m,k}\} \leq \exp(c(m)k^{14m}(\log k)^{8m})$, where $c(m)$ is an effectively computable constant that depends only on m .

A computer search suggests that the following table lists completely the elements in \mathcal{K}_k for $2 \leq k \leq 5$:

k	\mathcal{K}_k	
2	$\{3, 11, 45, 119, 120\}$	$2^4 = 2 \cdot 3^2 - 3 + 1,$ $3^5 = 2 \cdot 11^2 + 1,$ $2^{12} = 2 \cdot 45^2 + 45 + 1,$ $13^4 = 2 \cdot 119^2 + 2 \cdot 119 + 1,$ $13^4 = 2 \cdot 120^2 - 2 \cdot 120 + 1.$
3	$\{5, 72, 555\}$	$3^4 = 3 \cdot 5^2 + 5 + 1,$ $5^6 = 3 \cdot 72^2 + 72 + 1,$ $31^4 = 3 \cdot 555^2 - 555 + 1.$
4	$\{1, 9, 23\}$	$2^3 = 4 \cdot 1^2 + 3 \cdot 1 + 1,$ $7^3 = 4 \cdot 9^2 + 2 \cdot 9 + 1,$ $2^{11} = 4 \cdot 23^2 - 3 \cdot 23 + 1.$
5	$\{1, 2, 4, 56, 126\}$	$2^3 = 5 \cdot 1^2 + 2 \cdot 1 + 1,$ $3^3 = 5 \cdot 2^2 + 3 \cdot 2 + 1,$ $3^4 = 5 \cdot 4^2 + 1,$ $5^6 = 5 \cdot 56^2 - 56 + 1,$ $43^3 = 5 \cdot 126^2 + 126 + 1.$

Theorem 26. *For every natural number m we have $\mathcal{M}_m = \mathbb{N}$. In other words, for any $n, m \in \mathbb{N}$ there is a prime p and an elliptic curve E defined over \mathbb{F}_{p^m} such that $E(\mathbb{F}_{p^m}) \cong \mathbb{Z}_n \times \mathbb{Z}_{kn}$ for some $k \in \mathbb{N}$.*

Proof. Let $m \in \mathbb{N}$ be fixed. If $m \geq 2$, then we have the identity

$$X^m = (X^{m-2} + 2X^{m-3} + \cdots + (m-2)X + m-1)(X-1)^2 + m(X-1) + 1.$$

For any $n \in \mathbb{N}$, let p be a prime in the arithmetic progression $1 \bmod n$ that does not divide m , and put $d = (p-1)/n$. Applying the above identity with $X = p$, we have $p^m = kn^2 + \ell n + 1$, where

$$k = (p^{m-2} + 2p^{m-3} + \cdots + (m-2)p + m-1)d^2 \quad \text{and} \quad \ell = md.$$

The condition $|\ell| \leq 2\sqrt{k}$ is easily verified since

$$4k \geq 2m(m-1)d^2 \geq m^2d^2 = \ell^2 \quad (m \geq 2).$$

Furthermore, $a = p^m + 1 - kn^2 = \ell n + 2 = m(p-1)$ is not divisible by p . Hence, Lemma 3 shows that $n \in \mathcal{M}_m$.

If $m = 1$, then for any $n \in \mathbb{N}$, let p be an odd prime in the arithmetic progression $1 \bmod n^2$. Then $p = dn^2 + 1$ for some natural number d , and since $a = p + 1 - dn^2 = 2$ is not divisible by p , Lemma 3 shows that $n \in \mathcal{M}_1$. \square

8 Missed group structures

We have already given in (2) several examples of pairs (n, k) for which the group $\mathbb{Z}_n \times \mathbb{Z}_{kn}$ cannot be realized as the group of points on an elliptic curve defined over a finite field.

Here we present more extensive numerical results.

In Figure 1 we plot the counting function

$$f(D) = D^2 - \#\mathcal{S}_\Pi(D, D)$$

of “missed” pairs (n, k) with $\max\{n, k\} \leq D$ for values of D up to 37550. We immediately derive from Corollary 13 that

$$\lim_{D \rightarrow \infty} f(D)/D = \infty,$$

but this statement seems weak in view of our computations.

In Figure 2 we plot the counting function

$$F(N, K) = NK - \#\mathcal{S}_\Pi(N, K)$$

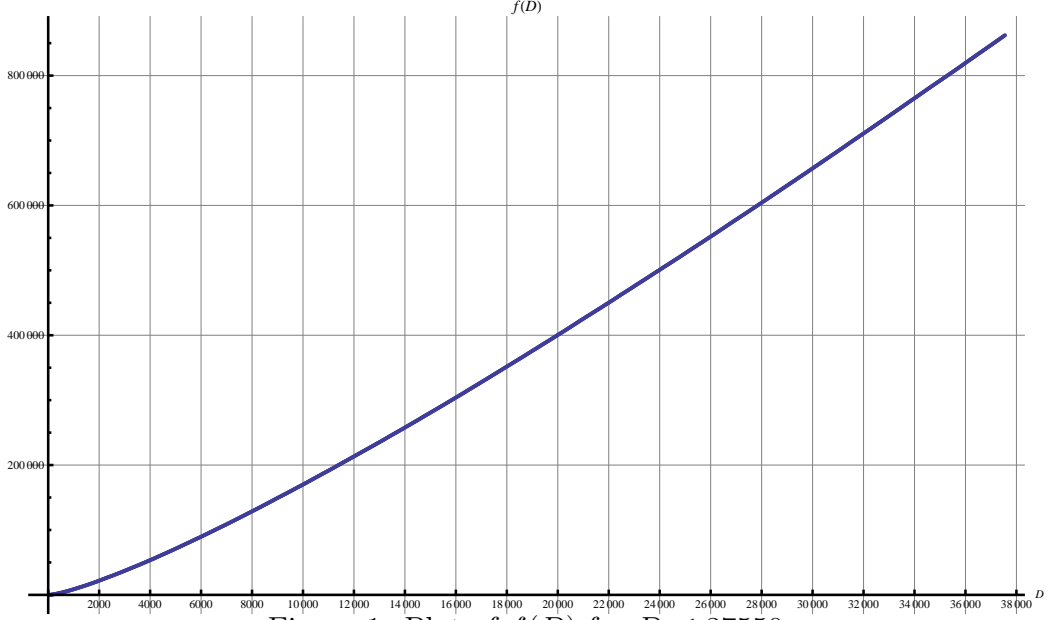


Figure 1: Plot of $f(D)$ for $D \leq 37550$

of “missed” pairs (n, k) with $n \leq N$ and $k \leq K$ for values of N and K up to 1000. For each fixed $N = N_0$ the function $G_{N_0}(K) = F(N_0, K)$ appears to be linear and increasing for modest values of K . Clearly, Corollary 13 implies that when $K = K_0$ is fixed then $H_{K_0}(N) = F(N, K_0) \sim K_0 N$ grows asymptotically linearly with the coefficient K_0 .

We now give some heuristic arguments to predict the behavior of $F(N, K)$. We note that a pair (n, k) contributes to $F(N, K)$ if $kn^2 + \ell n + 1$ is not a prime power for every ℓ such that $|\ell| \leq 2k^{1/2}$ (and in some other exceptional cases). Following the standard heuristic, $kn^2 + \ell n + 1$ is a prime power with “probability” about

$$\rho(n, k, \ell) = \begin{cases} \frac{n}{\varphi(n) \log(kn^2 + \ell n + 1)} & \text{if } kn^2 + \ell n + 1 > 1 \\ 0 & \text{otherwise.} \end{cases}$$

(where the ratio $n/\varphi(n)$ accounts for the fact that we seek prime powers in the arithmetic progression $1 \pmod n$). So $(n, k) \in [1, N] \times [1, K]$ contributes to $F(N, K)$ with “probability” about

$$\vartheta(n, k) = \prod_{|\ell| \leq 2k^{1/2}} (1 - \rho(n, k, \ell))$$

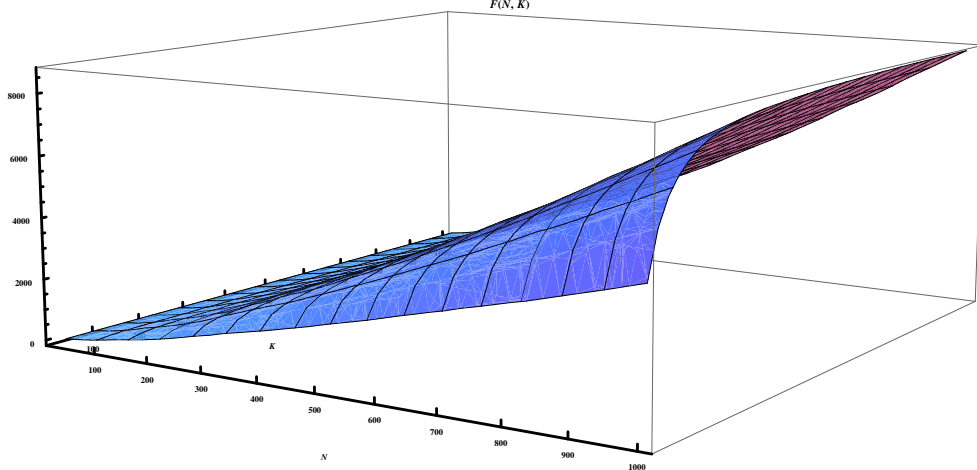


Figure 2: 3D plot of $F(N, K)$ for $N, K \leq 1000$

Thus, we expect that $F(N, K)$ is close to

$$B(N, K) = \sum_{n \leq N} \sum_{k \leq K} \vartheta(n, k).$$

We have not studied the function $B(N, K)$ analytically, but we note that for any fixed $\varepsilon > 0$ we have

$$\vartheta(n, k) \approx \begin{cases} 1 & \text{if } k \leq (\log n)^{2-\varepsilon}, \\ 0 & \text{if } k \geq (\log n)^{2+\varepsilon}. \end{cases}$$

Thus, it seems reasonable to expect that

$$F(N, K) \approx B(N, K) \approx \begin{cases} NK & \text{if } K \leq (\log N)^{2-\varepsilon}, \\ o(NK) & \text{if } K \geq (\log N)^{2+\varepsilon}. \end{cases}$$

One can see on Figure 3 that the ratio

$$\beta(N, K) = \frac{F(N, K)}{B(N, K)}$$

seems to stabilise when N and K are large enough.

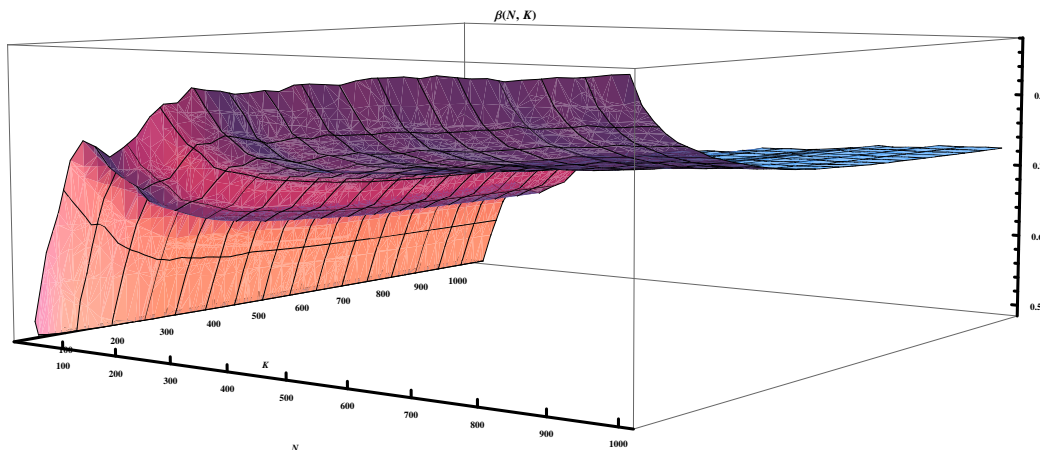


Figure 3: 3D plot of $\beta(N, K)$ for $N, K \leq 1000$

Acknowledgements

The authors are grateful to Karl Dilcher for pointing out the relevance of the result of Ljunggren [9] to this work, to Andrzej Schinzel for a discussion concerning Cramér's Conjecture for arithmetic progressions and to Corrado Falcolini for his help with Mathematica Plotting.

The second author was partially supported by GNSAGA from INDAM. The third author was supported in part by ARC Grant DP0881473, Australia and by NRF Grant CRP2-2007-03, Singapore.

References

- [1] P. T. BATEMAN AND R. A. HORN, *A heuristic asymptotic formula concerning the distribution of prime numbers*, Math. Comp. **16** (1962), 363–367
- [2] S. BAIER AND L. ZHAO, *Bombieri-Vinogradov type theorems for sparse sets of moduli*, Acta Arith. **125** (2006), 187–201.
- [3] Y. BUGEAUD, *Sur la distance entre deux puissances pures*, C. R. Acad. Sci. Paris Sér. I Math. **322** (1996), no. 12, 1119–1121.
- [4] Y. BUGEAUD, *Bounds for the solutions of superelliptic equations*, Compositio Math. **107** (1997), no. 2, 187–219.
- [5] A. GRANVILLE, *Harald Cramér and the distribution of prime numbers*, in *Harald Cramér Symposium (Stockholm, 1993)*, Scand. Actuar. J. (1995), no. 1, 12–28.
- [6] H. H. HALBERSTAM, H.-E. RICHERT, *Sieve methods*. Academic Press, London, 1974.
- [7] E. W. HOWE, *On the group orders of elliptic curves over finite fields*, Compositio Math. **85** (1993), 229–247.
- [8] V. A. LEBESGUE, *Sur l'impossibilité en nombres entiers de l'équation $x^m = y^2 + 1$* , Nouv. Ann. Math. **9** (1850), 178–181.
- [9] W. LJUNGGREN, *Einige Bemerkungen über die Darstellung ganzer Zahlen durch binäre kubische Formen mit positiver Diskriminante*, Acta Math. **75**, (1943), 1–21
- [10] T. NAGELL, *Des équations indéterminées $x^2 + x + 1 = y^m$ et $x^2 + x + 1 = 3y^m$* , Norsk Mat. Forenings Skr. Ser. I (1921), no. 2.
- [11] W. G. NOWAK, *On an error term involving the totient function*, Indian J. Pure Appl. Math. **20** (1989), 537–542.
- [12] R. REZAEIAN FARASHAHI AND I. E. SHPARLINSKI, *On group structures realized by elliptic curves over a fixed finite field*, Preprint, 2010.

- [13] H.-G. RÜCK, *A note on elliptic curves over finite fields*, Math. Comp. **49** (1987), 301–304.
- [14] A. SCHINZEL AND W. SIERPIŃSKI, *Sur certaines hypothèses concernant les nombres premiers*, Acta Arith. **4** (1958), 185–208; Erratum, **5** (1958), 259.
- [15] A. SCHINZEL AND R. TIJDEMAN, *On the equation $y^m = P(x)$* , Acta Arith. **31** (1976), no. 2, 199–204.
- [16] T. N. SHOREY AND R. TIJDEMAN, *Exponential Diophantine equations*, Cambridge University Press, Cambridge, 1986.
- [17] R. R. SITARAMACHANDRA, *On an error term of Landau*, Indian J. Pure Appl. Math. **13** (1982), 882–885.
- [18] R. R. SITARAMACHANDRA, *On an error term of Landau, II*, Number theory (Winnipeg, Man., 1983), Rocky Mountain J. Math. **15** (1985), no. 2, 579–588.
- [19] K. SOUNDARARAJAN, *The distribution of prime numbers*, in *Equidistribution in number theory, an introduction*, 59–83, NATO Sci. Ser. II Math. Phys. Chem. **237**, Springer, Dordrecht, 2007.
- [20] G. TENENBAUM, *Introduction to analytic and probabilistic number theory*, Cambridge University Press, 1995.
- [21] J. F. VOLOCH, *A note on elliptic curves over finite fields*, Bull. Soc. Math. Franc. **116** (1988), 455–458.
- [22] L. C. WASHINGTON, *Elliptic curves: Number theory and cryptography*, 2nd edition, Chapman & Hall/CRC Press, Boca Raton, FL, 2008.
- [23] W. C. WATERHOUSE, *Abelian varieties over finite fields*, Ann. Sci. Ecole Norm. Sup. **2** (1969), 521–560.