Experimental Mathematics

ISSN: 1058-6458 (Print) 1944-950X (Online) Journal homepage: http://www.tandfonline.com/loi/uexm20

On Group Structures Realized by Elliptic Curves over Arbitrary Finite Fields

William D. Banks , Francesco Pappalardi & Igor E. Shparlinski

To cite this article: William D. Banks , Francesco Pappalardi & Igor E. Shparlinski (2012) On Group Structures Realized by Elliptic Curves over Arbitrary Finite Fields, Experimental Mathematics, 21:1, 11-25, DOI: 10.1080/10586458.2011.606075

To link to this article: http://dx.doi.org/10.1080/10586458.2011.606075

Experimental Mathematics, 21(1):11–25, 2012

ISSN: 1058-6458 print / 1944-950X online
DOI: 10.1080/10586458.2011.606075

Taylor & Francis
Taylor & Francis Group

On Group Structures Realized by Elliptic Curves over Arbitrary Finite Fields

William D. Banks, Francesco Pappalardi, and Igor E. Shparlinski

CONTENTS

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 that 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,

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

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 [Howe 93, Rück 87, Voloch 88, Waterhouse 69]. 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. Besides being of intrinsic interest, various questions about the existence and frequency of various group structures of elliptic curves over finite fields are of primary importance for elliptic curve cryptography; see [Avanzi et al. 05] for a comprehensive treatise on such applications.

Our aim in the present paper is to introduce and investigate the set

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

Since we are interested in groups $\mathbb{Z}_n \times \mathbb{Z}_{kn}$ with a realization (1–1) in which $q = p$ is a prime number, we also

2000 AMS Subject Classification: Primary 11D45; Secondary 11P05, 14G05
Keywords: elliptic curve, finite field, group structure

study the subset $\mathcal{S}_\pi \subset \mathcal{S}_\Pi$ defined by

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

Although one can expect $\mathcal{S}_\pi$ and $\mathcal{S}_\Pi$ 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 $n,k \le 25$:

$$(11,1),\ (11,14),\ (13,6),\ (13,25),\ (15,4),\ (19,7),\ (19,10),\ (19,14),\ (19,15),\ (19,18),\ (21,18),\ (23,1),\ (23,5),\ (23,8),\ (23,19),\ (25,5),\ (25,14). \tag{1–2}$$

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

$$\mathcal{S}_\pi(N,K) = \big\{(n,k) \in \mathcal{S}_\pi : n \le N, k \le K\big\},$$
$$\mathcal{S}_\Pi(N,K) = \big\{(n,k) \in \mathcal{S}_\Pi : n \le N, k \le K\big\}.$$

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

For natural numbers n and k, we also put

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

The set $\mathcal{P}(n,k)$ parameterizes 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

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

for which we obtain an asymptotic formula in certain ranges.

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

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

We remark that the distribution of group structures generated by elliptic curves over a *fixed* finite field $\mathbb{F}_q$ has been studied in [Rezaeian Farashahi and Shparlinski 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 primes less than or equal to x. For coprime integers a and $m \ge 1$, we put

$$\pi(x;m,a) = \#\big\{p \le x : p \equiv a \pmod m\big\},$$
$$\Pi(x;m,a) = \#\big\{q \le x : q \equiv a \pmod m\big\}.$$

We also set

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

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

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

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 $|F| \le cG$ 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, $\ll_\rho$, and $\gg_\rho$ to indicate that the implied constant depends on a given parameter ρ.

3. PRELIMINARIES

Lemma 3.1. *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 $|\ell| \le 2\sqrt{k}$.*

Proof. By the Hasse bound, we can write $kn^2 = q + 1 - a$ for some integer a that satisfies the bound $a^2 \le 4q$. 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

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

it follows that $|\ell| \le 2\sqrt{k}$ as required. □

The following result from [Waterhouse 69] (see also [Washington 08, Theorem 4.3]) is a characterization of the natural numbers $\mathbb{N}$ 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 3.2. *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 $|a| \le 2\sqrt{q}$ 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 from [Rück 87] (see also [Washington 08, Theorem 4.4]) describes the group structures that are possible for $E(\mathbb{F}_q)$ given that $\#E(\mathbb{F}_q) = N$; see also [Howe 93, Voloch 88].

Lemma 3.3. *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 3.2;*

(ii) *$n_1 \mid q - 1$ in all other cases of Lemma 3.2.*

Combining Lemmas 3.2 and 3.3, we get the following corollary.

Corollary 3.4. *If p is prime and $N \in \mathbb{N}$ with $|p + 1 - N| \le 2\sqrt{p}$, 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 3.5. *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 $|\ell| \le 2\sqrt{k}$.*

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 3.1, $p = kn^2 + \ell n + 1$ with some integer ℓ such that $|\ell| \le 2\sqrt{k}$.

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

$$\begin{aligned} |p + 1 - N|^2 &= (\ell n + 2)^2 = \ell^2 n^2 + 4\ell n + 4 \\ &\le 4kn^2 + 4\ell n + 4 = 4p\,; \end{aligned}$$

hence $|p + 1 - N| \le 2\sqrt{p}$. Applying Corollary 3.4 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 $p \in \mathcal{P}(n,k)$. □

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

Lemma 3.6. *For all $N, K \in \mathbb{N}$, we have*

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

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

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

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

(iii) $\frac{1}{4}\ell^2 n^2 + \ell n + 1 \le p \le Kn^2 + \ell n + 1$.

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

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

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

(iv) $k \le K$;

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

By Lemma 3.5, condition (v) is equivalent to the assertion that $p \in \mathcal{P}(n,k)$, whence

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

Since

$$\begin{aligned} &\sum_{n \le N} \#\mathcal{T}_1(n) \\ &\quad = \sum_{\substack{n \le N \\ |\ell| \le 2\sqrt{K}}} \Big(\pi(Kn^2 + \ell n + 1; n^2, \ell n + 1) - \pi\left(\frac{1}{4}\ell^2 n^2 + \ell n; n^2, \ell n + 1\right)\Big) \end{aligned}$$

and

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

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

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

$$\frac{1}{4}\ell^2 n^2 + \ell n + 1 \le kn^2 + \ell n + 1 \le Kn^2 + \ell n + 1,$$

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

Next, suppose that $(k, p) \in \mathcal{T}_2(n)$, and let ℓ be as in (v). By (iv), we have $|\ell| \le 2\sqrt{k} \le 2\sqrt{K}$, and $p \equiv \ell n + 1 \pmod{n^2}$ by (v). Furthermore, since $\frac{1}{4}\ell^2 \le k \le K$, 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 $\mathcal{T}_2(n) \to \mathcal{T}_1(n)$ given by $(k, p) \mapsto (\ell, p)$ is injective, we have $\#\mathcal{T}_2(n) \le \#\mathcal{T}_1(n)$, and the proof is complete. □

4. PRIMES IN SPARSE PROGRESSIONS

Below, we use the following result of [Baker 10] (see also [Baier and Zhao 06, Baier and Zhao 08]), which is a variant of the Bombieri–Vinogradov theorem, which deals with primes in arithmetic progressions to square moduli. In fact, we present it in a slightly modified (but equivalent) form that suits better our purpose.

Lemma 4.1. *For fixed $\varepsilon > 0$ and $C > 0$, we have*

$$\sum_{m \le x^{43/180-\varepsilon}} m \max_{y \le x} \max_{\gcd(a,m)=1} \left| \psi(y; m^2, a) - \frac{y}{\varphi(m^2)} \right| \ll \frac{x}{(\log x)^C},$$

where the implied constant depends only on ε and C.

By partial summation one obtains the following corollary.

Corollary 4.2. *For fixed $\varepsilon > 0$ and $C > 0$, we have*

$$\sum_{m \le x^{43/180-\varepsilon}} m \max_{y \le x} \max_{\gcd(a,m)=1} \left| \pi(y; m^2, a) - \frac{\pi(y)}{\varphi(m^2)} \right| \ll \frac{x}{(\log x)^C},$$

where the implied constant depends only on ε and C.

Furthermore, it is shown in [Baker 10] that for progressions modulo squares of primes, the range can be extended to the same level as in the classical Bombieri–Vinogradov theorem.

Lemma 4.3. *For fixed $\varepsilon > 0$ and $C > 0$, we have*

$$\sum_{p \le x^{1/4-\varepsilon}} p \max_{y \le x} \max_{\gcd(a,p)=1} \left| \psi(y; p^2, a) - \frac{y}{p(p-1)} \right| \ll \frac{x}{(\log x)^C},$$

where the implied constant depends only on ε and C.

For our applications of Corollary 4.2, we also need the well-known asymptotic formula

$$\sum_{n \le X} \frac{n}{\varphi(n)} = \frac{315\,\zeta(3)}{2\pi^4} X + O(\log X); \tag{4–1}$$

for more precise results, we refer the reader to [Nowak 89, Sitaramachandra 82, Sitaramachandra 85].

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

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

Lemma 4.4. *Fix $\varepsilon \in (0, 43/94)$. 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 $3 \le X \le \lambda^{43/94-\varepsilon}$, 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

$$J = \left\lfloor \frac{2 \log\log X}{\log 2} \right\rfloor.$$

Put

$$X_j = X 2^{j-J} \quad (j = 0, 1, \dots, J).$$

Note that

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

and we have

$$\log X_j \gg \log X.$$

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

$$\mathcal{P}(\mathcal{A}; \lambda, X) = \sum_{j=0}^{J-1} S_j + O\left(\frac{\lambda X}{(\log X)^2} \right), \tag{4–3}$$

where

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

In view of the hypothesis $3 \le X \le \lambda^{43/94-\varepsilon}$, we can apply Corollary 4.2 with $C = 2$ to derive the bound

$$\left| S_j - \sum_{X_j < n \le X_{j+1}} \frac{\pi(\lambda n^2)}{\varphi(n^2)} \right| \ll \frac{\lambda X}{(\log X)^2}. \tag{4–4}$$

Using the prime number theorem (see [Tenenbaum 95, Chapter II.4, Theorem 1], the well-known lower bound on the Euler function (see [Tenenbaum 95, Chapter I.5, Theorem 4]), and the trivial inequalities

$$\begin{aligned}\log(\lambda X^2) &\ge \log(\lambda n^2) \ge \log(\lambda X_0^2)\\ &= \log(\lambda X^2) + O(\log\log X),\end{aligned}$$

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

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

Combining this result with (4–4), we see that

$$S_j - \frac{\lambda}{\log(\lambda X^2)} \sum_{X_j < n \le X_{j+1}} \frac{n}{\varphi(n)} \ll \frac{\lambda X_j (\log\log X)^2}{(\log X)^2}.$$

We insert this estimate in (4–3) and deduce that

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

Recalling (4–1), we conclude the proof. □

We are certain that the error term of Lemma 4.4 can be improved easily, but we have not attempted to do so, since we require only the asymptotic behavior of $\mathcal{P}(\mathcal{A}; \lambda, X)$ stated in the next corollary.

Corollary 4.5. *Fix $\varepsilon \in (0, 43/94)$. 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 $\lambda^{\varepsilon} \le X \le \lambda^{43/94-\varepsilon}$, 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

$$\#\mathcal{S}_\pi(N, K) \ge \sum_{n \le N} \pi(Kn^2; n^2, 1). \tag{5–1}$$

Indeed, if $p = kn^2 + 1$ is a prime that does not exceed Kn^2, then the pair $(n, (p-1)/n^2)$ lies in $\mathcal{S}_\pi(N, K)$. Clearly, Corollary 4.5 can be applied to the sum on the right-hand side of (5–1) to derive the lower bound

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

provided that $K^{\varepsilon} \le N \le K^{43/94-\varepsilon}$. Moreover, even without the condition $N \ge K^{\varepsilon}$, we are able to get a lower bound of the same strength.

Theorem 5.1. *Fix $\varepsilon \in (0, 43/94)$, and suppose that $N \le K^{43/94-\varepsilon}$. Then the following bound holds:*

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

Proof. Using (5–1) together with the elementary bound

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

we have

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

where

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

Applying Lemma 4.1 with $x = \frac{1}{4}KN^2$ and $C = 1$ (which is permissible, since our assumption $N \le K^{43/94-\varepsilon}$

implies that $N \le (\frac{1}{4}KN^2)^{43/180-\delta}$ for a suitable $\delta > 0$ that depends only on ε), we see that

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

and therefore,

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

Since

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

the result follows. □

It is also clear that considering only prime values of n and using Lemma 4.1 instead of Lemma 4.3, we obtain a slightly weaker estimate but in a wider range.

Theorem 5.2. *Fix $\varepsilon \in (0, 1/2)$, and suppose that $N \le K^{1/2-\varepsilon}$. Then the following bound holds:*

$$\#\mathcal{S}_\pi(N,K) \gg \frac{KN}{(\log K)^2}.$$

We now turn to upper bounds on $\#\mathcal{S}_\pi(N,K)$.

Theorem 5.3. *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 74, Theorem 5.3] for a more general statement):

$$\#\{n \le x : F(n) \text{ is prime}\} \le 2\prod_p \left(1 - \frac{\chi_p(b^2-4a)}{p-1}\right) \times \frac{x}{\log x}\left(1 + O_F\left(\frac{\log\log 3x}{\log x}\right)\right), \tag{5–2}$$

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

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

Applying (5–2) with $F(n) = kn^2 + \ell n + 1$, the result is immediate. □

Corollary 5.4. *For any fixed $K \in \mathbb{N}$ we have*

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

Proof. We have

$$\#\mathcal{S}_\Pi(N,K) \le \#\mathcal{S}_\pi(N,K) + \sum_{j=2}^{\infty} \#\mathcal{S}_\Pi^{(j)}(N,K), \tag{5–3}$$

where for each $j \ge 2$, 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

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

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, $\mathcal{S}_\Pi^{(j)}(N,K) = \varnothing$ for all but $O(\log(KN))$ choices of j. Thus from (5–3), we deduce that

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

and the result follows from Theorem 5.3. □

An immediate consequence of Corollary 5.4 is that there are infinitely many pairs (n,k) that do not lie in $\mathcal{S}_\Pi$. 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 $\mathcal{S}_\pi$ for all but finitely many $k \in \mathbb{N}$. To prove this, one needs to show that

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

for all sufficiently large k. Although this problem is intractable at present, the probabilistic model of Cramér (see, for example, [Granville 95, Soundararajan 07])

predicts that

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

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 $\mathcal{S}_\Pi$?

We conclude this section with the following theorem.

Theorem 5.5. *The set $\mathcal{S}_\Pi\setminus\mathcal{S}_\pi$ is infinite. In fact, we have*

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

Proof. For fixed $j\in\{\pm1\}$, let $\widetilde{\mathcal{S}}_j(N)$ be the set of natural numbers $n\le N$ such that $n+j$ is prime, but n^2+1, n^2+n+1, and n^2-n+1 are all composite. Using the prime number theorem together with a standard upper bound from sieve theory such as [Halberstam and Richert 74, Theorem 5.3], one has

$$\#\widetilde{\mathcal{S}}_j(N)\ge(1+o(1))N/\log N\quad(j\in\{\pm1\},N\to\infty),$$

whereas

$$\#\big(\widetilde{\mathcal{S}}_{+1}(N)\cap\widetilde{\mathcal{S}}_{-1}(N)\big)\ll N/(\log N)^2.$$

For each $n\in\widetilde{\mathcal{S}}_j(N)$, $(n+j)^2$ is a prime power, and thus $(n,1)\in\mathcal{S}_\Pi$; on the other hand, $n^2+\ell n+1$ is clearly composite for $-2\le\ell\le2$, and thus $(n,1)\notin\mathcal{S}_\pi$. Therefore,

$$\#\{n\le N:(n,1)\in\mathcal{S}_\Pi\setminus\mathcal{S}_\pi\}\ge\#\widetilde{\mathcal{S}}_{+1}(N)+\#\widetilde{\mathcal{S}}_{-1}(N)-\#\big(\widetilde{\mathcal{S}}_{+1}(N)\cap\widetilde{\mathcal{S}}_{-1}(N)\big),$$

and the result follows from the bounds above. □

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 3.6. Our main result is the following.

Theorem 6.1. *Fix $\varepsilon\in(0,43/94)$, and suppose that $K^\varepsilon\le N\le K^{43/94-\varepsilon}$. Then the estimate*

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

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 3.6 that $\mathcal{N}_\mathcal{P}(N,K)$ is equal to

$$\begin{aligned}&\sum_{\substack{n\le N\\|\ell|\le2\sqrt{K}}}\left(\pi\left(Kn^2;n^2,\ell n+1\right)-\pi\left(\frac14\ell^2n^2;n^2,\ell n+1\right)+O\left(\frac{\ell}{n}+1\right)\right)\\&=\sum_{\substack{n\le N\\|\ell|\le2\sqrt{K}}}\left(\pi\left(Kn^2;n^2,\ell n+1\right)-\pi\left(\frac14\ell^2n^2;n^2,\ell n+1\right)\right)+O(K\log N+K^{1/2}N)\\&=\sum_{|\ell|\le2\sqrt{K}}\left(\mathcal{P}(\mathcal{A}_\ell;K,N)-\mathcal{P}\left(\mathcal{A}_\ell;\frac14\ell^2,N\right)\right)+O(K\log N),\end{aligned}$$

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 (4–2). Note that we have used the bound $K^{1/2}N\ll K\log N$, which follows from our hypothesis that $N\le K^{43/94-\varepsilon}$.

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

$$\mathcal{N}_\mathcal{P}(N,K)=S_1+S_2+O(K\log N),\tag{6–1}$$

where

$$S_1=\sum_{|\ell|\le L}\left(\mathcal{P}(\mathcal{A}_\ell;K,N)-\mathcal{P}\left(\mathcal{A}_\ell;\frac14\ell^2,N\right)\right),$$

$$S_2=\sum_{L<|\ell|\le2\sqrt{K}}\left(\mathcal{P}(\mathcal{A}_\ell;K,N)-\mathcal{P}\left(\mathcal{A}_\ell;\frac14\ell^2,N\right)\right).$$

For S_1 we use the trivial estimate

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

together with Corollary 4.5 to derive the bound

$$S_1\ll\frac{LKN}{\log K}\ll\frac{K^{3/2}N}{(\log K)^2}.\tag{6–2}$$

For S_2 we apply Corollary 4.5 to both terms in the summation. Writing $\Theta=315\,\zeta(3)/(2\pi^4)$, and taking into account that

$$\log(\ell^2N^2/4)=(1+o(1))\log(KN^2)\quad(L<|\ell|\le2\sqrt{K}),$$

we see that

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

Using this bound and (6–2) in (6–1), we finish the proof. □

7. THE SETS $\mathcal{N}_{m,k}$ AND $\widetilde{\mathcal{N}}_{m,k}$

7.1. A General Observation

In this section, we study the sets $\mathcal{N}_{m,k}$ and $\widetilde{\mathcal{N}}_{m,k}$ introduced in Section 1. We begin with a lemma.

Lemma 7.1. *For all $m,k\in\mathbb{N}$ we have $\mathcal{N}_{m,k}\subseteq\widetilde{\mathcal{N}}_{m,k}$.*

Proof. For every $n\in\mathcal{N}_{m,k}$, there exist 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 3.1, $p^m=kn^2+\ell n+1$ for some integer ℓ that satisfies $|\ell|\le 2\sqrt{k}$, that is, $n\in\widetilde{\mathcal{N}}_{m,k}$. □

7.2. Results with Fixed Values of m

In the case that $m=1$, the set inclusion of Lemma 7.1 is an equality.

Theorem 7.2. *For all $k\in\mathbb{N}$ we have $\mathcal{N}_{1,k}=\widetilde{\mathcal{N}}_{1,k}$.*

Proof. In view of Lemma 7.1, it suffices to show that $\widetilde{\mathcal{N}}_{1,k}\subseteq\mathcal{N}_{1,k}$. For every $n\in\widetilde{\mathcal{N}}_{1,k}$, there is a prime p such that $p=kn^2+\ell n+1$. Put $a=n\ell+2$, and note that $|a|\le 2\sqrt{p}$, since

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

If $\gcd(a,p)=1$, then by Lemma 3.2 (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 $|a|\le 2\sqrt{p}$ implies that either $p\le 3$ and $a=\pm p$, or $a=0$. Applying Lemma 3.2 (vi) in the former case and Lemma 3.2 (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 3.3 (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, $n\in\mathcal{N}_{1,k}$. □

Lemma 7.3. *For natural numbers n,k the set*

$$\begin{aligned}\widetilde{\mathcal{P}}(n,k)=\{&\text{primes } p: p^2=kn^2+\ell n+1 \text{ for some } \ell\in\mathbb{Z}\\ &\text{with } |\ell|\le 2\sqrt{k}\,\}\end{aligned}$$

contains at most one prime except for the following cases:

(i) $\widetilde{\mathcal{P}}(n,k)=\{2,3\}$ *if* $n=1$ *and* $4\le k\le 9$;

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

Proof. It is easy to see that

$$\begin{aligned}\widetilde{\mathcal{P}}(n,k)=\{&\text{primes } p\in\left[n\sqrt{k}-1,n\sqrt{k}+1\right]: \\ &p^2\equiv 1 \pmod n\}.\end{aligned} \tag{7–1}$$

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

When $m=2$, the inclusion of Lemma 7.1 can be proper. Fortunately, we are able to classify those natural numbers k for which this happens.

Theorem 7.4. *For all $k\in\mathbb{N}$ we have $\mathcal{N}_{2,k}=\widetilde{\mathcal{N}}_{2,k}$ 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* $\widetilde{\mathcal{N}}_{2,k}\setminus\mathcal{N}_{2,k}=\{1\}$, *and in case* (iii), *we have*

$$\widetilde{\mathcal{N}}_{2,k}\setminus\mathcal{N}_{2,k}=\{n\in\mathbb{N}: hn-1 \text{ or } hn+1 \text{ is prime}\}. \tag{7–2}$$

Proof. Let k be fixed, and suppose that $n\in\widetilde{\mathcal{N}}_{2,k}$. Let p and ℓ be such that $p^2=kn^2+\ell n+1$, $|\ell|\le 2\sqrt{k}$, and put $a=\ell n+2$. Then $|a|\le 2p$, and using Lemmas 3.2 and 3.3, 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 $j\ge 0$.

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 $\widetilde{\mathcal{N}}_{2,k}\setminus\mathcal{N}_{2,k}\subseteq\{1\}$ and that k satisfies condition (i). Since $k\ge 26$ and $k\ne h^2$ for any $h>1$, we have $\widetilde{\mathcal{P}}(n,k)=\{p\}$ by Lemma 7.3. It remains to show that $1\notin\mathcal{N}_{2,k}$ in this case. Suppose to the contrary that $1\in\mathcal{N}_{2,k}$. Then

there exist 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 3.1, we see that $p_0^2 = k + \ell + 1$ for some integer ℓ such that $|\ell| \le 2\sqrt{k}$; that is, $p_0 \in \widetilde{\mathcal{P}}(n,k)$. Therefore, $p_0 = p$ and $\#E(\mathbb{F}_{p^2}) = k$. But this is impossible by Lemma 3.2 (v), since $p \equiv 1 \pmod 4$.

In case (2), we have $p = \pm(\ell n + 2) \equiv \pm 2 \pmod n$, whence $p^2 \equiv 4 \pmod n$. 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, because 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 $\widetilde{\mathcal{N}}_{2,k} \setminus \mathcal{N}_{2,k} \subseteq \{1\}$ and that k satisfies 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 condition (iii). It remains to establish (7–2).

Fix $h > 1$, and suppose that $n \in \widetilde{\mathcal{N}}_{2,h^2}$. Then $\widetilde{\mathcal{P}}(n,h^2) \neq \varnothing$, where by (7–1) we have

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

First, suppose $\widetilde{\mathcal{P}}(n,h^2)$ contains a prime p in the open interval $(hn-1, hn+1)$. Then, using Lemma 7.3, we deduce that $\widetilde{\mathcal{P}}(n,h^2) = \{p\}$, and thus case (3) does not occur for any prime in $\widetilde{\mathcal{P}}(n,h^2)$. Also, 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 $\widetilde{\mathcal{P}}(n,h^2)$ does not contain a prime p in the open interval $(hn-1, hn+1)$. If $p \in \widetilde{\mathcal{P}}(n,h^2)$, then $p = hn \pm 1$ for some choice of 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 $E(\mathbb{F}_{p^2}) \cong \mathbb{Z}_n \times \mathbb{Z}_{h^2 n}$, then by Lemma 3.2 (ii) and Lemma 3.3 (i), it would follow that $n = h^2 n$, which is impossible since $h > 1$. This argument shows that $n \notin \mathcal{N}_{2,h^2}$ in this case. □

Corollary 7.5. *Suppose that k is not a perfect square. Then*

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

Proof. In view of Lemma 7.1, it is enough to show that $\#\widetilde{\mathcal{N}}_{2,k}(T) \ll_k \log T$.

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

$$\max\{2kn + \ell, 2p\} \ll_k T. \tag{7–3}$$

Since

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

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

$$X^2 - kY^2 = \ell^2 - 4k. \tag{7–4}$$

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 (7–4) has the form

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

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

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

In view of (7–3) we have $t \ll_k \log T$ for every solution $(x,y) = (2kn + \ell, 2p)$ to (7–4), and the result follows. □

We remark that Theorem 7.4 implies

$$\begin{aligned}\#\mathcal{N}_{2,1}(T) &= \pi(T-1) + \pi(T+1) \\ &\quad - \#\big\{p \le T - 1 : p + 2 \text{ is prime}\big\} \\ &\sim \frac{2T}{\log T}.\end{aligned}$$

For $m \ge 3$, the situation is more complicated. For example, it is easy to see that $3 \in \widetilde{\mathcal{N}}_{3,237} \setminus \mathcal{N}_{3,237}$. Indeed, since $13^3 = 3^2 \cdot 237 + 3 \cdot 21 + 1$, we have $3 \in \widetilde{\mathcal{N}}_{3,237}$. 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 $|\ell| < 2\sqrt{237} = 30.79\ldots$ 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 3.2.

7.3. 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 respectively to Lebesgue, Nagell, and Ljunggren.

Lemma 7.6. [Lebesgue 50] *For any $m \in \mathbb{N}$, the Diophantine equation $y^m = x^2 + 1$ has only the trivial solutions $(0, \pm 1)$.*

Lemma 7.7. [Nagell 21] *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 7.8. [Ljunggren 43] *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 7.9. *If m is even, then*

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

If $m \geq 5$ and m is odd, then $\mathcal{N}_{m,1} = \varnothing$. 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 7.6 and 7.7. 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

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

and the proof is complete when m is even.

Next suppose that $m = 2r + 1 \geq 5$. Combining Lemmas 7.6, 7.7, and 7.8, one sees that there is no integer n for which any one of the numbers $n^2 + 1$, $n^2 + n + 1$, $n^2 - n + 1$ is the mth power of a prime. Since the relation $(n \pm 1)^2 = p^{2r+1}$ is also impossible, it follows that $\mathcal{N}_{m,1} = \varnothing$, as stated.

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

$$y^3 = x^2 + 1, \quad y^3 = x^2 + x + 1, \quad y^3 = x^2 - x + 1.$$

The first equation has no nontrivial solution by Lemma 7.6, the second only the nontrivial solution $(18, 7)$ by Lemma 7.8, and the third only the nontrivial solution $(19, 7)$ by Lemma 7.8. Since $\gcd(7, 20) = \gcd(7, -17) = 1$, we conclude using Lemmas 3.2 and 3.3 that $\mathcal{N}_{3,1} = \{18, 19\}$.

As an application of Theorem 7.2, we deduce that

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

Using the Brun sieve (see [Tenenbaum 95, Chapter I.4, Theorem 3]) or the Selberg sieve (see (5–2)), we see that $\#\mathcal{N}_{1,1}(T) \ll T/\log T$ as required. □

Remark 7.10. The asymptotic version of Schinzel's Hypothesis H (see [Schinzel and Sierpiński 58]) given in [Bateman and Horn 62], leads us to conjecture that

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

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.

7.4. 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 the following result.

Theorem 7.11. *For every $k \geq 2$, the set $\mathcal{K}_k$ is finite.*

Proof. For any $n \in \mathcal{K}_k$, there exist 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 from [Schinzel and Tijdeman 76] that 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 [Shorey and Tijdeman 86, 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 [Shorey and Tijdeman 86, Theorem 6.1]), we conclude that there are only finitely many possibilities for the pair (n, p).

k	$\mathcal{K}_k$	Elements in $\mathcal{K}_k$
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	$\{1, 2, 4, 56, 126\}$	$2^3 = 5\cdot 1^2 + 2\cdot 1 + 1,\ 3^3 = 5\cdot 2^2 + 3\cdot 2 + 1,\ 3^4 = 5\cdot 4^2 + 1,\ 5^6 = 5\cdot 56^2 - 56 + 1, 43^3 =$ $5\cdot 126^2 + 126 + 1$

TABLE 1. The elements in $\mathcal{K}_k$ for $2 \le k \le 5$ found by computer search.

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 3.3 (i), it follows that $kn = h^2 n = n$; this contradicts our hypothesis that $k \ge 2$ and shows that the case $\ell = \pm 2\sqrt{k}$ does not occur. □

Remark 7.12. All of the underlying ingredients in the proof of Theorem 7.11 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 96, Theorem 2], it can be shown that $\mathcal{N}_{m,k} = \varnothing$ for any $m > 2^{137} k^{3/2} (\log_2 4k)^6$. Furthermore, the result [Bugeaud 97, Theorem 2] on solutions of superelliptic equations implies the bound $\max\{n \in \mathcal{N}_{m,k}\} \le \exp\big(c(m) k^{14m} (\log k)^{8m}\big)$, where $c(m)$ is an effectively computable constant that depends only on m.

A computer search suggests that Table 1 lists completely the elements in $\mathcal{K}_k$ for $2 \le k \le 5$.

Theorem 7.13. *For every natural number m we have $\mathcal{M}_m = \mathbb{N}$. In other words, for any $n, m \in \mathbb{N}$ there exist 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 $m \ge 2$, then we have the identity

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

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

$$k = \big(p^{m-2} + 2p^{m-3} + \cdots + (m-2)p + m - 1\big)\, d^2$$

and $\ell = md$. The condition $|\ell| \le 2\sqrt{k}$ is easily verified, since

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

Furthermore, $a = p^m + 1 - kn^2 = \ell n + 2 = m(p-1)$ is not divisible by p. Hence, Lemma 3.3 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 3.3 shows that $n \in \mathcal{M}_1$. □

8. MISSED GROUP STRUCTURES

We have already given in (1–2) 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 $\max\{n, k\} \le D$ for values of D up to 37,550. We immediately derive from Corollary 5.4 that

$$\lim_{D\to\infty} f(D)/D = \infty.$$

But this statement seems weak in view of our computations.

In Figure 2 we plot the counting function

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

of "missed" pairs (n, k) with $n \le N$ and $k \le K$ 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 5.4 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.

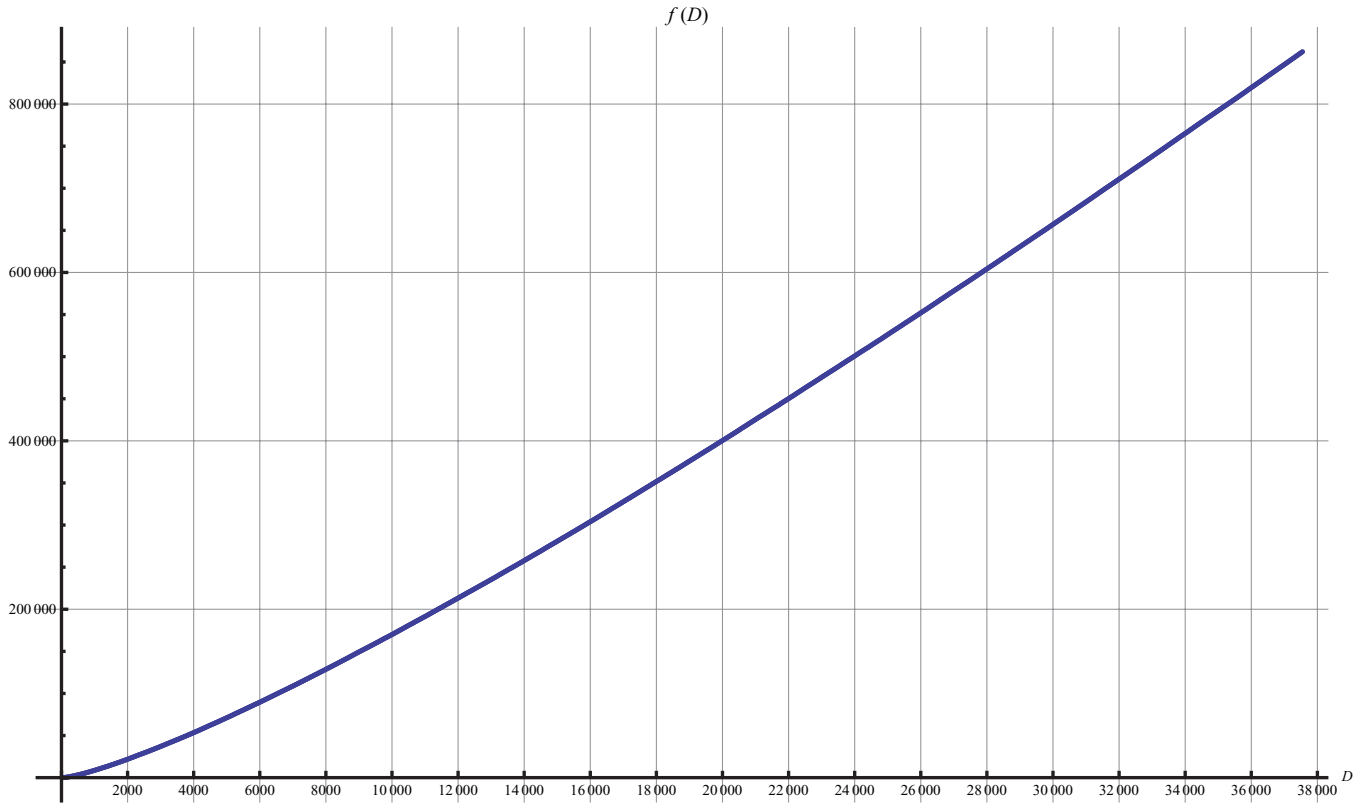

FIGURE 1. Plot of $f(D)$ for $D \leq 37550$ (color figure available online).

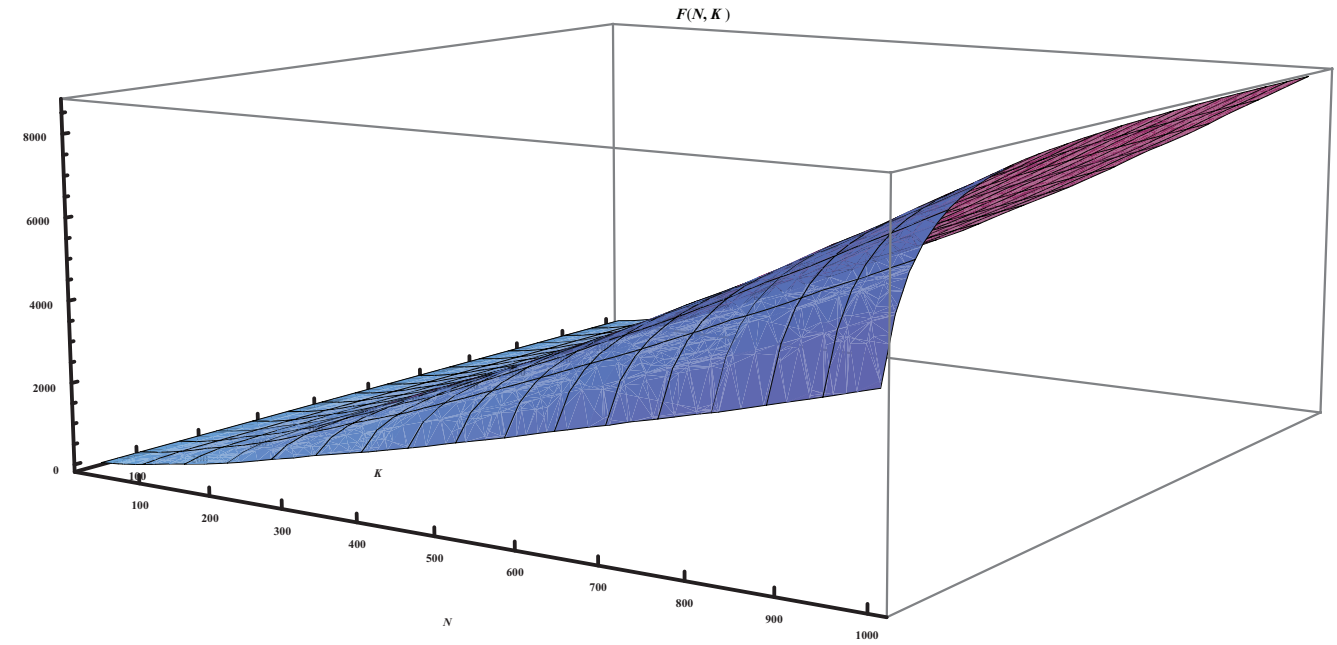

FIGURE 2. 3D plot of $F(N, K)$ for $N, K \leq 1000$ (color figure available online).

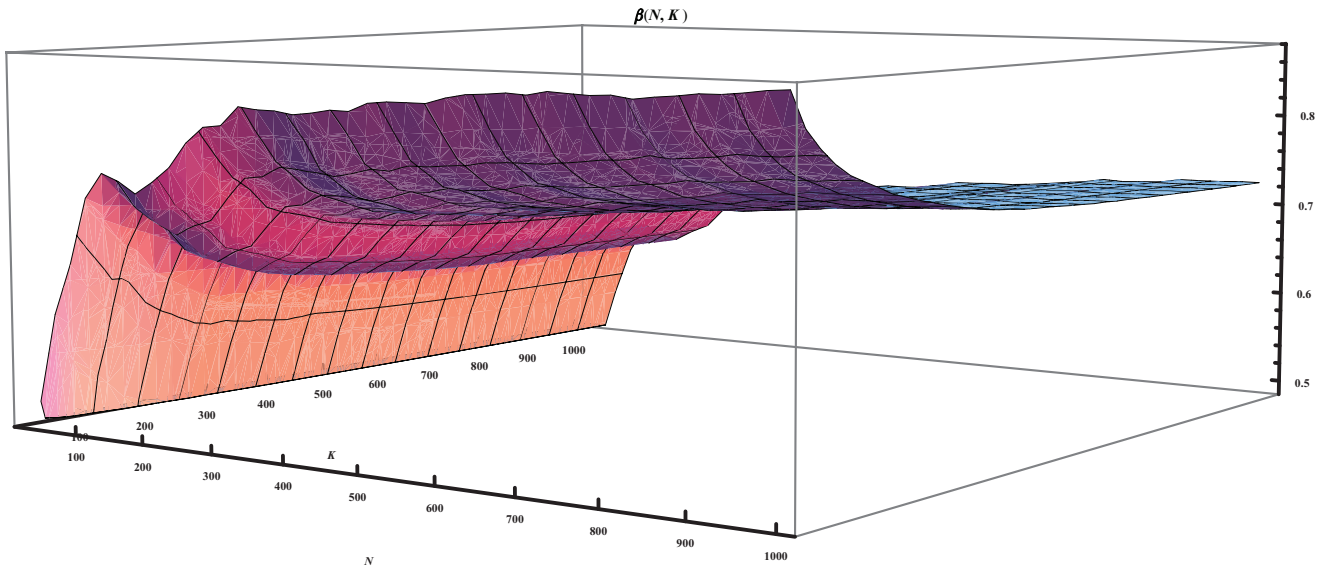

FIGURE 3. 3D plot of $\beta(N,K)$ for $N,K \le 1000$ (color figure available online).

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 $|\ell| \le 2k^{1/2}$ (and in some other exceptional cases). Following the standard heuristic, $kn^2+\ell n+1$ is a prime power with "probability" about

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

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

$$\vartheta(n,k) = \prod_{|\ell|\le 2k^{1/2}} (1-\rho(n,k,\ell)). \qquad (8\text{–}1)$$

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

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

Above, we have considered the primality events in the sequence $kn^2+\ell n+1$, $|\ell| \le 2k^{1/2}$, to be independent. This is not quite correct, however, so both the formula (8–1) and the expression for $B(N,K)$ should contain a correction factor to reflect such local dependencies. Nevertheless, we do believe that $B(N,K)$ is of the same order of magnitude as $F(N,K)$.

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

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

Thus, it seems reasonable to expect that

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

One can see in Figure 3 that the ratio

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

seems to stabilize at around $0.71\ldots$ when N and K are large enough.

We also leave as an open problem the task of deriving an analytic expression for the adjustment factor in (8–1) that may explain the limiting behavior of $\beta(N,K)$.

ACKNOWLEDGMENTS

The authors are grateful to Roger Baker for sending us a preliminary version of his work [Baker 10], to Karl Dilcher for pointing out the relevance of the results of [Ljunggren 43] 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

[Avanzi et al. 05] R. Avanzi, H. Cohen, C. Doche, G. Frey, T. Lange, K. Nguyen, and F. Vercauteren. *Handbook of Elliptic and Hyperelliptic Curve Cryptography*. CRC Press, 2005.

[Baier and Zhao 06] S. Baier and L. Zhao. "Bombieri–Vinogradov Type Theorems for Sparse Sets of Moduli." *Acta Arith.* 125 (2006), 187–201.

[Baier and Zhao 08] S. Baier and L. Zhao. "An Improvement for the Large Sieve for Square Moduli." *J. Number Theory* 128 (2008), 154–174.

[Baker 10] R. Baker. "Primes in Arithmetic Progressions to Spaced Moduli." Preprint, 2010.

[Bateman and Horn 62] P. T. Bateman and R. A. Horn. "A Heuristic Asymptotic Formula Concerning the Distribution of Prime Numbers." *Math. Comp.* 16 (1962), 363–367.

[Bugeaud 96] Y. Bugeaud. "Sur la distance entre deux puissances pures." *C. R. Acad. Sci. Paris Sér. I Math.* 322 (1996), 1119–1121.

[Bugeaud 97] Y. Bugeaud. "Bounds for the Solutions of Superelliptic Equations." *Compositio Math.* 107 (1997), 187–219.

[Granville 95] A. Granville. "Harald Cramér and the Distribution of Prime Numbers," Harald Cramér Symposium (Stockholm, 1993). *Scand. Actuar. J.* (1995), 12–28.

[Halberstam and Richert 74] H. H. Halberstam and H.-E. Richert. *Sieve Methods*. Academic Press, 1974.

[Howe 93] E. W. Howe. "On the Group Orders of Elliptic Curves over Finite Fields." *Compositio Math.* 85 (1993), 229–247.

[Lebesgue 50] V. A. Lebesgue. "Sur l'impossibilité en nombres entiers de l'équation $x^m = y^2 + 1$." *Nouv. Ann. Math.* 9 (1850), 178–181.

[Ljunggren 43] W. Ljunggren. "Einige Bemerkungen über die Darstellung ganzer Zahlen durch binäre kubische Formen mit positiver Diskriminante." *Acta Math.* 75, (1943), 1–21.

[Nagell 21] 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), 1–14.

[Nowak 89] W. G. Nowak. "On an Error Term Involving the Totient Function." *Indian J. Pure Appl. Math.* 20 (1989), 537–542.

[Rezaeian Farashahi and Shparlinski 12] R. Rezaeian Farashahi and I. E. Shparlinski. "On Group Structures Realized by Elliptic Curves over a Fixed Finite Field." Exp. Math. 21 (2012), 1–10.

[Rück 87] H.-G. Rück. "A Note on Elliptic Curves over Finite Fields." *Math. Comp.* 49 (1987), 301–304.

[Schinzel and Sierpiński 58] A. Schinzel and W. Sierpiński. "Sur certaines hypothèses concernant les nombres premiers." *Acta Arith.* 4 (1958), 185–208; Erratum *Acta Arith.* 5 (1958), 259.

[Schinzel and Tijdeman 76] A. Schinzel and R. Tijdeman. "On the Equation $y^m = P(x)$." *Acta Arith.* 31 (1976), 199–204.

[Shorey and Tijdeman 86] T. N. Shorey and R. Tijdeman. "Exponential Diophantine Equations." Cambridge University Press, 1986.

[Sitaramachandra 82] R. R. Sitaramachandra. "On an Error Term of Landau." *Indian J. Pure Appl. Math.* 13 (1982), 882–885.

[Sitaramachandra 85] R. R. Sitaramachandra. "On an Error Term of Landau, II," Number Theory (Winnipeg, Man., 1983). *Rocky Mountain J. Math.* 15 (1985), 579–588.

[Soundararajan 07] K. Soundararajan. "The Distribution of Prime Numbers." In *Equidistribution in Number Theory, an Introduction*, NATO Sci. Ser. II Math. Phys. Chem. 237, pp. 59–83. Springer, 2007.

[Tenenbaum 95] G. Tenenbaum. *Introduction to Analytic and Probabilistic Number Theory*. Cambridge University Press, 1995.

[Voloch 88] J. F. Voloch. "A Note on Elliptic Curves over Finite Fields." *Bull. Soc. Math. Franc.* 116 (1988), 455–458.

[Washington 08] L. C. Washington. *Elliptic Curves: Number Theory and Cryptography*, 2nd edition. Chapman & Hall/CRC Press, 2008.

[Waterhouse 69] W. C. Waterhous. "Abelian Varieties over Finite Fields." *Ann. Sci. Ecole Norm. Sup.* 2 (1969), 521–560.

William D. Banks, Department of Mathematics, University of Missouri, Columbia, MO 65211, USA (bankswd@missouri.edu)

Francesco Pappalardi, Dipartimento di Matematica, Universitá Roma Tre, Roma, I–00146, Italy (pappa@mat.uniroma3.it)

Igor E. Shparlinski, Department of Computing, Macquarie University, Sydney, NSW 2109, Australia (igor.shparlinski@mq.edu.au)

Received March 14, 2010; accepted October 19, 2010.