Finite Fields and Their Applications 91 (2023) 102264

Contents lists available at ScienceDirect

Finite Fields and Their Applications

journal homepage: www.elsevier.com/locate/ffa

# Arithmetic progressions in certain subsets of finite fields

Sadık Eyidoğan [a,*], Haydar Göral [b], Mustafa Kutay Kutlu [b]

[a] Department of Mathematics, Faculty of Science, Çukurova University, 01330 Adana, Turkey
[b] Department of Mathematics, Izmir Institute of Technology, 35430 Urla, Izmir, Turkey

ARTICLE INFO

*Article history:*
Received 23 February 2023
Received in revised form 30 April 2023
Accepted 30 June 2023
Available online 20 July 2023
Communicated by Gary L. Mullen

*MSC:*
11B25
11T24

*Keywords:*
Arithmetic progressions
Szemerédi's theorem
Arithmetic geometry
Weil estimates
Sato-Tate conjecture

ABSTRACT

In this note, we focus on how many arithmetic progressions we have in certain subsets of finite fields. For this purpose, we consider the sets $S_p = \{t^2 : t \in \mathbb{F}_p\}$ and $C_p = \{t^3 : t \in \mathbb{F}_p\}$, and we use the results on Gauss and Kummer sums. We prove that for any integer $k \geq 3$ and for an odd prime number $p$, the number of $k$-term arithmetic progressions in $S_p$ is given by

$$\frac{p^2}{2^k} + R,$$

where

$$|R| \leq \left(\frac{k-2}{4} - \frac{k-2}{2^{k-1}}\right) \cdot p^{\frac{3}{2}} + c_k \cdot p$$

and $c_k$ is a computable constant depending only on $k$. The proof also uses finite Fourier analysis and certain types of Weil estimates. Also, we obtain some formulas that give the exact number of arithmetic progressions of length $\ell$ in the set $S_p$ when $\ell \in \{3, 4, 5\}$ and $p$ is an odd prime number. For $\ell = 4, 5$, our formulas are based on the number of points on

[*] Corresponding author.
*E-mail addresses:* seyidogan@cu.edu.tr (S. Eyidoğan), haydargoral@iyte.edu.tr (H. Göral), mustafakutlu@iyte.edu.tr (M.K. Kutlu).

https://doi.org/10.1016/j.ffa.2023.102264

certain elliptic curves, and the error term is best possible due to the Sato-Tate conjecture.

## 1. Introduction

In 1927, van der Waerden [46] proved a celebrated theorem regarding the existence of arithmetic progressions in any partition of the positive integers with finitely many classes. This is one of the fundamental results of Ramsey theory, and this theorem has been strengthened in many different directions. In 1936, a strengthening of van der Waerden's theorem was conjectured by Erdős and Turán [19], which states that any subset of positive integers with a positive upper density contains arbitrarily long arithmetic progressions. For a subset $A$ of positive integers, its upper density is defined as

$$\bar{d}(A) = \limsup_{N\to\infty} \frac{|A \cap \{1, \ldots, N\}|}{N}.$$

In 1953, this conjecture was confirmed by Roth [35] for arithmetic progressions of length three. Actually, his proof shows not only the conjecture is true for arithmetic progressions of length three, but it also provides an explicit upper bound for the largest size of a subset of $\{1, \ldots, N\}$ with no non-trivial arithmetic progressions of length three (which is denoted by $r_3(N)$). In 1969, Szemerédi [42] extended the aforementioned result to arithmetic progressions of length four, and then in 1975 he developed his combinatorial method to resolve the conjecture for arbitrarily long arithmetic progressions, see [43]. The affirmative answer to Erdős and Turán's conjecture is now known as Szemerédi's theorem, which is one of the cornerstones of additive combinatorics. There is also a finitary version of Szemerédi's theorem which is equivalent to Szemerédi's theorem itself. Let $\varepsilon > 0$, and let $k$ be a positive integer. Then, there is some $N(\varepsilon, k)$ such that if $n \geq N(\varepsilon, k)$, then any subset of $\{1, 2, \ldots, n\}$ with at least $\varepsilon n$ elements contains a $k$-term arithmetic progression. The smallest such $N(\varepsilon, k)$ is called the Szemerédi number denoted by $S(\varepsilon, k)$.

A second proof of Szemerédi's theorem was given by Furstenberg [20] using ergodic theory in 1977. Furstenberg's proof was a major breakthrough in terms of both his techniques, which gave rise to many natural generalizations of the theorem, for example the density version of the Hales-Jewett theorem [21] and the polynomial Szemerédi theorem [4]. Despite their depths and impacts, the proofs of Szemerédi and Furstenberg fail to give upper bounds for $r_k(N)$ (which is the largest size of a subset of $\{1, \ldots, N\}$ with no non-trivial $k$-term arithmetic progressions), since Szemerédi's proof applies van der Waerden's theorem and Furstenberg's proof uses the axiom of choice.

Gowers developed new Fourier analytic methods to reprove Szemerédi's theorem for arithmetic progressions of length four [22] in 1998, and arbitrarily long arithmetic progressions [23] in 2001. In fact, he gave not only a proof of the full Szemerédi's theorem but also a quantitative bound for $r_k(N)$.

As well as in the integers, Szemerédi-type problems have been extensively studied in subsets of finite fields. While much work has been done on the problem of whether subsets of finite fields contain arithmetic progressions, in this study we concentrate on how many arithmetic progressions we have in certain subsets of finite fields. Here, we consider the set $S_p = \{t^2 : t \in \mathbb{F}_p\}$ and we obtain the exact asymptotic for the number of $k$-term arithmetic progressions in this set. Our approach relies on the estimation of character sums, which has been a recurrent topic in number theory. A typical exponential and character sum is of the form

$$T_1 = \sum_{(x_1,...,x_n)\in\mathbb{F}_p^n} \psi(q(x_1,...,x_n))$$

and

$$T_2 = \sum_{(x_1,...,x_n)\in\mathbb{F}_p^n} \chi(q(x_1,...,x_n)),$$

where $q(x_1,...,x_n) \in \mathbb{F}_p[x_1,...,x_n]$ of degree $d$, $\psi(x)$ is a non-trivial additive character and $\chi(x)$ is a non-trivial multiplicative character on the finite field $\mathbb{F}_p$. The expectation is the estimate

$$|T_i| \leq cp^{n/2}, \tag{1}$$

where $c$ is a constant depending on $n$ and the degree $d$ of the polynomial $q(x_1,...,x_n)$, and this is sort of a randomness. The above estimation corresponds to the Riemann hypothesis in finite fields. The estimation (1) was first achieved by Hasse [25] for single-variable smooth cubics and then generalized by Weil [48]. For each odd prime number $p$ and for each non-linear polynomial $f \in \mathbb{Z}[X]$, we denote the Weil sum by

$$s(f,p) = \sum_{x\in\mathbb{F}_p} e_p(f(x)),$$

where $e_p(x) = e^{2\pi ix/p}$. In 1948, Weil proved as a consequence of his work [48] in algebraic geometry that if $p$ is an odd prime number and $f \in \mathbb{Z}[X]$ is a non-linear polynomial with $f \notin p\mathbb{Z}[X]$, then we have

$$|s(f,p)| \leq (\deg f - 1) \cdot \sqrt{p}.$$

The higher dimensional version for the estimation of the exponential sum $T_1$ was obtained in the seminal works of Deligne [16,17] where he proved the Riemann hypothesis for finite fields that was also conjectured by Weil. More precisely, Deligne [16] proved that if $p$ does not divide $d$ and if the homogeneous part $q_d$ with degree $d$ of $q$ defines a smooth hypersurface in $\mathbb{P}^{n-1}$, then the expected estimation for $T_1$ holds with $c = (d-1)^n$. Later on, Katz [28] proved the multiplicative version of Deligne's result and obtained an estimation for the sum $T_2$. In this article, our algebraic sets that we encounter are highly singular and this is why we need singular character sum estimations. An estimation of this type was proved by Rojas-León [33], extending the work of Katz. In a very recent work, Rojas-León [34] deduced an estimation for multi-variable multiplicative character sums, which extends the well-known estimates for both classical Jacobi sums and one-variable polynomial multiplicative character sums. The result of Rojas-León [34] will be crucial to prove our first theorem of this paper, and we obtain an asymptotic for the number of $k$-term arithmetic progressions ($k$-APs) in $S_p$ with a better error term. Moreover, our error term is sharp and best possible when $k \in \{4, 5\}$, owing to the celebrated Sato-Tate conjecture (a theorem now), see [3,8,24,45]. Observe that our estimate in the next theorem is reminiscent of the Riemann hypothesis in the sense of finite fields.

**Theorem 1.1.** *Let $k \geq 4$ be a positive integer and $p > 3$ be a prime number. The number of $k$-APs in $S_p$ is given by*

$$\frac{p^2}{2^k} + R,$$

*where*

$$|R| \leq \left(\frac{k-2}{4} - \frac{k-2}{2^{k-1}}\right) \cdot p^{\frac{3}{2}} + c_k \cdot p$$

*and $c_k$ is an explicit computable constant depending only on $k$.*

When we look at the historical process of the distribution of quadratic residues or counting quadratic residue patterns in finite fields, it is seen that it has been widely handled by different mathematicians. Over the past 100 years, for $k \geq 1$ and an odd prime $p > k$, it has been desirable to count how many $k$-tuples of consecutive numbers $a, a+1, \ldots, a+k-1$ in $\mathbb{F}_p^\times$ have predetermined quadratic residue or nonresidue behavior. For a choice of $k$ signs $\varepsilon_1, \ldots, \varepsilon_k \in \{\pm 1\}$, set

$$N_p(\varepsilon_1, \ldots, \varepsilon_k) = \left| \left\{ a \in \mathbb{F}_p^\times : \left(\frac{a+i-1}{p}\right) = \varepsilon_i \text{ for } i = 1, \ldots, k \right\} \right|,$$

where $\left(\frac{\cdot}{p}\right)$ is the Legendre symbol. In 1896, Aladov [1] counted each quadratic residue patterns of length 2, and some quadratic residue patterns of length 3 in $\mathbb{F}_p^\times$. In the 1930s, Davenport [13,14] considered this counting problem for $k \geq 4$. It was shown [27, Chapter 9] that for $k$ signs $\varepsilon_1, \ldots, \varepsilon_k \in \{\pm 1\}$, and an odd prime $p > k$,

$$\left| N_p(\varepsilon_1, \dots, \varepsilon_k) - \frac{p}{2^k} \right| < (k-1)\sqrt{p} + \frac{k}{2}.$$

Moreover, quadratic residue patterns with gaps that are not necessarily consecutive was also counted: if $p > k$ and $c_1, \dots, c_k$ are distinct in $\mathbb{F}_p$, the set

$$\left\{ a \in \mathbb{F}_p^\times : \left( \frac{a + c_i}{p} \right) = \varepsilon_i \text{ for } i = 1, \dots, k \right\}$$

has a size $N_p$, and it satisfies

$$\left| N_p - \frac{p}{2^k} \right| < (k-1)\sqrt{p} + \frac{k}{2}.$$

See also [9]. When we fix $\varepsilon_i = 1$ for each $i$, this yields that the number of $k$-APs in $S_p$ is given by

$$\frac{p^2}{2^k} + H, \tag{2}$$

where

$$|H| \leq (k-1) \cdot p^{\frac{3}{2}} + O_k(p).$$

In this case, it is seen that the estimate in Theorem 1.1 for the number of $k$-APs in $S_p$ has a better error term than that of (2).

In our following result, we obtain the formulas which give the exact number of non-trivial arithmetic progressions of length 3 in the set $S_p$. Note that "non-trivial" means that the common difference of the arithmetic progression is not zero.

**Proposition 1.2.** *Let $p$ be an odd prime number. The number of non-trivial 3-APs in $S_p$ is given by the following table:*

| *The formula* | *The prime number p* |
|---|---|
| $\frac{1}{8}(p+3)(p-1)$ | $p \equiv 1 \pmod 8$ |
| $\frac{1}{8}(p-3)(p-1)$ | $p \equiv 3 \pmod 8$ |
| $\frac{1}{8}(p-1)(p-1)$ | $p \equiv 5 \pmod 8$ |
| $\frac{1}{8}(p+1)(p-1)$ | $p \equiv 7 \pmod 8$ |

A formula that determines the number of non-trivial 4-APs in $S_p$ can be given in the following result, and it depends on the number of points on the elliptic curve

$$E : y^2 = x(x+3)(x+4).$$

Since we will use elliptic curves in some of our formulas, it would be necessary to point out that an important aspect of the study of elliptic curves is devising effective ways of counting points on the curve. There are several approaches to do so, and the algorithms devised have been proved to be useful tools in the study of various fields, see [29,38,39].

We also note that the error term in the following theorem is sharp.

**Theorem 1.3.** *Let $p > 3$ be a prime number. The number of non-trivial 4-APs in $S_p$ is given by the following formula:*

$$\begin{aligned}&\frac{(p+1)^4}{16p^2}+\frac{(p-1)(5p+1)}{16p^2}-\frac{p+1}{2}\\&\qquad+\frac{p-1}{16}\cdot\left(\frac{-1}{p}\right)\cdot\left(2\cdot\left(\frac{-6}{p}\right)+4\cdot\left(\frac{-2}{p}\right)+2\cdot\left(\frac{2}{p}\right)+2\cdot\left(\frac{-3}{p}\right)+2\right)\\&\qquad+\frac{1}{16}\cdot\left(\frac{-1}{p}\right)\cdot(p-1)\left(\#E(\mathbb{F}_p)-p-1\right),\end{aligned}$$

*where the elliptic curve $E$ over $\mathbb{F}_p$ is defined by*

$$E : y^2 = x(x+3)(x+4),$$

*and $\left(\frac{\cdot}{p}\right)$ is the Legendre symbol. Moreover, the number of non-trivial 4-APs in $S_p$ is given by*

$$\frac{p^2}{16}+R_p,$$

*where*

$$|R_p| \leq \frac{1}{8}\cdot p^{\frac{3}{2}}+O(p),$$

*and the error term $R_p$ and the above coefficient $\frac{1}{8}$ are best possible in the sense that $O(p^{\frac{3}{2}})$ cannot be replaced by a smaller function of $p$, and $\frac{1}{8}$ cannot be replaced by a smaller constant.*

The resulting formula for the number of 5-APs in $S_p$ is quite long and involves more elliptic curves. The exact formula can be found in its proof.

**Theorem 1.4.** *Let $p > 3$ be a prime number. There exist explicitly computable polynomials $f \in \mathbb{Z}[X]$, $g \in \mathbb{Z}[X_1, X_2, X_3, X_4]$ and $h_i \in \mathbb{Z}[X_1, X_2]$ with $\deg f = 3$, $\deg_{X_1} g = 3$ and $\deg_{X_1} h_i = 2$ for $i \in \{1,2,3\}$ such that the number of 5-APs in $S_p$ is given by*

$$\frac{(p+1)^5}{32p^3}+\frac{f(p)}{32p^3}+\frac{g\left(p,\left(\frac{-1}{p}\right),\left(\frac{2}{p}\right),\left(\frac{3}{p}\right)\right)}{32p^2}+\sum_{i=1}^{3}\frac{h_i\left(p,\left(\frac{-1}{p}\right)\right)}{32p}\left(\#E_i(\mathbb{F}_p)-p-1\right),$$

*where the elliptic curve $E_i$ over $\mathbb{F}_p$ is defined by*

$$\begin{aligned} E_1 : y^2 &= x(x+3)(x+4), \\ E_2 : y^2 &= x(x+4)(x+6), \\ E_3 : y^2 &= x(x+8)(x+9), \end{aligned}$$

*and $\left(\frac{\cdot}{p}\right)$ is the Legendre symbol. Moreover, the number of 5-APs in $S_p$ is given by*

$$\frac{p^2}{32} + O(p^{\frac{3}{2}}),$$

*and the error term is best possible in the sense that $O(p^{\frac{3}{2}})$ cannot be replaced by a smaller function of $p$.*

Can we take our question above a step further, and give a formula that calculates the number of 3-APs in the set of cubes

$$C_p = \{t^3 : t \in \mathbb{F}_p\}$$

in $\mathbb{F}_p$? In addition to our results which make use of Gauss sums for the number of non-trivial 3-APs in $S_p$, we give the following result using Kummer sums for the number of non-trivial 3-APs in $C_p$.

**Theorem 1.5.** *Let $p$ be a prime number with $p \equiv 1 \pmod 3$. Let $Q_p$ denote the number of non-trivial 3-APs in $C_p$. Then,*

$$Q_p = \frac{(p+2)^3}{27p} + \frac{p-1}{27p}(pc_p + 4p + 8) - \frac{p+2}{3},$$

*where $c_p \in \mathbb{Z}$ with $c_p = O(\sqrt{p})$ is a computable constant which depends on $p$. If $p$ is of the form $u^2 + 27v^2$ for some integers $u$ and $v$ with $u \equiv 2 \pmod 3$, then*

$$Q_p = \frac{(p+2)^3}{27p} + \frac{p-1}{27p}(2up + 12p + 8) - \frac{p+2}{3}.$$

In Table 1, using SageMath [36], we give the calculations of the formulas, we obtained in our theorems above, for some certain values. Note that if $p \not\equiv 1 \pmod 3$, then $C_p = \mathbb{F}_p$ and so $Q_p = p(p-1)$.

**Short Outline of the Paper:** In the next section, we will give the basics of finite Fourier analysis and some fundamental theorems of arithmetic geometry and exponential sums that we will use frequently in the manuscript. Section 3 contains the proof of Theorem 1.1. In Section 4, we will prove Proposition 1.2. The proof of Theorem 1.3 and the proof of Theorem 1.4 are contained in Section 5 and Section 6, respectively. Section 7 consists

**Table 1**
The number of non-trivial 3 and 4-APs in $S_p$ and $C_p$ for prime numbers $p$ between 20 and 50.

| $20 < p < 50$ | #3-APs in $S_p$ | #4-APs in $S_p$ | #3-APs in $C_p$ |
|---|---|---|---|
| 23 | 66 | 44 | 23×22 |
| 29 | 98 | 28 | 29×28 |
| 31 | 120 | 30 | 50 |
| 37 | 162 | 54 | 60 |
| 41 | 220 | 120 | 41×40 |
| 43 | 210 | 84 | 70 |
| 47 | 276 | 138 | 47×46 |

of the proof of Theorem 1.5. Finally, in Section 8, we will give some results concerning Salem sets and the Sárközy problem.

## 2. Preliminaries

In this paper, we make use of Fourier analysis on finite abelian groups. In particular, our main tool will be the Fourier transform of functions which are defined on the finite cyclic group $\mathbb{Z}_N$. Throughout this note, $e_N : \mathbb{Z}_N \to \mathbb{C}$ is defined as $e_N(x) = e^{2\pi i x/N}$ for any $x \in \mathbb{Z}_N$. This function has the following well-known property, which is known as *orthogonality*:

$$\sum_{m \in \mathbb{Z}_N} e_N(mu) = \begin{cases} 0 & \text{if } u \neq 0, \\ N & \text{if } u = 0. \end{cases} \tag{3}$$

Given a function $f : \mathbb{Z}_N \to \mathbb{C}$, its *Fourier transform* $\widehat{f}$ at $m \in \mathbb{Z}_N$ is defined by

$$\widehat{f}(m) = N^{-1} \sum_{x \in \mathbb{Z}_N} e_N(-xm) f(x). \tag{4}$$

Basically, the Fourier transform of $f : \mathbb{Z}_N \to \mathbb{C}$ is another function that is defined as the average of the values $f(x)$ multiplied by the corresponding roots of unity, namely $e_N(-xm)$, $x \in \mathbb{Z}_N$. It has numerous useful properties. Among them, the one we require is the inversion formula. The *inversion formula* states that with the above definition of the Fourier transform, we can recover $f$ from its Fourier coefficients via the formula

$$f(x) = \sum_{m \in \mathbb{Z}_N} e_N(xm) \widehat{f}(m). \tag{5}$$

This basic feature of the Fourier transform appears frequently in the proof of our results. Surely, there are much more practical properties of the Fourier transform. For more detailed information about Fourier analysis on $\mathbb{Z}_N$, one can consult [41].

In the following definition, we describe an arithmetic progression in $\mathbb{Z}$.

**Definition 2.1** *(Arithmetic progressions).* An arithmetic progression of length $k$ ($k$-AP) in $\mathbb{Z}$ is a sequence of $k$-integers such that each difference between two consecutive terms is the same constant.

We say that a set $A \subseteq \mathbb{Z}$ contains arbitrarily long arithmetic progressions if for any $k \in \mathbb{N}$, there is a non-trivial $k$-AP in $A$. There are some distinctions between arithmetic progressions in $\mathbb{Z}$ and $\mathbb{Z}_N$. We define an arithmetic progression in $\mathbb{Z}_N$ in the following way.

**Definition 2.2.** A $k$-term arithmetic progression in $\mathbb{Z}_N$, $x_0, x_2, ..., x_{k-1}$, is a sequence of integers satisfying

$$2x_i \equiv x_{i-1} + x_{i+1} \pmod{N},$$

for all $i = 1, ..., k-2$.

The disadvantage is that arithmetic progressions in $\mathbb{Z}_N$ are not necessarily arithmetic progressions in $\mathbb{Z}$ (they might "wrap around"). For instance, in $\mathbb{Z}_{102}$, $\{65, 100, 33\}$ is a 3-term arithmetic progression but not in $\mathbb{Z}$. Nevertheless, there is a relation between lengths of arithmetic progressions in $\mathbb{Z}$ and $\mathbb{Z}_N$. The following proposition was stated by Bourgain without proof in [7]. Thus, we felt the need to prove this proposition.

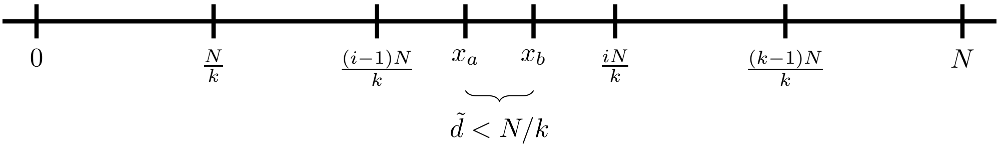
**Proposition 2.3.** *If there exists a non-trivial $\left(2k^2 - 2k + 1\right)$-AP in $\mathbb{Z}_N$, then there is a non-trivial arithmetic progression in $\mathbb{Z}$ of length $k$ contained in this given arithmetic progression in $\mathbb{Z}_N$.*

**Proof.** Let $x_1, x_2, \ldots, x_{2k^2-2k+1}$ be a non-trivial $\left(2k^2 - 2k + 1\right)$-AP in $\mathbb{Z}_N$. Now, we divide the interval $[0, N)$ into $k$ disjoint parts

$$[0, N) = \bigcup_{i=1}^{k} \left[\frac{(i-1)N}{k}, \frac{iN}{k}\right).$$

By the pigeonhole principle, there exist an interval $\left[\frac{(i-1)N}{k}, \frac{iN}{k}\right)$ and $a, b \in \{1, \ldots, k+1\}$ with $a < b$ and $i \in \{1, \ldots, k\}$ such that

$$x_a, x_b \in \left[\frac{(i-1)N}{k}, \frac{iN}{k}\right) \text{ and } x_a \neq x_b :$$



Note that after choosing the appropriate representation of the elements of the arithmetic progression on classes modulo $N$, we consider the elements as integers.

Hence, the set $\left\{x_a, x_b, x_{a+2(b-a)}, \dots, x_{a+(2k-2)(b-a)}\right\}$ has the following properties:

- $|x_{a+(i-1)(b-a)} - x_{a+i(b-a)}| = \tilde{d}$ for $i \in \{1, 2, \dots, 2k-2\}$,
- $a + (2k-2) \cdot (b-a) \leq 2k^2 - 2k + 1$ for $a, b \in \{1, \dots, k+1\}$.

It means that $\left\{x_{a+i(b-a)}\right\}_{i=0}^{2k-2}$ is an arithmetic progression on one of the intervals $(-2N, N)$ and $[0, 3N)$. Now, without loss of generality, we assume that $\left\{x_{a+i(b-a)}\right\}_{i=0}^{2k-2}$ is a $(2k-1)$-AP on $[0, 3N)$. If $\left\{x_{a+i(b-a)}\right\}_{i=0}^{2k-2} \cap [2N, 3N) \neq \emptyset$, then there exists an arithmetic progression of length at least $k$ on the intervals $[N, 2N)$ since $\tilde{d} < N/k$. For the other case, if $\left\{x_{a+i(b-a)}\right\}_{i=0}^{2k-2} \cap [2N, 3N) = \emptyset$, by the pigeonhole principle and as $\tilde{d} < N/k$, there exists an arithmetic progression of length at least $k$ on one of the intervals $[0, N)$ and $[N, 2N)$. Thus, we conclude that there exists a non-trivial $k$-AP in $\mathbb{Z}$ obtained from the $k$-AP on one of the intervals $[0, N)$ and $[N, 2N)$. □

The proposition above provides a way to connect $\mathbb{Z}_N$-progressions to $\mathbb{Z}$-progressions. In particular, finding a $\left(2k^2 - k + 1\right)$-AP in $\mathbb{Z}_N$ gives rise to the existence of a $k$-AP in $\{1, ..., N\}$. In the further parts of this note, we prove that there are long APs in some special subsets of $\mathbb{Z}_N$. Hence, if we lift those sets up to $\mathbb{Z}$, that is to say, see them as a subset of $\mathbb{Z}$, then we obtain $\mathbb{Z}$-APs.

The characteristic function $A(x)$ of a set $A \subseteq \mathbb{Z}_N$ is defined as

$$A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

We will often use the following lemma, which determines the number of $k$-APs in the set $A \subseteq \mathbb{Z}_N$.

**Lemma 2.4.** *Let $A$ be any subset of $\mathbb{Z}_N$ and $k \geq 3$. Then, the number of $k$-APs in the set $A$ is*

$$\frac{N^2 |A|^k}{N^k} + H,$$

*where*

$$\begin{aligned} H = & N^2 \sum_{(x_1, x_2, \dots, x_{k-2}) \neq 0} \widehat{A}(x_1) \cdots \widehat{A}(x_{k-2}) \widehat{A}(x_1 + 2x_2 + \cdots + (k-2)x_{k-2}) \\ & \cdot \widehat{A}(-2x_1 - 3x_2 - \cdots - (k-1)x_{k-2}). \end{aligned}$$

**Proof.** Let $A$ be any subset of $\mathbb{Z}_N$ and $k \geq 3$. Define

$$Q_N(t) = |\{(y_1, y_2, \ldots, y_k) \in A^k : y_{i+1} - y_i = t \text{ for } i \in \{1, \ldots, k-1\}\}|$$

as the number of $k$-term arithmetic progressions in $A$ with common difference $t$.

So, the number of $k$-APs in $A$ is equal to

$$\sum_{t \in \mathbb{Z}_N} Q_N(t) = \sum_{y_1 \in \mathbb{Z}_N} \sum_{t \in \mathbb{Z}_N} A(y_1) A(y_1 + t) \cdots A(y_1 + (k-1)t). \tag{6}$$

If we use the Fourier inversion formula (5) for each of the terms in Equation (6), namely for

$$A(y_1 + t), A(y_1 + 2t), \ldots, A(y_1 + (k-1)t),$$

we obtain the following sums which depend on the Fourier coefficients of $A$:

$$\begin{aligned}\sum_{t \in \mathbb{Z}_N} Q_N(t) = & \sum_{y_1, t \in \mathbb{Z}_N} A(y_1) \sum_{x_0 \in \mathbb{Z}_N} e_N(x_0(y_1 + t)) \widehat{A}(x_0) \\ & \cdots \sum_{x_{k-2} \in \mathbb{Z}_N} e_N(x_{k-2}(y_1 + (k-1)t)) \widehat{A}(x_{k-2}) \\ = & \sum_{(y_1, x_0, x_1, \ldots, x_{k-2})} A(y_1) \widehat{A}(x_0) \cdots \widehat{A}(x_{k-2}) e_N(y_1(x_0 + x_1 + \cdots + x_{k-2})) \\ & \cdot \sum_{t \in \mathbb{Z}_N} e_N(t(x_0 + 2x_1 + \cdots + (k-1)x_{k-2})).\end{aligned}$$

By orthogonality, we have

$$\sum_{t \in \mathbb{Z}_N} e_N(t(x_0 + 2x_1 + \cdots + (k-1)x_{k-2})) = \begin{cases} N & \text{if } x_0 + 2x_1 + \cdots + (k-1)x_{k-2} = 0, \\ 0 & \text{otherwise.} \end{cases}$$

From this orthogonality relation, we obtain that

$$\begin{aligned}\sum_{t \in \mathbb{Z}_N} Q_N(t) = N & \sum_{(y_1, x_1, x_2, \ldots, x_{k-2})} A(y_1) \widehat{A}(x_1) \cdots \widehat{A}(x_{k-2}) \widehat{A}(-2x_1 - \cdots - (k-1)x_{k-2}) \\ & \cdot e_N(y_1(-x_1 - 2x_2 - \cdots - (k-2)x_{k-2})).\end{aligned}$$

Then, one can conclude that

$$\begin{aligned}\sum_{t \in \mathbb{Z}_N} Q_N(t) = N^2 & \sum_{x_1, x_2, \ldots, x_{k-2}} \widehat{A}(x_1) \cdots \widehat{A}(x_{k-2}) \widehat{A}(x_1 + 2x_2 + 3x_3 + \cdots + (k-2)x_{k-2}) \\ & \cdot \widehat{A}(-2x_1 - 3x_2 - \cdots - (k-1)x_{k-2}).\end{aligned}$$

We denote this again by

$$\sum_{t\in\mathbb{Z}_N} Q_N(t) = \frac{N^2|A|^k}{N^k} + H$$

where

$$\begin{aligned} H =& N^2 \sum_{(x_1,x_2,\ldots,x_{k-2})\neq 0} \widehat{A}(x_1)\cdots\widehat{A}(x_{k-2})\widehat{A}(x_1+2x_2+\cdots+(k-2)x_{k-2}) \\ & \cdot \widehat{A}(-2x_1-3x_2-\cdots-(k-1)x_{k-2}), \end{aligned}$$

and the proof is complete. □

Now, we introduce the notion of a Salem family. These are certain families of sets that lie in $\mathbb{Z}_N$. We call the sets in a Salem family as Salem sets. Lott [30] showed that it is possible to guarantee the existence of a 3-AP in Salem sets. In the last section, we will focus on the existence of long arithmetic progressions in Salem sets.

**Definition 2.5** *(Generalized Salem family).* Let $\{A_N\}_{N\in\mathcal{B}}$ be a family of sets with $A_N \subseteq \mathbb{Z}_N$ and $\alpha \in (0,1)$, where $\mathcal{B} \subseteq \mathbb{Z}_{>0}$ is infinite. The family $\{A_N\}_{N\in\mathcal{B}}$ is said to be an $\alpha$-Salem family, if there exists a constant $C$ depending on $\alpha$ such that for all $N \in \mathcal{B}$ and all nonzero $m \in \mathbb{Z}_N$,

$$|\widehat{A}_N(m)| \leq C \cdot N^{-1}|A_N|^{\alpha},$$

and the constant $C$ is called the $\alpha$-Salem constant.

The definition of a Salem family is based on the magnitude of the values of its Fourier transform. For instance, for a positive integer $n \geq 2$, one can consider

$$\Omega_p^n = \{t^n : t \in \mathbb{F}_p\},$$

and the sets $\{\Omega_p^n\}_{p\in\mathbb{P}}$ constitute a $\frac{1}{2}$-Salem family, where $\mathbb{P}$ is the set of prime numbers. Proof of this can be obtained from [48]. Now, we continue to give the necessary background in order to prove our results and obtain some properties of the families $\{S_p\}_{p\in\mathbb{P}}$ and $\{\Omega_p^n\}_{p\in\mathbb{P}}$.

**Definition 2.6.** Let $p$ be an odd prime number. An integer $a$ which is not divisible by $p$ is said to be a quadratic residue modulo $p$ if it is congruent to a perfect square modulo $p$ and is a quadratic nonresidue modulo $p$ otherwise. The Legendre symbol is a function of $a$ and $p$ and it is defined as

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & \text{if } a \text{ is a quadratic residue modulo } p, \\ -1 & \text{if } a \text{ is a quadratic nonresidue modulo } p, \\ 0 & \text{if } p \mid a. \end{cases}$$

Below we list three important properties of the Legendre symbol which will be frequently used.

**Proposition 2.7** *([32]). For integers $b$ and $c$ with $p \nmid b$,*

$$\sum_{\ell=0}^{p-1} \left(\frac{b\ell + c}{p}\right) = 0.$$

**Proposition 2.8** *([32]). Let $a, b$ and $c$ be integers, and let $p$ be an odd prime. Then*

$$\sum_{\ell=0}^{p-1} \left(\frac{a\ell^2 + b\ell + c}{p}\right) = \begin{cases} -\left(\frac{a}{p}\right) & \text{if } p \nmid \left(b^2 - 4ac\right), \\ \left(\frac{a}{p}\right)(p-1) & \text{if } p \mid \left(b^2 - 4ac\right). \end{cases}$$

**Remark 2.9.** Let $k$ be an odd positive integer and $m \geq 1$. Let $a_{ij} \in \mathbb{F}_p$ for $1 \leq i \leq k$ and $1 \leq j \leq m$. Then

$$\sum_{x_1, x_2, \ldots, x_m \in \mathbb{F}_p} \left(\frac{a_{11}x_1 + \cdots + a_{1m}x_m}{p}\right) \cdots \left(\frac{a_{k1}x_1 + \cdots + a_{km}x_m}{p}\right) = 0.$$

This equation is quickly obtained by defining new variables $y_j = ax_j$ for a chosen $a \in \mathbb{F}_p$ with $\left(\frac{a}{p}\right) = -1$. This property will be used frequently without being specified in the following sections.

In 1924, Artin estimated the correctness of the following theorem on elliptic curves. However, Artin was not able to prove his estimate. In 1933, Hasse proved the estimate of Artin. Then, Weil generalized the result of Hasse, as we mentioned in the previous section. The following two theorems will play an important role in finding the number of arithmetic progressions of length 4 and 5 in $S_p$.

**Theorem 2.10** *(Hasse [47, Theorem 4.2]). Let $E$ be an elliptic curve over the finite field $\mathbb{F}_p$. Then, the order of $E(\mathbb{F}_p)$ satisfies*

$$|p + 1 - \#E(\mathbb{F}_p)| \leq 2\sqrt{p}.$$

**Theorem 2.11** *([47, Theorem 4.14]). Let $E$ be an elliptic curve defined by $y^2 = x^3 + Ax + B$ over the finite field $\mathbb{F}_p$ where $p$ is an odd prime. Then,*

$$\#E(\mathbb{F}_p) = p + 1 + \sum_{x \in \mathbb{F}_p} \left(\frac{x^3 + Ax + B}{p}\right).$$

Let $k \geq 6$ be a positive integer. When estimating the number of arithmetic progressions of length $k$ in $S_p$, we need a conclusion that yields more than Hasse's theorem. In 2022, Rojas-León proved an estimate for multi-variable multiplicative character sums over affine subspaces of $\mathbb{A}_k^n$, which generalizes the well-known estimates for both classical Jacobi sums and one-variable polynomial multiplicative character sums [34]. The following theorem proved by Rojas-León is of fundamental importance in finding the number of arithmetic progressions of length $k$ in $S_p$ with a better error term.

**Theorem 2.12** *([34, Corollary 2]). Let $k = \mathbb{F}_q$ be a finite field, with $q = p^a$ a prime power. Let $\chi_1, \ldots, \chi_n : \mathbb{F}_q^\times \to \mathbb{C}^\times$ be $n$ non-trivial multiplicative characters. Let $L_1, \ldots, L_n : \mathbb{A}_k^d \to \mathbb{A}_k^1$ be affine linear forms, with $L_i(t) = a_{i,1}t_1 + \cdots + a_{i,d}t_d + b_i$, and let $V_i \subseteq \mathbb{A}_k^d$ be the hyperplane defined by $L_i(t) = 0$. Suppose that the affine map $\mathbb{A}_k^d \to \mathbb{A}_k^n$ defined by the $L_i$ is injective (that is, that the matrix $(a_{ij})$ has rank d), and that for every $I \subseteq \{1, \ldots, n\}$ with $|I| \leq d + 1$ we have $\dim(\cap_{i\in I} V_i) \leq d - |I|$. Then, we have the estimate*

$$\left| \sum_{t\in k^d} \chi_1 (L_1(t)) \cdots \chi_n (L_n(t)) \right| \leq D_L \cdot q^{d/2},$$

*where*

$$D_L := (-1)^d + \sum_{j=1}^{d} (-1)^{d+j} a_j,$$

*and $a_j$ is the number of subsets $I \subseteq \{1, \ldots, n\}$ with $|I| = j$ such that $\cap_{i\in I} V_i \neq \emptyset$.*

Observe that the algebraic sets occurring in the previous theorem are highly singular, so one cannot apply the results of [28] and [33] immediately. Although the character sum estimates are in the realm of analytic number theory, the technique behind them is the use of $\ell$-adic cohomology and Grothendieck's trace formula, see also the works of Deligne [15,18].

Given that $p$ and $q$ are two distinct odd primes, suppose we know whether $q$ is a quadratic residue of $p$ or not. The natural question is as follows: will $p$ be a quadratic residue of $q$? One of Gauss' favorite theorems, which is the law of quadratic reciprocity answers this question. The law of quadratic reciprocity is a very deep theorem with over two hundred fifty proofs.

**Proposition 2.13** *([6, Theorem 1.2.6] Law of quadratic reciprocity). Let $p$ and $q$ be two distinct odd prime numbers. Then,*

$$\left(\frac{p}{q}\right) \cdot \left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2} \cdot \frac{q-1}{2}}$$

*holds.*

For an odd prime number $p$, an integer $a$ and $k \in \mathbb{Z}_{>0}$, a general Gauss sum is defined as

$$G_k(a,p) = \sum_{m=0}^{p-1} e_p(am^k). \tag{7}$$

When $k = 1$ and $p \nmid a$, as mentioned before, the sum of all $p$-th roots of unity, which is a geometric sum and can be easily evaluated to be zero. When $k \geq 2$, the task of determining the sum then becomes considerably more difficult. In fact, even for the initial case $k = 2$, it took Gauss several years to accomplish this. In late May of 1801, Gauss conjectured that

$$G_2(1,p) = \begin{cases} \sqrt{p} & \text{if } p \equiv 1 \pmod 4, \\ i\sqrt{p} & \text{if } p \equiv 3 \pmod 4. \end{cases} \tag{8}$$

On August 30, 1805, Gauss wrote in his diary that he devoted some time to this problem every week for more than four years before he was able to prove his conjecture on the signs of these sums [5]. The sum $G_2(a,p)$ introduced by Gauss in 1801 is now called the quadratic Gauss sum.

**Theorem 2.14** *([6, Theorem 1.5.2]). Let $a$ be an integer not divisible by a prime $p > 2$. Then*

$$G_2(a,p) = \sum_{m=0}^{p-1} e_p(am^2) = \left(\frac{a}{p}\right) G_2(1,p) = \begin{cases} \left(\frac{a}{p}\right)\sqrt{p} & \textit{if } p \equiv 1 \pmod 4, \\ i\left(\frac{a}{p}\right)\sqrt{p} & \textit{if } p \equiv 3 \pmod 4. \end{cases}$$

We also recall Weil's theorem from the introduction.

**Theorem 2.15** *([48]). Let $p$ be an odd prime number. Let $f \in \mathbb{Z}[X]$ be a non-linear polynomial such that $f \notin p\mathbb{Z}[X]$. We denote the Weil sum by*

$$s(f,p) = \sum_{x \in \mathbb{F}_p} e_p\left(f(x)\right),$$

*where $e_p(x) = e^{2\pi i x/p}$. Then, we have*

$$|s(f,p)| \leq (\deg f - 1) \cdot \sqrt{p}.$$

In additive number theory, another theorem that comes on the scene, which is in a similar spirit of Szemerédi's theorem, is Sárközy's theorem. Sárközy [37] and Furstenberg [20] independently proved the following result in the late 1970s, now commonly known as Sárközy's theorem:

**Theorem 2.16** *(Sárközy's theorem, [37]). If $A$ is a subset of positive integers with a positive upper density, then there are two distinct elements of $A$ whose difference is a perfect square.*

In the last section, we will get some results considering the generalized Sárközy problem for $\alpha$-Salem families.

## 3. Proof of Theorem 1.1

**Proof of Theorem 1.1.** Let $p > k$ be an odd prime number. Let

$$Q_p(t) = |\{(x_1, \dots, x_k) \in S_p^k \mid x_{i+1} - x_i = t \text{ for } i \in \{1, \dots, k-1\}\}| \tag{9}$$

denote the number of $k$-term arithmetic progressions in $S_p$ with common difference $t$. By Lemma 2.4, the number of $k$-APs in $S_p$ is equal to

$$\frac{p^2 |S_p|^k}{p^k} + R,$$

where

$$R = p^2 \sum_{(x_1, x_2, \dots, x_{k-2}) \neq 0} \widehat{S_p}(x_1) \cdots \widehat{S_p}(x_{k-2}) \widehat{S_p}(x_1 + 2x_2 + \cdots + (k-2)x_{k-2})$$
$$\cdot \widehat{S_p}(-2x_1 - 3x_2 - \cdots - (k-1)x_{k-2}).$$

Note that when $0 \neq m \in \mathbb{F}_p$,

$$\begin{aligned}
\widehat{S}_p(m) &= \frac{1}{p} \sum_{x \in \mathbb{F}_p} e_p(-mx) S_p(x) \\
&= \frac{1}{p} \sum_{x \in S_p} e_p(-mx) \\
&= \frac{1}{2p} \left( \sum_{t \in \mathbb{F}_p} e_p(-mt^2) + 1 \right) \\
&= \frac{1}{2p} \left( \varepsilon \left( \frac{-m}{p} \right) \sqrt{p} + 1 \right) \\
&= \frac{1}{2p} + \frac{1}{2\sqrt{p}} \varepsilon \left( \frac{-m}{p} \right),
\end{aligned}$$

where

$$\varepsilon = \begin{cases} 1 & \text{if } p \equiv 1 \pmod 4 \\ i & \text{if } p \equiv 3 \pmod 4, \end{cases}$$

and $\widehat{S}_p(0) = \dfrac{|S_p|}{p} = \dfrac{p+1}{2p}$ in case $m = 0$.

Next, we find the upper bound mentioned for the error term $R$. By the expression of $R$ and the Fourier transform of $S_p$ which were given above, it is sufficient to find an upper bound for the following expressions in forms $A, B, C$ and $D$ since they are the largest ones:

$$A = p^2 \cdot \left(\frac{p+1}{2p}\right)^{k-2-m} \cdot \frac{1}{(2\sqrt{p})^{m+2}} \sum_{x_{a_1},\dots,x_{a_m}} \left(\frac{-x_{a_1}}{p}\right) \cdots \left(\frac{-x_{a_m}}{p}\right) \cdot \left(\frac{-a_1 x_{a_1} - \cdots - a_m x_{a_m}}{p}\right) \left(\frac{(a_1+1)\,x_{a_1} + \cdots + (a_m+1)\,x_{a_m}}{p}\right), \tag{10}$$

$$B = p^2 \cdot \left(\frac{p+1}{2p}\right)^{k-1-m} \cdot \frac{1}{(2\sqrt{p})^{m+1}} \sum_{\substack{x_{a_1},\dots,x_{a_m} \\ (a_1+1)x_{a_1}+\cdots+(a_m+1)x_{a_m}=0}} \left(\frac{-x_{a_1}}{p}\right) \cdots \cdot \left(\frac{-x_{a_m}}{p}\right) \left(\frac{-a_1 x_{a_1} - \cdots - a_m x_{a_m}}{p}\right), \tag{11}$$

$$C = p^2 \cdot \left(\frac{p+1}{2p}\right)^{k-1-m} \cdot \frac{1}{(2\sqrt{p})^{m+1}} \sum_{\substack{x_{a_1},\dots,x_{a_m} \\ -a_1 x_{a_1}-\cdots-a_m x_{a_m}=0}} \left(\frac{-x_{a_1}}{p}\right) \cdots \cdot \left(\frac{-x_{a_m}}{p}\right) \left(\frac{(a_1+1)\,x_{a_1} + \cdots + (a_m+1)\,x_{a_m}}{p}\right), \tag{12}$$

$$D = p^2 \cdot \left(\frac{p+1}{2p}\right)^{k-m} \cdot \frac{1}{(2\sqrt{p})^{m}} \sum_{\substack{x_{a_1},\dots,x_{a_m} \\ -a_1 x_{a_1}-\cdots-a_m x_{a_m}=0 \\ (a_1+1)x_{a_1}+\cdots+(a_m+1)x_{a_m}=0}} \left(\frac{-x_{a_1}}{p}\right) \cdots \left(\frac{-x_{a_m}}{p}\right), \tag{13}$$

where $a_i \in \{1, 2, \dots, k-2\}$ such that $a_i \neq a_j$ when $i \neq j$, and $m \in \{2, \dots, k-2\}$. Now, we observe that in case of $m = 1$, the equations from (10) to (13) are equal to zero. By the properties of the Legendre symbol, we get

$$\sum_{x_{a_1}} \left(\frac{-x_{a_1}}{p}\right) \left(\frac{-a_1 x_{a_1}}{p}\right) \left(\frac{(a_1+1)\,x_{a_1}}{p}\right) = \left(\frac{a_1\,(a_1+1)}{p}\right) \sum_{x_{a_1}} \left(\frac{x_{a_1}}{p}\right).$$

Then, it follows from orthogonality that

$$\left(\frac{a_1\,(a_1+1)}{p}\right) \sum_{x_{a_1}} \left(\frac{x_{a_1}}{p}\right) = 0.$$

Thus, equation (10) is equal to zero. As $a_1 \in \{1, 2, \ldots, k-2\}$, we also have

$$\sum_{\substack{x_{a_1} \\ (a_1+1)x_{a_1}=0}} \left(\frac{-x_{a_1}}{p}\right)\left(\frac{-a_1 x_{a_1}}{p}\right) = 0.$$

Similarly, the other equations are shown to be equal to zero.

We first find the upper bound for (10). Since we cannot apply Theorem 2.12 immediately to the sum, we bring the expressions to the appropriate forms. Using change of variables, namely $x_{a_i} = x_{a_m} x_{a_i}$ for $i \in \{1, \ldots, m-1\}$, and by the properties of the Legendre symbol, $A$ above becomes

$$\begin{aligned} &\frac{p^2 \cdot (p+1)^{k-2-m}}{2^k \cdot p^{k-2-m} \cdot p^{\frac{m+2}{2}}} \cdot p \sum_{x_{a_1}, \ldots, x_{a_{m-1}}} \left(\frac{-x_{a_1}}{p}\right) \cdots \left(\frac{-x_{a_{m-1}}}{p}\right) \\ &\quad \cdot \left(\frac{-a_1 x_{a_1} - \cdots - a_{m-1} x_{a_{m-1}} - a_m}{p}\right) \\ &\quad \cdot \left(\frac{(a_1+1)\, x_{a_1} + \cdots + (a_{m-1}+1)\, x_{a_{m-1}} + a_m + 1}{p}\right). \end{aligned} \tag{14}$$

Now, we calculate (14) with the help of Theorem 2.12.

Take affine linear forms $L_1, \ldots, L_{m+1} : \mathbb{A}_k^{m-1} \to \mathbb{A}_k^1$ as

$$\begin{aligned} L_1(x) &= -1 \cdot x_{a_1} \\ L_2(x) &= -1 \cdot x_{a_2} \\ &\ \ \vdots \\ L_{m-1}(x) &= -1 \cdot x_{a_{m-1}} \\ L_m(x) &= -a_1 \cdot x_{a_1} + \cdots + -a_{m-1} \cdot x_{a_{m-1}} - a_m \\ L_{m+1}(x) &= (a_1+1) \cdot x_{a_1} + \cdots + (a_{m-1}+1) \cdot x_{a_{m-1}} + (a_m+1)\,. \end{aligned}$$

The affine map defined by $L_i$ is injective since the matrix $(a_{ij})$ has rank $m-1$. Now, let $V_i \subseteq \mathbb{A}_k^{m-1}$ be the hyperplane defined by $L_i(x) = 0$ for each $i \in \{1, \ldots, m+1\}$, and $I \subset \{1, \ldots, m+1\}$ be a subset with $|I| \leq m$. When $|I| = m$,

$$\bigcap_{i \in I} V_i = \emptyset,$$

that is to say $\dim\left(\cap_{i \in I} V_i\right) = -1$. When $|I| \leq m-1$,

$$\dim\left(\bigcap_{i \in I} V_i\right) \leq m - 1 - |I|$$

holds since the intersections of the hyperplanes $V_i$ do not coincide with themselves. Hence, by applying Theorem 2.12, we get the following inequality

$$|A| \leq \frac{1}{2^k} \cdot D_L \cdot p^{\frac{3}{2}} + O_k(p^{\frac{1}{2}}), \tag{15}$$

where

$$D_L := (-1)^{m-1} + \sum_{j=1}^{m-1} (-1)^{m-1+j} c_j$$

and $c_j$ is the number of subsets $I \subset \{1, \ldots, m+1\}$ with $|I| = j$ such that

$$\bigcap_{i \in I} V_i \neq \emptyset.$$

Notice that

$$c_j = \binom{m+1}{j}$$

when $j \in \{1, ..., m-1\}$. Using the well-known identity

$$\sum_{j=0}^{m+1} (-1)^j \binom{m+1}{j} = 0,$$

the value $D_L$ becomes

$$\begin{aligned} &(-1)^{m-1} + \sum_{j=1}^{m-1} (-1)^{m-1+j} \binom{m+1}{j} \\ &\quad = (-1)^{m-1} + (-1)^m (1 + (-1)^m (m+1) + (-1)^{m+1}) = m. \end{aligned} \tag{16}$$

Therefore, (15) and the previous equality yield that

$$|A| \leq \frac{1}{2^k} \cdot m \cdot p^{\frac{3}{2}} + O_k(p^{\frac{1}{2}}).$$

Now, we rewrite the other forms and bring them to the form $A$. Note that the properties of the Legendre symbol and change of variables will be used again. If we arrange the indices on (11) and (12) using

$$(a_1 + 1)\, x_{a_1} + \cdots + (a_m + 1)\, x_{a_m} = 0 \ \text{ and } \ - a_1 x_{a_1} - \cdots - a_m x_{a_m} = 0,$$

we get the following sums:

$$\begin{aligned}
&\frac{p^2\cdot(p+1)^{k-1-m}}{2^k\cdot p^{k-1-m}\cdot p^{\frac{m+1}{2}}}\sum_{x_{a_1},\dots,x_{a_{m-1}}}\left(\frac{-x_{a_1}}{p}\right)\cdots\left(\frac{-x_{a_{m-1}}}{p}\right)\\
&\quad\cdot\left(\frac{(a_1+1)\,x_{a_1}+\cdots+(a_{m-1}+1)\,x_{a_{m-1}}}{p}\right)\\
&\quad\cdot\left(\frac{(a_m-a_1)\,x_{a_1}+\cdots+(a_m-a_{m-1})\,x_{a_{m-1}}}{p}\right),
\end{aligned}\tag{17}$$

$$\begin{aligned}
&\frac{p^2\cdot(p+1)^{k-1-m}}{2^k\cdot p^{k-1-m}\cdot p^{\frac{m+1}{2}}}\sum_{x_{a_1},\dots,x_{a_{m-1}}}\left(\frac{-x_{a_1}}{p}\right)\cdots\left(\frac{-x_{a_{m-1}}}{p}\right)\\
&\quad\cdot\left(\frac{a_1x_{a_1}+\cdots+a_{m-1}x_{a_{m-1}}}{p}\right)\left(\frac{(a_m-a_1)\,x_{a_1}+\cdots+(a_m-a_{m-1})\,x_{a_{m-1}}}{p}\right).
\end{aligned}\tag{18}$$

Similarly, if the same method as in form $A$ is applied for (17) and (18), we get

$$|B|\le\frac{m}{2^k}\cdot p^{\frac{3}{2}}+O_k(p^{\frac{1}{2}}),\tag{19}$$

$$|C|\le\frac{m}{2^k}\cdot p^{\frac{3}{2}}+O_k(p^{\frac{1}{2}}).\tag{20}$$

If we first use $-a_1x_{a_1}-\cdots-a_mx_{a_m}=0$ to rewrite the indices of the form $D$, we get the following sum in order to find the upper bound for (13):

$$\begin{aligned}
&\frac{p^2\cdot(p+1)^{k-m}}{2^k\cdot p^{k-m}\cdot p^{\frac{m}{2}}}\sum_{\substack{x_{a_1},\dots,x_{a_{m-1}}\\(1-a_m^{-1}a_1)x_{a_1}+\cdots+(1-a_m^{-1}a_{m-1})x_{a_{m-1}}=0}}\left(\frac{-x_{a_1}}{p}\right)\cdots\left(\frac{-x_{a_{m-1}}}{p}\right)\\
&\quad\cdot\left(\frac{a_1x_{a_1}+\cdots+a_{m-1}x_{a_{m-1}}}{p}\right).
\end{aligned}\tag{21}$$

Then, again if the same method is used as in forms $B$ and $C$, we deduce that

$$|D|\le\frac{m}{2^k}\cdot p^{\frac{3}{2}}+O_k(p^{\frac{1}{2}}).\tag{22}$$

There can be at most $2^{k-2}$ expressions for (10), (11), (12) and (13) in the error term $R$. We also know that half of these expressions are zero by Remark 2.9. Thus, we deduce the upper bound for $R$ as

$$|R|\le 2\cdot\sum_{m=2}^{k-2}\binom{k-2}{m}\cdot\frac{m}{2^k}\cdot p^{\frac{3}{2}}+O_k(p)=\left(\frac{k-2}{4}-\frac{k-2}{2^{k-1}}\right)\cdot p^{\frac{3}{2}}+O_k(p)$$

using the equality

$$\sum_{m=2}^{k-2}\binom{k-2}{m}\cdot m=\sum_{m=1}^{k-2}\binom{k-2}{m}\cdot m-(k-2)=(k-2)\cdot 2^{k-3}-(k-2).$$

Moreover, if the sums for the error term contributing to $p$ are determined and calculated, as in our proof, the constant $c_k$ can be found explicitly. □

**An Application of Theorem 1.1** Let $\{S_{p_i}\}_{i\in\mathbb{N}}$ be a sequence of sets such that $p_1 = 5$ and $p_i$ is a prime number. This time, we consider $S_{p_i}$ as a subset of $\{1, \dots, p_i\} \subset \mathbb{N}$. Let us make an assumption for now. Assume that when $i > j \geq 1$,

$$S_{p_i} \cap \{1, 2, \dots, p_j\} = S_{p_j}.$$

Let us define

$$A = \bigcup_{i\geq 1} S_{p_i}.$$

**Claim:** The set $A$ contains arbitrarily long arithmetic progressions.

**Proof.** Let $k \geq 3$ be a positive integer. By Theorem 1.1, for a sufficiently large prime number $p_i$, the set $S_{p_i}$ contains non-trivial arithmetic progressions of length $2k^2 - 2k + 1$ modulo $p_i$. It follows from Proposition 2.3 that $S_{p_i}$ contains non-trivial arithmetic progressions of length $k$ in $\mathbb{Z}$. Hence, $A$ contains arbitrarily long arithmetic progressions. (Thus, we proved the claim without using Szemerédi's theorem.) □

Now, let us prove the above assumption, namely the existence of such sequences.

**Proposition 3.1.** *Let $q$ be a prime number such that $q \equiv 5 \pmod 8$ and $S_q$ be the set of quadratic residues modulo $q$, that is $S_q = \{k \in \{1, 2, \dots, q\} : x^2 \equiv k \pmod q$ for some $x\}$. Then, there exists a prime number $p > q$ such that $p \equiv 5 \pmod 8$ with*

$$S_p \cap \{1, 2, \dots, q\} = S_q.$$

**Proof.** Let $q$ be a prime number such that $q \equiv 5 \pmod 8$. Now, let us divide the primes in $\{1, 2, \dots, q\}$ into two sets according to be quadratic or quadratic nonresidue modulo $q$. Let $p_1, \dots, p_k \in \{1, 2, \dots, q\}$ be the list of primes where $\left(\dfrac{p_i}{q}\right) = 1$ and $2 = q_1, \dots, q_r \in \{1, 2, \dots, q\}$ be the list of primes where $\left(\dfrac{q_i}{q}\right) = -1$.

For $i \in \{2, \dots, r\}$, choose $a_i \in \{1, \dots, q_i\}$ such that $\left(\dfrac{a_i}{q_i}\right) = -1$. Consider the following congruences:

$$\begin{aligned} X &\equiv 5 \pmod 8, \\ X &\equiv 1 \pmod{p_i}, \text{for each } i \in \{1, 2, \dots, k\}, \\ X &\equiv a_i \pmod{q_i}, \text{for each } i \in \{2, 3, \dots, r\}. \end{aligned}$$

By Chinese Remainder theorem, the solution set is an arithmetic progression

$$(a + n \cdot 8p_1 \cdots p_k q_2 \cdots q_r)_n ,$$

where $0 \leq a < 8p_1 \cdots p_k q_2 \cdots q_r$. Moreover, $\gcd(a, 8p_1 \cdots p_k q_2 \cdots q_r) = 1$. Recall Dirichlet's theorem on arithmetic progressions [2, Theorem 7.9], which states that if $a$ and $\ell$ are relatively prime positive integers, then there are infinitely many primes of the form $a + n\ell$ with $n \in \mathbb{N}$. By Dirichlet's theorem, the above arithmetic progression contains a prime number, say $p > q$. Combining the law of quadratic reciprocity and $p \equiv 5 \pmod 8$, we obtain that for any odd prime $s$,

$$\left(\frac{p}{s}\right) \cdot \left(\frac{s}{p}\right) = (-1)^{\frac{p-1}{2} \cdot \frac{s-1}{2}} = 1.$$

Thus, $\left(\frac{p}{s}\right) = 1$ if and only if $\left(\frac{s}{p}\right) = 1$. As $\left(\frac{p}{p_i}\right) = \left(\frac{1}{p_i}\right) = 1$ for each $i \in \{1, 2, \dots, k\}$, we get $\left(\frac{p_i}{p}\right) = 1$. Similarly, as $\left(\frac{p}{q_i}\right) = \left(\frac{a_i}{q_i}\right) = -1$ for each $i \in \{2, 3, \dots, r\}$, we have $\left(\frac{q_i}{p}\right) = -1$. Also, as $p \equiv 5 \pmod 8$, we obtain that $\left(\frac{2}{p}\right) = -1$. This completes the proof. □

## 4. Some applications of quadratic Gauss sums

In this section, we prove Proposition 1.2 using quadratic Gauss sums.

**Proof of Proposition 1.2.** Let $p$ be an odd prime. Let

$$Q_p(t) = |\{(x, y, z) \in S_p^3 \mid y - x = z - y = t\}| \tag{23}$$

denote the number of 3-term arithmetic progressions in $S_p$ with common difference $t$.

By Lemma 2.4, we have that

$$\sum_{t \in \mathbb{F}_p} Q_p(t) = \frac{p^2 |S_p|^3}{p^3} + p^2 \sum_{\ell \neq 0} \widehat{S}_p(\ell) \widehat{S}_p(\ell) \widehat{S}_p(-2\ell). \tag{24}$$

Recall that when $0 \neq m \in \mathbb{F}_p$,

$$\widehat{S}_p(m) = \frac{1}{2p} + \frac{1}{2\sqrt{p}} \varepsilon \left(\frac{-m}{p}\right),$$

where

$$\varepsilon = \begin{cases} 1 & \text{if } p \equiv 1 \pmod 4, \\ i & \text{if } p \equiv 3 \pmod 4. \end{cases}$$

Using $\sum_{\ell\neq 0}\left(\frac{\ell}{p}\right)=0$, we conclude that

$$\begin{aligned} p^2\sum_{\ell\neq 0}\widehat{S}_p(\ell)\widehat{S}_p(\ell)\widehat{S}_p(-2\ell) &= p^2\sum_{\ell\neq 0}\left(\frac{1}{2p}\left(1+\varepsilon\left(\frac{-\ell}{p}\right)\sqrt{p}\right)\right)^2\left(\frac{1}{2p}\left(1+\varepsilon\left(\frac{2\ell}{p}\right)\sqrt{p}\right)\right) \\ &= \frac{1}{8p}\sum_{\ell\neq 0}\left(2\varepsilon^2 p\left(\frac{-2}{p}\right)+\varepsilon^2 p+1\right). \end{aligned}$$

Note that

$$\left(\frac{-1}{p}\right)=(-1)^{\frac{p-1}{2}}=\begin{cases}1 & \text{if } p\equiv 1 \pmod 4,\\ -1 & \text{if } p\equiv 3 \pmod 4\end{cases} \tag{25}$$

and

$$\left(\frac{2}{p}\right)=(-1)^{\frac{p^2-1}{8}}=\begin{cases}1 & \text{if } p\equiv 1 \text{ or } 7 \pmod 8,\\ -1 & \text{if } p\equiv 3 \text{ or } 5 \pmod 8.\end{cases} \tag{26}$$

When we assume $p\equiv 1 \pmod 8$, it follows from equations (25) and (26) that

$$p^2\sum_{\ell\neq 0}\widehat{S}_p(\ell)\widehat{S}_p(\ell)\widehat{S}_p(-2\ell)=\frac{1}{8p}\left(3p+1\right)\left(p-1\right).$$

Hence, we find that the number of non-trivial 3-APs in $S_p$ is

$$\begin{aligned} \sum_{t\in\mathbb{Z}_p} Q_N(t)-|S_p| &= \left(p^2\left(\frac{p+1}{2p}\right)^3+\frac{(p-1)\left(3p+1\right)}{8p}\right)-\left(\frac{p+1}{2}\right) \\ &= \frac{(p-1)(p+3)}{8}. \end{aligned}$$

In addition, if the processes are done by considering, respectively, conditions $p\equiv 3 \pmod 8$, $p\equiv 5 \pmod 8$ and $p\equiv 7 \pmod 8$ in the same way, we obtain the following formulas

$$\frac{(p-1)(p-3)}{8},\ \frac{(p-1)(p-1)}{8}\ \text{ and }\ \frac{(p-1)(p+1)}{8}.\quad\square$$

The above proposition actually gives the number of non-trivial solutions of the Diophantine congruence $x^2+y^2\equiv 2z^2 \pmod p$. Now, we consider this situation from another perspective. When we look at the significant developments on arithmetic progressions in recent years, it can be seen that the Diophantine equation $x^n+y^n=2z^n$ has no non-trivial primitive solutions in $\mathbb{Z}_{>0}$ when $n\geq 3$, and this was proved by Darmon and

Merel [11]. An integer solution $(x, y, z)$ is called primitive if $\gcd(x, y, z) = 1$. In contrast to the Darmon-Merel Theorem, we will observe that for $n \geq 3$, the congruence $x^n + y^n \equiv 2z^n \pmod{p}$ has a non-trivial solution when $p$ is a sufficiently large prime number. In order to get this observation, it is enough to show that there exist non-trivial arithmetic progressions of length 3 in $\Omega_p^n = \{x^n : x \in \mathbb{F}_p\}$, and this can be achieved by van der Waerden's theorem [46].

**Remark 4.1.** For $n \geq 3$, $\Omega_p^n = \{x^n : x \in \mathbb{F}_p\}$ contains non-trivial arithmetic progressions of length 3 when $p$ is a sufficiently large prime number.

**Proof.** Let $m = \left|\mathbb{F}_p^\times / \left(\mathbb{F}_p^\times\right)^n\right|$. Note that $1 \leq m \leq n$. Let $\{g_1, g_2, \ldots, g_m\}$ be a set of representatives of $\mathbb{F}_p^\times / \left(\mathbb{F}_p^\times\right)^n$, in other words

$$\mathbb{F}_p^\times = \bigsqcup_{i=1}^{m} g_i \left(\mathbb{F}_p^\times\right)^n .$$

Define a coloring $\psi$ of $\{1, 2, \ldots, p-1\}$ by $m$-many colors as follows. For each $a \in \{1, 2, \ldots, p-1\}$, there is a unique $g_i$ such that $a \in g_i \left(\mathbb{F}_p^\times\right)^n$. Set $\psi(a) = i$. By van der Waerden's theorem [46], if $p$ is large enough, there are distinct elements $x, y, z \in \mathbb{F}_p^\times$ such that

$$x + y = 2z \text{ and } \psi(x) = \psi(y) = \psi(z) = i.$$

As we can write $x = g_i x_1^n$, $y = g_i y_1^n$, $z = g_i z_1^n$, we obtain nonzero distinct elements $x_1^n, y_1^n, z_1^n \in \left(\mathbb{F}_p^\times\right)^n = \Omega_p^n \setminus \{0\}$ such that $x_1^n + y_1^n = 2z_1^n$. □

## 5. Proof of Theorem 1.3

First, let us calculate the sums specified in the following lemma, which we will need in the proofs of Theorem 1.3 and Theorem 1.4.

**Lemma 5.1.** *Let $p > 3$ be a prime number and for nonzero $m \in \mathbb{F}_p$,*

$$\widehat{S}_p(m) = \frac{1}{2p} + \frac{1}{2\sqrt{p}} \varepsilon \left(\frac{-m}{p}\right),$$

*where*

$$\varepsilon = \begin{cases} 1 & \text{if } p \equiv 1 \pmod{4}, \\ i & \text{if } p \equiv 3 \pmod{4}. \end{cases}$$

*Then, for $a, b, c, d \in \mathbb{F}_p$ with $abcd(ad - bc) \neq 0$, we have*

***(I)*** $\displaystyle\sum_{m\neq 0}\widehat{S}_p(m)\widehat{S}_p(am)\widehat{S}_p(bm)=\left(\frac{1}{8p^3}+\frac{1}{8p^2}\varepsilon^2\left(\left(\frac{a}{p}\right)+\left(\frac{b}{p}\right)+\left(\frac{ab}{p}\right)\right)\right)(p-1).$

***(II)***
$$\sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2=0\\ m_2\neq 0,\ m_1\neq 0}}\widehat{S}_p(m_1)\widehat{S}_p(m_2)\widehat{S}_p(cm_1+dm_2)$$
$$=\left(\frac{1}{8p^3}+\frac{1}{8p^2}\varepsilon^2\left(\left(\frac{-ab}{p}\right)+\left(\frac{b(bc-ad)}{p}\right)+\left(\frac{-a(bc-ad)}{p}\right)\right)\right)(p-1)\,.$$

***(III)***
$$\sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}}\widehat{S}_p(m_1)\widehat{S}_p(m_2)\widehat{S}_p(am_1+bm_2)\widehat{S}_p(cm_1+dm_2)$$
$$\begin{aligned}
&=\frac{(p-1)(p-3)}{16p^4}+\frac{1}{16p^2}(p-1)\left(\frac{-1}{p}\right)(\#E(\mathbb{F}_p)-p-1)\\
&-\frac{1}{16p^3}\varepsilon^2\left(\left(\frac{a}{p}\right)+\left(\frac{b}{p}\right)+\left(\frac{c}{p}\right)+\left(\frac{d}{p}\right)\right)(p-1)\\
&-\frac{1}{16p^3}\varepsilon^2\left(\left(\frac{a\cdot(ad-bc)}{p}\right)+\left(\frac{d\cdot(ad-bc)}{p}\right)\right)(p-1)\\
&-\frac{1}{16p^3}\varepsilon^2\left(\left(\frac{-b\cdot(ad-bc)}{p}\right)+\left(\frac{-c\cdot(ad-bc)}{p}\right)\right)(p-1)\\
&-\frac{1}{16p^3}\varepsilon^2\left(\left(\frac{-ab}{p}\right)+\left(\frac{-cd}{p}\right)+\left(\frac{ac}{p}\right)+\left(\frac{bd}{p}\right)\right)(p-1)
\end{aligned}$$

*where the elliptic curve $E$ over $\mathbb{F}_p$ is defined by*

$$E:y^2=x(x-bc)(x-ad).$$

**Proof.** ***(I).*** As $ab\neq 0$, we see that

$$\begin{aligned}
&\sum_{m\neq 0}\widehat{S}_p(m)\widehat{S}_p(am)\widehat{S}_p(bm)\\
&\quad=\sum_{m\neq 0}\left(\frac{1}{2p}+\frac{1}{2\sqrt{p}}\varepsilon\left(\frac{-m}{p}\right)\right)\left(\frac{1}{2p}+\frac{1}{2\sqrt{p}}\varepsilon\left(\frac{-am}{p}\right)\right)\left(\frac{1}{2p}+\frac{1}{2\sqrt{p}}\varepsilon\left(\frac{-bm}{p}\right)\right).
\end{aligned}$$

Since we know that the sum of product of terms containing an odd number of Legendre symbols is zero, we just need to calculate the product of terms containing an even number of Legendre symbols. Thus, we obtain that

$$\begin{aligned}
\sum_{m\neq 0}\widehat{S}_p(m)\widehat{S}_p(am)\widehat{S}_p(bm)&=\sum_{m\neq 0}\left(\frac{1}{8p^3}+\frac{1}{8p^2}\varepsilon^2\left(\left(\frac{ab}{p}\right)+\left(\frac{a}{p}\right)+\left(\frac{b}{p}\right)\right)\right)\\
&=\left(\frac{1}{8p^3}+\frac{1}{8p^2}\varepsilon^2\left(\left(\frac{ab}{p}\right)+\left(\frac{a}{p}\right)+\left(\frac{b}{p}\right)\right)\right)(p-1).
\end{aligned}$$

***(II).*** Since $am_1 + bm_2 = 0$, we have

$$\sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2=0\\ m_2\neq 0,\ m_1\neq 0}} \widehat{S}_p(m_1)\widehat{S}_p(m_2)\widehat{S}_p(cm_1+dm_2) = \sum_{m_1\neq 0} \widehat{S}_p(m_1)\widehat{S}_p(-b^{-1}am_1)\widehat{S}_p((c-db^{-1}a)m_1).$$

Then, by *(I)* the desired equality is obtained.

***(III).*** Now, let us calculate the last sum:

$$\sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \widehat{S}_p(m_1)\widehat{S}_p(m_2)\widehat{S}_p(am_1+bm_2)\widehat{S}_p(cm_1+dm_2) = T(1)+T(2)+T(3)+T(4),$$

where

$$T(1) = \sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \frac{1}{16p^3}\varepsilon^2\left(\frac{-m_1}{p}\right)\left(\frac{-cm_1-dm_2}{p}\right) \tag{27}$$

$$+ \sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \frac{1}{16p^3}\varepsilon^2\left(\frac{-m_1}{p}\right)\left(\frac{-am_1-bm_2}{p}\right) \tag{28}$$

$$+ \sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \frac{1}{16p^3}\varepsilon^2\left(\frac{-m_2}{p}\right)\left(\frac{-cm_1-dm_2}{p}\right) \tag{29}$$

$$+ \sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \frac{1}{16p^3}\varepsilon^2\left(\frac{-m_2}{p}\right)\left(\frac{-am_1-bm_2}{p}\right) \tag{30}$$

$$+ \sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \frac{1}{16p^3}\varepsilon^2\left(\frac{-am_1-bm_2}{p}\right)\left(\frac{-cm_1-dm_2}{p}\right) \tag{31}$$

$$+ \sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \frac{1}{16p^3}\varepsilon^2\left(\frac{-m_1}{p}\right)\left(\frac{-m_2}{p}\right), \tag{32}$$

$$T(2) = \sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \frac{1}{16p^4} \tag{33}$$

$$+ \sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \frac{1}{16p^2}\left(\frac{-m_1}{p}\right)\left(\frac{-m_2}{p}\right)\left(\frac{-am_1-bm_2}{p}\right)\left(\frac{-cm_1-dm_2}{p}\right), \tag{34}$$

$$T(3) = \sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \frac{1}{16p^{7/2}}\varepsilon\left(\frac{-m_1}{p}\right) \tag{35}$$

$$+ \sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \frac{1}{16p^{7/2}}\varepsilon\left(\frac{-m_2}{p}\right) \tag{36}$$

$$+ \sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \frac{1}{16p^{7/2}}\varepsilon\left(\frac{-cm_1-dm_2}{p}\right) \tag{37}$$

$$+ \sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \frac{1}{16p^{7/2}}\varepsilon\left(\frac{-am_1-bm_2}{p}\right), \tag{38}$$

$$T(4) = \sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \frac{1}{16p^{5/2}}\varepsilon^3\left(\frac{-m_1}{p}\right)\left(\frac{-m_2}{p}\right)\left(\frac{-cm_1-dm_2}{p}\right) \tag{39}$$

$$+ \sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \frac{1}{16p^{5/2}}\varepsilon^3\left(\frac{-m_1}{p}\right)\left(\frac{-m_2}{p}\right)\left(\frac{-am_1-bm_2}{p}\right) \tag{40}$$

$$+ \sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \frac{1}{16p^{5/2}}\varepsilon^3\left(\frac{-m_1}{p}\right)\left(\frac{-am_1-bm_2}{p}\right)\left(\frac{-cm_1-dm_2}{p}\right) \tag{41}$$

$$+ \sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \frac{1}{16p^{5/2}}\varepsilon^3\left(\frac{-m_2}{p}\right)\left(\frac{-am_1-bm_2}{p}\right)\left(\frac{-cm_1-dm_2}{p}\right). \tag{42}$$

We start by calculating (27). First, we edit the index of the sum by the inclusion-exclusion principle:

$$\sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \frac{1}{16p^3}\varepsilon^2\left(\frac{-m_1}{p}\right)\left(\frac{-cm_1-dm_2}{p}\right)$$

$$= \sum_{m_2\neq 0, m_1\neq 0} \frac{1}{16p^3}\varepsilon^2\left(\frac{-m_1}{p}\right)\left(\frac{-cm_1-dm_2}{p}\right)$$

$$- \sum_{\substack{cm_1+dm_2=0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \frac{1}{16p^3}\varepsilon^2\left(\frac{-m_1}{p}\right)\left(\frac{-cm_1-dm_2}{p}\right)$$

$$-\sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2=0\\ m_2\neq 0,\ m_1\neq 0}} \frac{1}{16p^3}\varepsilon^2\left(\frac{-m_1}{p}\right)\left(\frac{-cm_1-dm_2}{p}\right). \tag{43}$$

Then, by Proposition 2.8 or by a change of variable, we obtain that

$$\sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \frac{1}{16p^3}\varepsilon^2\left(\frac{-m_1}{p}\right)\left(\frac{-cm_1-dm_2}{p}\right)$$
$$= -\frac{1}{16p^3}\varepsilon^2\left(\left(\frac{c}{p}\right)+\left(\frac{-b\cdot(ad-bc)}{p}\right)\right)(p-1). \tag{44}$$

By arranging the coefficients in equation (44), we calculate (28), (29) and (30), respectively:

$$\sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \frac{1}{16p^3}\varepsilon^2\left(\frac{-m_1}{p}\right)\left(\frac{-am_1-bm_2}{p}\right)$$
$$= -\frac{1}{16p^3}\varepsilon^2\left(\left(\frac{a}{p}\right)+\left(\frac{d\cdot(ad-bc)}{p}\right)\right)(p-1),$$
$$\sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \frac{1}{16p^3}\varepsilon^2\left(\frac{-m_2}{p}\right)\left(\frac{-cm_1-dm_2}{p}\right)$$
$$= -\frac{1}{16p^3}\varepsilon^2\left(\left(\frac{d}{p}\right)+\left(\frac{a\cdot(ad-bc)}{p}\right)\right)(p-1),$$
$$\sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \frac{1}{16p^3}\varepsilon^2\left(\frac{-m_2}{p}\right)\left(\frac{-am_1-bm_2}{p}\right)$$
$$= -\frac{1}{16p^3}\varepsilon^2\left(\left(\frac{b}{p}\right)+\left(\frac{-c\cdot(ad-bc)}{p}\right)\right)(p-1).$$

If the same method in (43) is applied for (31) and (32), then we deduce that

$$\sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \frac{1}{16p^3}\varepsilon^2\left(\frac{-am_1-bm_2}{p}\right)\left(\frac{-cm_1-dm_2}{p}\right)$$
$$= -\frac{1}{16p^3}\varepsilon^2\left(\left(\frac{ac}{p}\right)+\left(\frac{bd}{p}\right)\right)(p-1),$$
$$\sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \frac{1}{16p^3}\varepsilon^2\left(\frac{-m_1}{p}\right)\left(\frac{-m_2}{p}\right) = -\frac{1}{16p^3}\varepsilon^2\left(\left(\frac{-ab}{p}\right)+\left(\frac{-cd}{p}\right)\right)(p-1).$$

Also, it is quickly obtained that

$$\sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \frac{1}{16p^4} = \frac{(p-1)(p-3)}{16p^4}.$$

Now, we compute (34) with the help of Theorem 2.11. Observe that

$$\begin{aligned}&\sum_{\substack{cm_1+dm_2\neq 0\\ am_1+bm_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \frac{1}{16p^2}\left(\frac{-m_1}{p}\right)\left(\frac{-m_2}{p}\right)\left(\frac{-am_1-bm_2}{p}\right)\left(\frac{-cm_1-dm_2}{p}\right)\\ &= \sum_{m_2\neq 0, m_1\neq 0} \frac{1}{16p^2}\left(\frac{-m_1}{p}\right)\left(\frac{-m_2}{p}\right)\left(\frac{-am_1-bm_2}{p}\right)\left(\frac{-cm_1-dm_2}{p}\right). \qquad (45)\end{aligned}$$

First, let us make (45) convenient to use Theorem 2.11. The change of variable

$$m_2 = m\cdot m_1$$

is bijective, so we have

$$\begin{aligned}&\frac{1}{16p^2}\sum_{m_1\neq 0, m\neq 0}\left(\frac{-m_1}{p}\right)\left(\frac{-m\cdot m_1}{p}\right)\left(\frac{-am_1-bm\cdot m_1}{p}\right)\left(\frac{-cm_1-dm\cdot m_1}{p}\right)\\ &= \frac{1}{16p^2}(p-1)\left(\frac{-1}{p}\right)\sum_{m\neq 0}\left(\frac{-m}{p}\right)\left(\frac{-bm-a}{p}\right)\left(\frac{-dm-c}{p}\right).\end{aligned}$$

Next, we deal with the sum

$$\sum_{m\neq 0}\left(\frac{-m}{p}\right)\left(\frac{-bm-a}{p}\right)\left(\frac{-dm-c}{p}\right). \qquad (46)$$

Consider the curve $y^2 = -x(-bx-a)(-dx-c) = -bdx^3-(bc+ad)x^2-acx$, which can be rewritten as $(-bdy)^2 = (-bdx)^3-(bc+ad)(-bdx)^2+acbd(-bdx)$. By replacing $-bdx$ with $x$ and $-bdy$ with $y$, we arrive at the elliptic curve

$$E: y^2 = x^3-(bc+ad)x^2+acbdx = x(x-bc)(x-ad). \qquad (47)$$

Hence, by Theorem 2.11, we obtain that

$$\sum_{m\in\mathbb{F}_p}\left(\frac{-m(bm+a)(dm+c)}{p}\right) = \#E(\mathbb{F}_p)-p-1 \qquad (48)$$

where the elliptic curve $E$ over $\mathbb{F}_p$ is defined as above (47).

By Remark 2.9, we compute that each sum in $T(3)$ and $T(4)$ is zero. Thus, we conclude that

$$\begin{aligned}
T(1)+T(2)+T(3)+T(4) = &-\frac{1}{16p^3}\varepsilon^2\left(\left(\frac{c}{p}\right)+\left(\frac{-b\cdot(ad-bc)}{p}\right)\right)(p-1)\\
&-\frac{1}{16p^3}\varepsilon^2\left(\left(\frac{a}{p}\right)+\left(\frac{d\cdot(ad-bc)}{p}\right)\right)(p-1)\\
&-\frac{1}{16p^3}\varepsilon^2\left(\left(\frac{d}{p}\right)+\left(\frac{a\cdot(ad-bc)}{p}\right)\right)(p-1)\\
&-\frac{1}{16p^3}\varepsilon^2\left(\left(\frac{b}{p}\right)+\left(\frac{-c\cdot(ad-bc)}{p}\right)\right)(p-1)\\
&-\frac{1}{16p^3}\varepsilon^2\left(\left(\frac{ac}{p}\right)+\left(\frac{bd}{p}\right)\right)(p-1)\\
&-\frac{1}{16p^3}\varepsilon^2\left(\left(\frac{-ab}{p}\right)+\left(\frac{-cd}{p}\right)\right)(p-1)\\
&+\frac{(p-1)(p-3)}{16p^4}\\
&+\frac{1}{16p^2}(p-1)\left(\frac{-1}{p}\right)(\#E(\mathbb{F}_p)-p-1). \quad \square
\end{aligned}$$

Now, we are ready to prove Theorem 1.3.

**Proof of Theorem 1.3.** Let $p>3$ be a prime number. Let

$$Q_p(t)=|\{(x,y,z,v)\in S_p^4 \mid y-x=z-y=v-z=t\}| \tag{49}$$

denote the number of 4-term arithmetic progressions in $S_p$ with common difference $t$.

By Lemma 2.4, the number of 4-APs in $S_p$ is equal to

$$\sum_{t\in\mathbb{F}_p}Q_p(t)=\frac{|S_p|^4}{p^2}+p^2\sum_{(m_1,m_2)\neq(0,0)}\widehat{S}_p(m_1)\widehat{S}_p(m_2)\widehat{S}_p(-2m_1-3m_2)\widehat{S}_p(m_1+2m_2). \tag{50}$$

Note that the following system of equations

$$\begin{aligned}-2x_1-3x_2&=0\\ x_1+2x_2&=0\end{aligned}$$

has a unique solution since we have

$$\begin{vmatrix}-2 & -3\\ 1 & 2\end{vmatrix}=-1,$$

which is invertible in $\mathbb{F}_p$. Denote the last splitted term by

$$\begin{aligned} H &= p^2 \sum_{(m_1,m_2)\neq(0,0)} \widehat{S}_p(m_1)\widehat{S}_p(m_2)\widehat{S}_p(-2m_1-3m_2)\widehat{S}_p(m_1+2m_2) \\ &= p|S_p| \sum_{m_2\neq 0,\ m_1=0} \widehat{S}_p(m_2)\widehat{S}_p(-3m_2)\widehat{S}_p(2m_2) \\ &+ p|S_p| \sum_{m_1\neq 0,\ m_2=0} \widehat{S}_p(m_1)\widehat{S}_p(-2m_1)\widehat{S}_p(m_1) \\ &+ p|S_p| \sum_{\substack{m_1+2m_2\neq 0 \\ -2m_1-3m_2=0 \\ m_2\neq 0,\ m_1\neq 0}} \widehat{S}_p(m_1)\widehat{S}_p(m_2)\widehat{S}_p(m_1+2m_2) \\ &+ p|S_p| \sum_{\substack{-2m_1-3m_2\neq 0 \\ m_1+2m_2=0 \\ m_2\neq 0,\ m_1\neq 0}} \widehat{S}_p(m_1)\widehat{S}_p(m_2)\widehat{S}_p(-2m_1-3m_2) \\ &+ p^2 \sum_{\substack{-2m_1-3m_2\neq 0 \\ m_1+2m_2\neq 0 \\ m_2\neq 0,\ m_1\neq 0}} \widehat{S}_p(m_1)\widehat{S}_p(m_2)\widehat{S}_p(-2m_1-3m_2)\widehat{S}_p(m_1+2m_2). \end{aligned}$$

By Lemma 5.1, we compute the five sums mentioned above respectively:

$$\begin{aligned} T_1 &= \sum_{m_2\neq 0,\ m_1=0} \widehat{S}_p(m_2)\widehat{S}_p(-3m_2)\widehat{S}_p(2m_2) \\ &= \left(\frac{1}{8p^3} + \frac{1}{8p^2}\varepsilon^2\left(\left(\frac{-6}{p}\right) + \left(\frac{2}{p}\right) + \left(\frac{-3}{p}\right)\right)\right)(p-1), \\ T_2 &= \sum_{m_1\neq 0,\ m_2=0} \widehat{S}_p(m_1)\widehat{S}_p(-2m_1)\widehat{S}_p(m_1) \\ &= \left(\frac{1}{8p^3} + \frac{1}{8p^2}\varepsilon^2\left(\left(\frac{-2}{p}\right) + \left(\frac{1}{p}\right) + \left(\frac{-2}{p}\right)\right)\right)(p-1)\,, \\ T_3 &= \sum_{\substack{m_1+2m_2\neq 0 \\ -2m_1-3m_2=0 \\ m_2\neq 0,\ m_1\neq 0}} \widehat{S}_p(m_1)\widehat{S}_p(m_2)\widehat{S}_p(m_1+2m_2) \\ &= \left(\frac{1}{8p^3} + \frac{1}{8p^2}\varepsilon^2\left(\left(\frac{-6}{p}\right) + \left(\frac{-3}{p}\right) + \left(\frac{2}{p}\right)\right)\right)(p-1)\,. \\ T_4 &= \sum_{\substack{-2m_1-3m_2\neq 0 \\ m_1+2m_2=0 \\ m_2\neq 0,\ m_1\neq 0}} \widehat{S}_p(m_1)\widehat{S}_p(m_2)\widehat{S}_p(-2m_1-3m_2) \\ &= \left(\frac{1}{8p^3} + \frac{1}{8p^2}\varepsilon^2\left(\left(\frac{-2}{p}\right) + \left(\frac{-2}{p}\right) + \left(\frac{1}{p}\right)\right)\right)(p-1)\,. \end{aligned}$$

Now, let us calculate the last sum by $(III)$ of Lemma 5.1:

$$\begin{aligned}
T_5 &= \sum_{\substack{-2m_1-3m_2\neq 0\\ m_1+2m_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \widehat{S}_p(m_1)\widehat{S}_p(m_2)\widehat{S}_p(-2m_1-3m_2)\widehat{S}_p(m_1+2m_2)\\
&= \frac{(p-1)(p-3)}{16p^4} + \frac{1}{16p^2}(p-1)\left(\frac{-1}{p}\right)(\#E(\mathbb{F}_p)-p-1)\\
&- \frac{1}{16p^3}\varepsilon^2\left(\left(\frac{1}{p}\right)+\left(\frac{2}{p}\right)+\left(\frac{-2}{p}\right)+\left(\frac{-3}{p}\right)\right)(p-1)\\
&- \frac{1}{16p^3}\varepsilon^2\left(\left(\frac{1}{p}\right)+\left(\frac{-3}{p}\right)\right)(p-1)\\
&- \frac{1}{16p^3}\varepsilon^2\left(\left(\frac{-2}{p}\right)+\left(\frac{2}{p}\right)\right)(p-1)\\
&- \frac{1}{16p^3}\varepsilon^2\left(\left(\frac{-2}{p}\right)+\left(\frac{-6}{p}\right)+\left(\frac{-2}{p}\right)+\left(\frac{-6}{p}\right)\right)(p-1)
\end{aligned}$$

where the elliptic curve $E$ over $\mathbb{F}_p$ is defined by

$$E : y^2 = x(x+4)(x+3).$$

Thus, we conclude that the number of non-trivial 4-APs in $S_p$ is given by the following formula:

$$\begin{aligned}
\sum_{t\in\mathbb{F}_p} Q_p(t) - |S_p| = \frac{|S_p|^4}{p^2}&\\
&+ p|S_p|\left(\frac{1}{8p^3}+\frac{1}{8p^2}\varepsilon^2\left(\left(\frac{-6}{p}\right)+\left(\frac{2}{p}\right)+\left(\frac{-3}{p}\right)\right)\right)(p-1)\\
&+ p|S_p|\left(\frac{1}{8p^3}+\frac{1}{8p^2}\varepsilon^2\left(\left(\frac{-2}{p}\right)+\left(\frac{1}{p}\right)+\left(\frac{-2}{p}\right)\right)\right)(p-1)\\
&+ p|S_p|\left(\frac{1}{8p^3}+\frac{1}{8p^2}\varepsilon^2\left(\left(\frac{-6}{p}\right)+\left(\frac{-3}{p}\right)+\left(\frac{2}{p}\right)\right)\right)(p-1)\\
&+ p|S_p|\left(\frac{1}{8p^3}+\frac{1}{8p^2}\varepsilon^2\left(\left(\frac{-2}{p}\right)+\left(\frac{-2}{p}\right)+\left(\frac{1}{p}\right)\right)\right)(p-1)\\
&+ \frac{(p-1)(p-3)}{16p^2} + \frac{1}{16}(p-1)\left(\frac{-1}{p}\right)(\#E(\mathbb{F}_p)-p-1)\\
&- \frac{1}{16p}\varepsilon^2\left(\left(\frac{1}{p}\right)+\left(\frac{2}{p}\right)+\left(\frac{-2}{p}\right)+\left(\frac{-3}{p}\right)\right)(p-1)\\
&- \frac{1}{16p}\varepsilon^2\left(\left(\frac{1}{p}\right)+\left(\frac{-3}{p}\right)\right)(p-1)
\end{aligned}$$

$$
\begin{aligned}
&- \frac{1}{16p}\varepsilon^2 \left( \left(\frac{-2}{p}\right) + \left(\frac{2}{p}\right) \right)(p-1) \\
&- \frac{1}{16p}\varepsilon^2 \left( \left(\frac{-2}{p}\right) + \left(\frac{-6}{p}\right) + \left(\frac{-2}{p}\right) + \left(\frac{-6}{p}\right) \right)(p-1) \\
&- |S_p|.
\end{aligned}
$$

When the above equation is rearranged, the desired formula is deduced.

Next, we prove the second part of the theorem by applying the Sato-Tate conjecture, which is a theorem now. Recall by Hasse's theorem that

$$|p + 1 - \#E(\mathbb{F}_p)| \leq 2\sqrt{p}.$$

Write

$$2\cos\theta_p = \frac{\#E(\mathbb{F}_p) - p - 1}{\sqrt{p}},$$

where $\theta_p \in [0, \pi]$. By SageMath [36], our elliptic curve

$$E : y^2 = x(x+3)(x+4)$$

has no complex multiplication and its $j$-invariant is $35152/9 \notin \mathbb{Z}$. Then by the Sato-Tate conjecture [3,24], we know that

$$\lim_{N\to\infty} \frac{|\{p \leq N : 0 \leq \theta_p \leq \alpha\}|}{|\{p \leq N\}|} = \frac{2}{\pi}\int_0^{\alpha} \sin^2\theta \; d\theta = \frac{1}{\pi}(\alpha - \sin\alpha \cdot \cos\alpha)$$

for any $\alpha \in [0, \pi]$. Let $\varepsilon > 0$ be given. Then, choosing $\alpha \in (0, \pi]$ sufficiently small with respect to $\varepsilon$ and assembling the first part of the theorem, the Hasse bound and the previous consequence of the Sato-Tate conjecture, the number of non-trivial 4-APs in $S_p$ is given by

$$\frac{p^2}{16} + R_p,$$

where

$$\left(\frac{1}{8} - \varepsilon\right)p^{\frac{3}{2}} \leq |R_p| \leq \left(\frac{1}{8} + \varepsilon\right)p^{\frac{3}{2}}$$

holds for infinitely many prime numbers $p$. Hence, the error term $O(p^{\frac{3}{2}})$ and the constant $\frac{1}{8}$ are both best possible. □

## 6. Proof of Theorem 1.4

We will make use of the following two lemmas in the proof Theorem 1.4.

**Lemma 6.1.** *Let $p > 3$ be a prime number and for nonzero $m \in \mathbb{F}_p$,*

$$\widehat{S}_p(m) = \frac{1}{2p} + \frac{1}{2\sqrt{p}}\varepsilon\left(\frac{-m}{p}\right),$$

*where*

$$\varepsilon = \begin{cases} 1 & \text{if } p \equiv 1 \pmod 4, \\ i & \text{if } p \equiv 3 \pmod 4. \end{cases}$$

*Then, for $a, b, c, d, e \in \mathbb{F}_p$ with*

$$abcd(ad - bc) \neq 0 \text{ and } e(a - cd)(b - ce) \neq 0,$$

*respectively, we have*

$$\begin{aligned}
&\boldsymbol{(I')} \sum_{m_1 \neq 0,\ m_2 \neq 0} \widehat{S}_p(m_1)\widehat{S}_p(m_2)\widehat{S}_p(am_1 + bm_2)\widehat{S}_p(cm_1 + dm_2) \\
&= \left(\frac{1}{8p^3} + \frac{1}{8p^2}\varepsilon^2\left(\left(\frac{-ab}{p}\right) + \left(\frac{b(bc - ad)}{p}\right) + \left(\frac{-a(bc - ad)}{p}\right)\right)\right)\frac{p^2 - 1}{2p} \\
&+ \left(\frac{1}{8p^3} + \frac{1}{8p^2}\varepsilon^2\left(\left(\frac{-cd}{p}\right) + \left(\frac{-d(bc - ad)}{p}\right) + \left(\frac{c(bc - ad)}{p}\right)\right)\right)\frac{p^2 - 1}{2p} \\
&+ \frac{(p - 1)(p - 3)}{16p^4} + \frac{1}{16p^2}(p - 1)\left(\frac{-1}{p}\right)(\#E(\mathbb{F}_p) - p - 1) \\
&- \frac{1}{16p^3}\varepsilon^2\left(\left(\frac{a}{p}\right) + \left(\frac{b}{p}\right) + \left(\frac{c}{p}\right) + \left(\frac{d}{p}\right)\right)(p - 1) \\
&- \frac{1}{16p^3}\varepsilon^2\left(\left(\frac{a \cdot (ad - bc)}{p}\right) + \left(\frac{d \cdot (ad - bc)}{p}\right)\right)(p - 1) \\
&- \frac{1}{16p^3}\varepsilon^2\left(\left(\frac{-b \cdot (ad - bc)}{p}\right) + \left(\frac{-c \cdot (ad - bc)}{p}\right)\right)(p - 1) \\
&- \frac{1}{16p^3}\varepsilon^2\left(\left(\frac{-ab}{p}\right) + \left(\frac{-cd}{p}\right) + \left(\frac{ac}{p}\right) + \left(\frac{bd}{p}\right)\right)(p - 1)
\end{aligned}$$

*where the elliptic curve $E$ over $\mathbb{F}_p$ is defined by*

$$E : y^2 = x(x - bc)(x - ad).$$

$$\textbf{(}\boldsymbol{II'}\textbf{)} \quad \sum_{\substack{am_1+bm_2+cm_3\neq 0\\ dm_1+em_2+m_3=0\\ m_1\neq 0, m_2\neq 0, m_3\neq 0}} \widehat{S}_p(m_1)\widehat{S}_p(m_2)\widehat{S}_p(m_3)\widehat{S}_p(am_1+bm_2+cm_3)$$

$$= \sum_{\substack{(a-cd)m_1+(b-ce)m_2\neq 0\\ -dm_1-em_2\neq 0\\ m_2\neq 0,\ m_1\neq 0}} \widehat{S}_p(m_1)\widehat{S}_p(m_2)\widehat{S}_p(-dm_1-em_2)\widehat{S}_p((a-cd)m_1+(b-ce)m_2).$$

**Proof.** ***(I′).*** Combining $(II)$ and $(III)$ in Lemma 5.1, we obtain $(I')$.

***(II′).*** Using $dm_1+em_2+m_3=0$, we get $(II')$. □

**Lemma 6.2.** ***(I″)*** *For $a,b,c,\beta,\gamma\in\mathbb{F}_p$ with $abc\beta\gamma(c-a\gamma)\neq 0$, we have*

$$\sum_{\substack{am_1+bm_2+cm_3\neq 0\\ m_1+\beta m_2+\gamma m_3\neq 0\\ m_1\neq 0, m_2\neq 0, m_3\neq 0}} \left(\frac{-m_1}{p}\right)\left(\frac{-m_2}{p}\right)$$

$$= \left(\left(\frac{-ab}{p}\right)+\left(\frac{-\beta}{p}\right)+\left(\frac{-(\beta c-\gamma b)(c-a\gamma)}{p}\right)\right)(p-1).$$

***(II″)*** *For $a,b,c,\alpha,\beta,\gamma\in\mathbb{F}_p$ with $abc\alpha\beta\gamma\neq 0$, we have*

$$\sum_{\substack{\alpha m_1+\beta m_2+\gamma m_3\neq 0\\ m_1\neq 0, m_2\neq 0, m_3\neq 0}} \left(\frac{-m_2}{p}\right)\left(\frac{am_1+bm_2+cm_3}{p}\right)$$

$$= \left(\left(\frac{-b}{p}\right)+\left(\frac{-\alpha(b\alpha-a\beta)}{p}\right)+\left(\frac{-\gamma(b\gamma-c\beta)}{p}\right)\right)(p-1)\,.$$

**Proof.** ***(I″)*** By the inclusion-exclusion principle and the properties of the Legendre symbol,

$$\sum_{\substack{am_1+bm_2+cm_3\neq 0\\ m_1\neq 0, m_2\neq 0, m_3\neq 0}} \left(\frac{-m_1}{p}\right)\left(\frac{-m_2}{p}\right) - \sum_{\substack{am_1+bm_2+cm_3\neq 0\\ m_1+\beta m_2+\gamma m_3=0\\ m_1\neq 0, m_2\neq 0, m_3\neq 0}} \left(\frac{-m_1}{p}\right)\left(\frac{-m_2}{p}\right) \tag{51}$$

$$= \sum_{m_1\neq 0, m_2\neq 0, m_3\neq 0} \left(\frac{-m_1}{p}\right)\left(\frac{-m_2}{p}\right) - \sum_{\substack{am_1+bm_2+cm_3=0\\ m_1\neq 0, m_2\neq 0, m_3\neq 0}} \left(\frac{-m_1}{p}\right)\left(\frac{-m_2}{p}\right)$$

$$- \sum_{\substack{(b-a\beta)m_2+(c-a\gamma)m_3\neq 0\\ m_2\neq 0, m_3\neq 0}} \left(\frac{\beta m_2+\gamma m_3}{p}\right)\left(\frac{-m_2}{p}\right)$$

$$= -\sum_{m_2\neq 0, m_3\neq 0} \left(\frac{a^{-1}bm_2+a^{-1}cm_3}{p}\right)\left(\frac{-m_2}{p}\right) - \sum_{m_2\neq 0, m_3\neq 0} \left(\frac{\beta m_2+\gamma m_3}{p}\right)\left(\frac{-m_2}{p}\right)$$

$$+\sum_{\substack{(b-a\beta)m_2+(c-a\gamma)m_3=0\\ m_2\neq 0, m_3\neq 0}}\left(\frac{\beta m_2+\gamma m_3}{p}\right)\left(\frac{-m_2}{p}\right)$$

$$=-\sum_{m_2\neq 0, m_3\neq 0}\left(\frac{bm_2+cm_3}{p}\right)\left(\frac{-am_2}{p}\right)-\sum_{m_2\neq 0, m_3\neq 0}\left(\frac{\beta m_2+\gamma m_3}{p}\right)\left(\frac{-m_2}{p}\right)$$

$$+\sum_{m_2\neq 0}\left(\frac{(\beta c-\gamma b)m_2}{p}\right)\left(\frac{-(c-a\gamma)m_2}{p}\right).$$

It follows from Proposition 2.8 that

$$(51)=\sum_{m_3\neq 0}\left(\frac{-ab}{p}\right)+\sum_{m_3\neq 0}\left(\frac{-\beta}{p}\right)+\sum_{m_2\neq 0}\left(\frac{-(\beta c-\gamma b)(c-a\gamma)}{p}\right).$$

Then, this yields $(I'')$.

***(II'')*** By the inclusion-exclusion principle,

$$\sum_{\substack{\alpha m_1+\beta m_2+\gamma m_3\neq 0\\ m_1\neq 0, m_2\neq 0, m_3\neq 0}}\left(\frac{-m_2}{p}\right)\left(\frac{am_1+bm_2+cm_3}{p}\right)$$

$$=\sum_{m_1\neq 0, m_2\neq 0, m_3\neq 0}\left(\frac{-m_2}{p}\right)\left(\frac{am_1+bm_2+cm_3}{p}\right)$$

$$-\sum_{\substack{\alpha m_1+\beta m_2+\gamma m_3=0\\ m_1\neq 0, m_2\neq 0, m_3\neq 0}}\left(\frac{-m_2}{p}\right)\left(\frac{am_1+bm_2+cm_3}{p}\right)$$

$$=\sum_{m_2\neq 0, m_3\neq 0}\left(\frac{-m_2}{p}\right)\sum_{m_1}\left(\frac{am_1+bm_2+cm_3}{p}\right)-\sum_{m_2\neq 0, m_3\neq 0}\left(\frac{-m_2}{p}\right)\left(\frac{bm_2+cm_3}{p}\right)$$

$$-\sum_{\substack{-\alpha^{-1}\beta m_2-\alpha^{-1}\gamma m_3\neq 0\\ m_2\neq 0, m_3\neq 0}}\left(\frac{-m_2}{p}\right)\left(\frac{(b-a\alpha^{-1}\beta)m_2+(c-a\alpha^{-1}\gamma)m_3}{p}\right)$$

$$=\sum_{m_2\neq 0, m_3\neq 0}\left(\frac{-m_2}{p}\right)\sum_{m_1}\left(\frac{am_1+bm_2+cm_3}{p}\right)-\sum_{m_2\neq 0, m_3\neq 0}\left(\frac{-m_2}{p}\right)\left(\frac{bm_2+cm_3}{p}\right) \tag{52}$$

$$-\sum_{m_2\neq 0, m_3\neq 0}\left(\frac{-\alpha m_2}{p}\right)\left(\frac{(b\alpha-a\beta)m_2+(c\alpha-a\gamma)m_3}{p}\right)$$

$$+\sum_{m_2\neq 0}\left(\frac{-\alpha m_2}{p}\right)\left(\frac{(b\alpha-a\beta)m_2-(c\alpha-a\gamma)\gamma^{-1}\beta m_2}{p}\right).$$

Then, it follows Proposition 2.7, Proposition 2.8 and change of variables that

$$(52) = \left(\frac{-b}{p}\right)(p-1) + \left(\frac{-\alpha(b\alpha - a\beta)}{p}\right)(p-1) + \left(\frac{-\gamma(b\gamma - c\beta)}{p}\right)(p-1)\,. \quad \square$$

Now, we are ready to prove Theorem 1.4.

**Proof of Theorem 1.4.** Let $p > 3$ be a prime number. Let

$$Q_p(t) = |\{(x_1, x_2, x_3, x_4, x_5) \in S_p^5 \mid x_{i+1} - x_i = t \text{ for } i \in \{1, \dots, 4\}\}| \tag{53}$$

denote the number of 5-term arithmetic progressions in $S_p$ with common difference $t$.

By Lemma 2.4, the number of 5-APs in $S_p$ is equal to

$$\begin{aligned}\sum_{t\in\mathbb{F}_p} Q_p(t) = \frac{|S_p|^5}{p^3} \\ &+ p^2 \sum_{(m_1,m_2,m_3)\neq(0,0,0)} \widehat{S}_p(m_1)\widehat{S}_p(m_2)\widehat{S}_p(m_3)\widehat{S}_p(-2m_1 - 3m_2 - 4m_3) \\ &\cdot \widehat{S}_p(m_1 + 2m_2 + 3m_3).\end{aligned}$$

Denote the last splitted term by

$$\begin{aligned} H = p|S_p| &\sum_{m_1\neq 0, m_2\neq 0} \widehat{S}_p(m_1)\widehat{S}_p(m_2)\widehat{S}_p(-2m_1 - 3m_2)\widehat{S}_p(m_1 + 2m_2) \\ &+ p|S_p| \sum_{m_1\neq 0, m_3\neq 0} \widehat{S}_p(m_1)\widehat{S}_p(m_3)\widehat{S}_p(-2m_1 - 4m_3)\widehat{S}_p(m_1 + 3m_3) \\ &+ p|S_p| \sum_{m_2\neq 0, m_3\neq 0} \widehat{S}_p(m_2)\widehat{S}_p(m_3)\widehat{S}_p(-3m_2 - 4m_3)\widehat{S}_p(2m_2 + 3m_3) \\ &+ |S_p|^2 \sum_{m_3\neq 0} \widehat{S}_p(m_3)\widehat{S}_p(-4m_3)\widehat{S}_p(3m_3) \\ &+ |S_p|^2 \sum_{m_2\neq 0} \widehat{S}_p(m_2)\widehat{S}_p(-3m_2)\widehat{S}_p(2m_2) \\ &+ |S_p|^2 \sum_{m_1\neq 0} \widehat{S}_p(m_1)\widehat{S}_p(-2m_1)\widehat{S}_p(m_1) \\ &+ p|S_p| \sum_{\substack{-2m_1-3m_2-4m_3\neq 0 \\ m_1+2m_2+3m_3=0 \\ m_1\neq 0, m_2\neq 0, m_3\neq 0}} \widehat{S}_p(m_1)\widehat{S}_p(m_2)\widehat{S}_p(m_3)\widehat{S}_p(-2m_1 - 3m_2 - 4m_3) \\ &+ p|S_p| \sum_{\substack{-2m_1-3m_2-4m_3=0 \\ m_1+2m_2+3m_3\neq 0 \\ m_1\neq 0, m_2\neq 0, m_3\neq 0}} \widehat{S}_p(m_1)\widehat{S}_p(m_2)\widehat{S}_p(m_3)\widehat{S}_p(m_1 + 2m_2 + 3m_3) \\ &+ |S_p|^2 \sum_{\substack{-2m_1-3m_2-4m_3=0 \\ m_1+2m_2+3m_3=0 \\ m_1\neq 0, m_2\neq 0, m_3\neq 0}} \widehat{S}_p(m_1)\widehat{S}_p(m_2)\widehat{S}_p(m_3) \end{aligned} \tag{54}$$

$$+ p^2 \sum_{\substack{-2m_1-3m_2-4m_3\neq 0\\ m_1+2m_2+3m_3\neq 0\\ m_1\neq 0, m_2\neq 0, m_3\neq 0}} \widehat{S}_p(m_1)\widehat{S}_p(m_2)\widehat{S}_p(m_3)\widehat{S}_p(-2m_1-3m_2-4m_3)$$
$$\cdot \widehat{S}_p(m_1+2m_2+3m_3).$$

By $(I')$ in Lemma 6.1, we compute the first three sums mentioned above, respectively:

$$\begin{aligned} p|S_p| &\sum_{m_1\neq 0, m_2\neq 0} \widehat{S}_p(m_1)\widehat{S}_p(m_2)\widehat{S}_p(-2m_1-3m_2)\widehat{S}_p(m_1+2m_2) \\ &= \frac{f_1(p)}{32p^3} + \frac{g_1\left(p, \left(\frac{-1}{p}\right), \left(\frac{2}{p}\right), \left(\frac{3}{p}\right)\right)}{32p^2} \\ &+ \frac{1}{32p}\left(p^2-1\right)\left(\frac{-1}{p}\right)\left(\#E_1(\mathbb{F}_p)-p-1\right), \end{aligned} \tag{55}$$

where $f_1$ and $g_1$ are two polynomials of degree 3 with respect to $p$, and the elliptic curve $E_1$ over $\mathbb{F}_p$ is defined by

$$E_1 : y^2 = x^3 + 7x^2 + 12x = x(x+3)(x+4).$$

$$\begin{aligned} p|S_p| &\sum_{m_1\neq 0, m_3\neq 0} \widehat{S}_p(m_1)\widehat{S}_p(m_3)\widehat{S}_p(-2m_1-4m_3)\widehat{S}_p(m_1+3m_3) \\ &= \frac{f_2(p)}{32p^3} + \frac{g_2\left(p, \left(\frac{-1}{p}\right), \left(\frac{2}{p}\right), \left(\frac{3}{p}\right)\right)}{32p^2} \\ &+ \frac{1}{32p}\left(p^2-1\right)\left(\frac{-1}{p}\right)\left(\#E_2(\mathbb{F}_p)-p-1\right), \end{aligned} \tag{56}$$

where $f_2$ and $g_2$ are two polynomials of degree 3 with respect to $p$, and the elliptic curve $E_2$ over $\mathbb{F}_p$ is defined by

$$E_2 : y^2 = x^3 + 10x^2 + 24x = x(x+4)(x+6).$$

$$\begin{aligned} p|S_p| &\sum_{m_2\neq 0, m_3\neq 0} \widehat{S}_p(m_2)\widehat{S}_p(m_3)\widehat{S}_p(-3m_2-4m_3)\widehat{S}_p(2m_2+3m_3) \\ &= \frac{f_3(p)}{32p^3} + \frac{g_3\left(p, \left(\frac{-1}{p}\right), \left(\frac{2}{p}\right), \left(\frac{3}{p}\right)\right)}{32p^2} \\ &+ \frac{1}{32p}\left(p^2-1\right)\left(\frac{-1}{p}\right)\left(\#E_3(\mathbb{F}_p)-p-1\right), \end{aligned} \tag{57}$$

where $f_3$ and $g_3$ are two polynomials of degree 3 with respect to $p$, and the elliptic curve $E_3$ over $\mathbb{F}_p$ is defined by

$$E_3 : y^2 = x^3 + 17x^2 + 72x = x(x+8)(x+9).$$

By $(I)$ in Lemma 5.1, we calculate the other three following sums in (54), respectively.

$$|S_p|^2 \sum_{m_3 \neq 0} \widehat{S}_p(m_3)\widehat{S}_p(-4m_3)\widehat{S}_p(3m_3) = \frac{f_4(p)}{32p^3} + \frac{g_4\left(p, \left(\frac{-1}{p}\right), \left(\frac{2}{p}\right), \left(\frac{3}{p}\right)\right)}{32p^2}, \tag{58}$$

where $f_4$ and $g_4$ are two polynomials of degree 3 with respect to $p$.

$$|S_p|^2 \sum_{m_2 \neq 0} \widehat{S}_p(m_2)\widehat{S}_p(-3m_2)\widehat{S}_p(2m_2) = \frac{f_5(p)}{32p^3} + \frac{g_5\left(p, \left(\frac{-1}{p}\right), \left(\frac{2}{p}\right), \left(\frac{3}{p}\right)\right)}{32p^2}, \tag{59}$$

where $f_5$ and $g_5$ are two polynomials of degree 3 with respect to $p$.

$$|S_p|^2 \sum_{m_1 \neq 0} \widehat{S}_p(m_1)\widehat{S}_p(-2m_1)\widehat{S}_p(m_1) = \frac{f_6(p)}{32p^3} + \frac{g_6\left(p, \left(\frac{-1}{p}\right), \left(\frac{2}{p}\right), \left(\frac{3}{p}\right)\right)}{32p^2}, \tag{60}$$

where $f_6$ and $g_6$ are two polynomials of degree 3 with respect to $p$.

Combining $(II')$ in Lemma 6.1 with $(III)$ in Lemma 5.1, we arrive at

$$\begin{aligned} & p|S_p| \sum_{\substack{-2m_1-3m_2-4m_3 \neq 0 \\ m_1+2m_2+3m_3=0 \\ m_1\neq 0, m_2 \neq 0, m_3 \neq 0}} \widehat{S}_p(m_1)\widehat{S}_p(m_2)\widehat{S}_p(m_3)\widehat{S}_p(-2m_1-3m_2-4m_3) \\ &= \frac{f_1(p)}{32p^3} + \frac{g_1\left(p, \left(\frac{-1}{p}\right), \left(\frac{2}{p}\right), \left(\frac{3}{p}\right)\right)}{32p^2} \\ &+ \frac{1}{32p}\left(p^2-1\right)\left(\frac{-1}{p}\right)\left(\#E_1(\mathbb{F}_p)-p-1\right), \end{aligned} \tag{61}$$

$$\begin{aligned} & p|S_p| \sum_{\substack{-2m_1-3m_2-4m_3 = 0 \\ m_1+2m_2+3m_3\neq 0 \\ m_1\neq 0, m_2 \neq 0, m_3 \neq 0}} \widehat{S}_p(m_1)\widehat{S}_p(m_2)\widehat{S}_p(m_3)\widehat{S}_p(m_1+2m_2+3m_3) \\ &= p|S_p| \sum_{\substack{-3m_2-4m_3 \neq 0 \\ m_2+2m_3\neq 0 \\ m_3\neq 0,\ m_2 \neq 0}} \widehat{S}_p(2m_2)\widehat{S}_p(2m_3)\widehat{S}_p(-3m_2-4m_3)\widehat{S}_p(m_2+2m_3) \\ &= \frac{f_7(p)}{32p^3} + \frac{g_7\left(p, \left(\frac{-1}{p}\right), \left(\frac{2}{p}\right), \left(\frac{3}{p}\right)\right)}{32p^2} \\ &+ \frac{1}{32p}\left(p^2-1\right)\left(\frac{-1}{p}\right)\left(\#E_4(\mathbb{F}_p)-p-1\right), \end{aligned} \tag{62}$$

where $f_7$ and $g_7$ are two polynomials of degree 3 with respect to $p$, and the elliptic curve $E_4$ over $\mathbb{F}_p$ is defined by

$$E_4 : y^2 = x^3 + 40x^2 + 384x = x(x+16)(x+24).$$

Notice that

$$\begin{aligned}\#E_4(\mathbb{F}_p) - p - 1 &= \sum_{x=0}^{p-1} \left(\frac{x(x+16)(x+24)}{p}\right) = \sum_{x=0}^{p-1} \left(\frac{4x(4x+16)(4x+24)}{p}\right) \\ &= \sum_{x=0}^{p-1} \left(\frac{x(x+4)(x+6)}{p}\right) = \#E_2(\mathbb{F}_p) - p - 1.\end{aligned}$$

By (60), we have

$$|S_p|^2 \sum_{\substack{-2m_1-3m_2-4m_3=0 \\ m_1+2m_2+3m_3=0 \\ m_1\neq 0, m_2\neq 0, m_3\neq 0}} \widehat{S}_p(m_1)\widehat{S}_p(m_2)\widehat{S}_p(m_3) = |S_p|^2 \sum_{m\neq 0} \widehat{S}_p(m)\widehat{S}_p(-2m)\widehat{S}_p(m) \tag{63}$$

$$= \frac{f_6(p)}{32p^3} + \frac{g_6\left(p, \left(\frac{-1}{p}\right), \left(\frac{2}{p}\right), \left(\frac{3}{p}\right)\right)}{32p^2}.$$

Now, we calculate the last sum in (54) using Lemma 6.2:

$$p^2 \sum_{\substack{-2m_1-3m_2-4m_3\neq 0 \\ m_1+2m_2+3m_3\neq 0 \\ m_1\neq 0, m_2\neq 0, m_3\neq 0}} \widehat{S}_p(m_1)\widehat{S}_p(m_2)\widehat{S}_p(m_3)\widehat{S}_p(-2m_1-3m_2-4m_3)\widehat{S}_p(m_1+2m_2+3m_3) \tag{64}$$

$$= R(1) + R(2) + R(3),$$

where

$$R(1) = p^2 \sum_{\substack{-2m_1-3m_2-4m_3\neq 0 \\ m_1+2m_2+3m_3\neq 0 \\ m_1\neq 0, m_2\neq 0, m_3\neq 0}} \frac{1}{32p^3} \left(\frac{-m_2}{p}\right)\left(\frac{-m_3}{p}\right)\left(\frac{-m_1-2m_2-3m_3}{p}\right)$$

$$\cdot \left(\frac{2m_1+3m_2+4m_3}{p}\right) \tag{65}$$

$$+ p^2 \sum_{\substack{-2m_1-3m_2-4m_3\neq 0 \\ m_1+2m_2+3m_3\neq 0 \\ m_1\neq 0, m_2\neq 0, m_3\neq 0}} \frac{1}{32p^3} \left(\frac{-m_1}{p}\right)\left(\frac{-m_3}{p}\right)\left(\frac{-m_1-2m_2-3m_3}{p}\right)$$

$$\cdot \left(\frac{2m_1+3m_2+4m_3}{p}\right) \tag{66}$$

$$+ p^2 \sum_{\substack{-2m_1-3m_2-4m_3\neq 0 \\ m_1+2m_2+3m_3\neq 0 \\ m_1\neq 0, m_2\neq 0, m_3\neq 0}} \frac{1}{32p^3} \left(\frac{-m_1}{p}\right)\left(\frac{-m_2}{p}\right)\left(\frac{-m_1-2m_2-3m_3}{p}\right)$$

$$\cdot\left(\frac{2m_1+3m_2+4m_3}{p}\right) \tag{67}$$

$$+\,p^2\sum_{\substack{-2m_1-3m_2-4m_3\neq 0\\ m_1+2m_2+3m_3\neq 0\\ m_1\neq 0,m_2\neq 0,m_3\neq 0}}\frac{1}{32p^3}\left(\frac{-m_1}{p}\right)\left(\frac{-m_2}{p}\right)\left(\frac{-m_3}{p}\right)\left(\frac{2m_1+3m_2+4m_3}{p}\right) \tag{68}$$

$$+\,p^2\sum_{\substack{-2m_1-3m_2-4m_3\neq 0\\ m_1+2m_2+3m_3\neq 0\\ m_1\neq 0,m_2\neq 0,m_3\neq 0}}\frac{1}{32p^3}\left(\frac{-m_1}{p}\right)\left(\frac{-m_2}{p}\right)\left(\frac{-m_3}{p}\right)\left(\frac{-m_1-2m_2-3m_3}{p}\right), \tag{69}$$

and

$$R(2)=p^2\sum_{\substack{-2m_1-3m_2-4m_3\neq 0\\ m_1+2m_2+3m_3\neq 0\\ m_1\neq 0,m_2\neq 0,m_3\neq 0}}\frac{1}{32p^5} \tag{70}$$

$$+\,p^2\sum_{\substack{-2m_1-3m_2-4m_3\neq 0\\ m_1+2m_2+3m_3\neq 0\\ m_1\neq 0,m_2\neq 0,m_3\neq 0}}\frac{1}{32p^4}\varepsilon^2\left(\frac{-m_1-2m_2-3m_3}{p}\right)\left(\frac{2m_1+3m_2+4m_3}{p}\right) \tag{71}$$

$$+\,p^2\sum_{\substack{-2m_1-3m_2-4m_3\neq 0\\ m_1+2m_2+3m_3\neq 0\\ m_1\neq 0,m_2\neq 0,m_3\neq 0}}\frac{1}{32p^4}\varepsilon^2\left(\frac{-m_1}{p}\right)\left(\frac{-m_2}{p}\right) \tag{72}$$

$$+\,p^2\sum_{\substack{-2m_1-3m_2-4m_3\neq 0\\ m_1+2m_2+3m_3\neq 0\\ m_1\neq 0,m_2\neq 0,m_3\neq 0}}\frac{1}{32p^4}\varepsilon^2\left(\frac{-m_1}{p}\right)\left(\frac{-m_3}{p}\right) \tag{73}$$

$$+\,p^2\sum_{\substack{-2m_1-3m_2-4m_3\neq 0\\ m_1+2m_2+3m_3\neq 0\\ m_1\neq 0,m_2\neq 0,m_3\neq 0}}\frac{1}{32p^4}\varepsilon^2\left(\frac{-m_2}{p}\right)\left(\frac{-m_3}{p}\right), \tag{74}$$

and

$$R(3)=p^2\sum_{\substack{-2m_1-3m_2-4m_3\neq 0\\ m_1+2m_2+3m_3\neq 0\\ m_1\neq 0,m_2\neq 0,m_3\neq 0}}\frac{1}{32p^4}\varepsilon^2\left(\frac{-m_3}{p}\right)\left(\frac{2m_1+3m_2+4m_3}{p}\right) \tag{75}$$

$$+\,p^2\sum_{\substack{-2m_1-3m_2-4m_3\neq 0\\ m_1+2m_2+3m_3\neq 0\\ m_1\neq 0,m_2\neq 0,m_3\neq 0}}\frac{1}{32p^4}\varepsilon^2\left(\frac{-m_3}{p}\right)\left(\frac{-m_1-2m_2-3m_3}{p}\right) \tag{76}$$

$$+ p^2 \sum_{\substack{-2m_1-3m_2-4m_3\neq 0\\ m_1+2m_2+3m_3\neq 0\\ m_1\neq 0, m_2\neq 0, m_3\neq 0}} \frac{1}{32p^4}\varepsilon^2\left(\frac{-m_2}{p}\right)\left(\frac{2m_1+3m_2+4m_3}{p}\right) \tag{77}$$

$$+ p^2 \sum_{\substack{-2m_1-3m_2-4m_3\neq 0\\ m_1+2m_2+3m_3\neq 0\\ m_1\neq 0, m_2\neq 0, m_3\neq 0}} \frac{1}{32p^4}\varepsilon^2\left(\frac{-m_2}{p}\right)\left(\frac{-m_1-2m_2-3m_3}{p}\right) \tag{78}$$

$$+ p^2 \sum_{\substack{-2m_1-3m_2-4m_3\neq 0\\ m_1+2m_2+3m_3\neq 0\\ m_1\neq 0, m_2\neq 0, m_3\neq 0}} \frac{1}{32p^4}\varepsilon^2\left(\frac{-m_1}{p}\right)\left(\frac{2m_1+3m_2+4m_3}{p}\right) \tag{79}$$

$$+ p^2 \sum_{\substack{-2m_1-3m_2-4m_3\neq 0\\ m_1+2m_2+3m_3\neq 0\\ m_1\neq 0, m_2\neq 0, m_3\neq 0}} \frac{1}{32p^4}\varepsilon^2\left(\frac{-m_1}{p}\right)\left(\frac{-m_1-2m_2-3m_3}{p}\right). \tag{80}$$

Let us calculate (65). Observe that

$$\frac{1}{32p} \sum_{\substack{-2m_1-3m_2-4m_3\neq 0\\ m_1+2m_2+3m_3\neq 0\\ m_1\neq 0, m_2\neq 0, m_3\neq 0}} \left(\frac{-m_2}{p}\right)\left(\frac{-m_3}{p}\right)\left(\frac{-m_1-2m_2-3m_3}{p}\right)\left(\frac{2m_1+3m_2+4m_3}{p}\right)$$

$$= \frac{1}{32p} \sum_{m_1\neq 0, m_2\neq 0, m_3\neq 0} \left(\frac{m_2m_3}{p}\right)$$

$$\cdot\left(\frac{-2m_1^2-(7m_2+10m_3)\,m_1-(2m_2+3m_3)\,(3m_2+4m_3)}{p}\right). \tag{81}$$

We rewrite (81) in order to use Proposition 2.8. Since

$$b^2-4ac=(7m_2+10m_3)^2-4\cdot 2\cdot(2m_2+3m_3)\cdot(3m_2+4m_3)=(m_2+2m_3)^2,$$

equation (81) becomes

$$\frac{1}{32p} \sum_{\substack{m_2+2m_3=0\\ m_2\neq 0, m_3\neq 0}} \sum_{m_1\neq 0} \left(\frac{m_2m_3}{p}\right)\left(\frac{-m_1-2m_2-3m_3}{p}\right)\left(\frac{2m_1+3m_2+4m_3}{p}\right) \tag{82}$$

$$+ \frac{1}{32p} \sum_{\substack{m_2+2m_3\neq 0\\ m_2\neq 0, m_3\neq 0}} \sum_{m_1\neq 0} \left(\frac{m_2m_3}{p}\right)\left(\frac{-m_1-2m_2-3m_3}{p}\right)\left(\frac{2m_1+3m_2+4m_3}{p}\right). \tag{83}$$

Now, we compute (82) with the help of Proposition 2.8:

$$(82) = \frac{1}{32p} \sum_{\substack{m_2+2m_3=0\\ m_2\neq 0, m_3\neq 0}} \sum_{m_1} \left(\frac{m_2m_3}{p}\right)\left(\frac{-m_1-2m_2-3m_3}{p}\right)\left(\frac{2m_1+3m_2+4m_3}{p}\right)$$

$$
\begin{aligned}
&- \frac{1}{32p} \sum_{\substack{m_2+2m_3=0\\ m_2\neq 0, m_3\neq 0}} \left(\frac{m_2m_3}{p}\right)\left(\frac{-2m_2-3m_3}{p}\right)\left(\frac{3m_2+4m_3}{p}\right)\\
&= \frac{1}{32p} \sum_{\substack{m_2+2m_3=0\\ m_2\neq 0, m_3\neq 0}} (p-1) - \frac{1}{32p} \sum_{\substack{m_2+2m_3=0\\ m_2\neq 0, m_3\neq 0}} 1\\
&= \frac{1}{32p}\left((p-1)^2-(p-1)\right).
\end{aligned}
$$

By Proposition 2.8, the inclusion-exclusion principle and Equation (48), we get that

$$
\begin{aligned}
(83) &= \frac{1}{32p} \sum_{\substack{m_2+2m_3\neq 0\\ m_2\neq 0, m_3\neq 0}} \sum_{m_1} \left(\frac{m_2m_3}{p}\right)\left(\frac{-m_1-2m_2-3m_3}{p}\right)\left(\frac{2m_1+3m_2+4m_3}{p}\right)\\
&- \frac{1}{32p} \sum_{\substack{m_2+2m_3\neq 0\\ m_2\neq 0, m_3\neq 0}} \left(\frac{m_2m_3}{p}\right)\left(\frac{-2m_2-3m_3}{p}\right)\left(\frac{3m_2+4m_3}{p}\right)\\
&= \frac{1}{32p} \sum_{\substack{m_2+2m_3\neq 0\\ m_2\neq 0, m_3\neq 0}} -\left(\frac{-2m_2m_3}{p}\right)\\
&- \frac{1}{32p} \sum_{\substack{m_2+2m_3\neq 0\\ m_2\neq 0, m_3\neq 0}} \left(\frac{m_2m_3}{p}\right)\left(\frac{-2m_2-3m_3}{p}\right)\left(\frac{3m_2+4m_3}{p}\right)\\
&= \frac{1}{32p}\left(2\,(p-1) - \left(\frac{-1}{p}\right)(p-1)\left(\#E_3(\mathbb{F}_p)-p-1\right)\right).
\end{aligned}
$$

Thus, we obtain that

$$
(65) = \frac{1}{32p}\left(p\,(p-1) - \left(\frac{-1}{p}\right)(p-1)\left(\#E_3(\mathbb{F}_p)-p-1\right)\right).
$$

The other terms are quickly determined if the method when calculating (65) is applied. Thus, the following calculations are obtained:

$$
\begin{aligned}
(66) &= \frac{1}{32p}\left(\left(\frac{-6}{p}\right)p\,(p-1) - \left(\frac{-1}{p}\right)(p-1)\left(\#E_2(\mathbb{F}_p)-p-1\right)\right),\\
(67) &= \frac{1}{32p}\left(\left(\frac{6}{p}\right)p\,(p-1) - \left(\frac{-1}{p}\right)(p-1)\left(\#E_1(\mathbb{F}_p)-p-1\right)\right),\\
(68) &= \frac{1}{32p}\left(\left(\frac{6}{p}\right)p\,(p-1) - \left(\frac{-1}{p}\right)(p-1)\left(\#E_1(\mathbb{F}_p)-p-1\right)\right),\\
(69) &= \frac{1}{32p}\left(\left(\frac{-6}{p}\right)p\,(p-1) - \left(\frac{-1}{p}\right)(p-1)\left(\#E_2(\mathbb{F}_p)-p-1\right)\right),
\end{aligned}
$$

$$(70) = \frac{(p-1)(p^2-4p+5)}{32p^3},$$

$$(71) = \frac{1}{32p^2}\left(\left(\frac{-6}{p}\right)+\left(\frac{-3}{p}\right)+\left(\frac{-2}{p}\right)\right)(p-1).$$

Also, applying $(I'')$ of Lemma 6.2 to $R(2)$ and applying $(II'')$ of Lemma 6.2 to $R(3)$, we obtain that

$$(72) = \frac{1}{32p^2}\left(\left(\frac{-6}{p}\right)+\left(\frac{-2}{p}\right)+\left(\frac{-2}{p}\right)\right)(p-1),$$

$$(73) = \frac{1}{32p^2}\left(\left(\frac{-2}{p}\right)+\left(\frac{-3}{p}\right)+1\right)(p-1),$$

$$(74) = \frac{1}{32p^2}\left(\left(\frac{-3}{p}\right)+\left(\frac{-6}{p}\right)+\left(\frac{-2}{p}\right)\right)(p-1)$$

and

$$(75) = \frac{1}{32p^2}\left(\left(\frac{-1}{p}\right)+\left(\frac{2}{p}\right)+\left(\frac{2}{p}\right)\right)(p-1),$$

$$(76) = \frac{1}{32p^2}\left(\left(\frac{3}{p}\right)+1+\left(\frac{3}{p}\right)\right)(p-1),$$

$$(77) = \frac{1}{32p^2}\left(\left(\frac{-3}{p}\right)+1+\left(\frac{-3}{p}\right)\right)(p-1),$$

$$(78) = \frac{1}{32p^2}\left(\left(\frac{2}{p}\right)+\left(\frac{2}{p}\right)+\left(\frac{-1}{p}\right)\right)(p-1),$$

$$(79) = \frac{1}{32p^2}\left(\left(\frac{-2}{p}\right)+\left(\frac{-2}{p}\right)+\left(\frac{-6}{p}\right)\right)(p-1),$$

$$(80) = \frac{1}{32p^2}\left(1+\left(\frac{-3}{p}\right)+\left(\frac{-2}{p}\right)\right)(p-1).$$

Considering the equations from (65) to (80), we have calculated Equation (64). Thus, we obtain an explicit formula as expressed in the theorem.

Next, we prove the second part of the theorem by making use of a version of the Sato-Tate conjecture, and for this we refer the reader to [3,24] and the generalized version of [31, Corollary 2]. By the first part of the theorem and the Hasse bound, observe that the contributions for the error term $O(p^{\frac{3}{2}})$ come from a subsum in Equations (55), (56), (57), (61) and (62), and they are all of the form

$$\frac{1}{32p}(p^2-1)\left(\frac{-1}{p}\right)(\#E_i(\mathbb{F}_p)-p-1)$$

for some $i \in \{1,2,3\}$, as

$$\#E_4(\mathbb{F}_p) - p - 1 = \#E_2(\mathbb{F}_p) - p - 1.$$

Let $V_1 = V_1(a,b)$ and $V_2 = V_2(a,b)$ be the following two elliptic curves

$$y^2 = x^3 + ax^2 + bx$$

and

$$y^2 = x^3 - 2ax^2 + (a^2 - 4b)x$$

respectively, where $a, b \in \mathbb{Z}$ and $b(a^2 - 4b) \neq 0$. Then the map

$$\varphi(x,y) = \left(x + a + \frac{1}{x}, y\left(1 - \frac{b}{x^2}\right)\right)$$

yields an isogeny from $V_1$ to $V_2$ with kernel $\{\mathcal{O}, (0,0)\}$, see [40, p. 110]. Thus, one infers that the elliptic curve $C_1$ defined by $y^2 = x(x+1)(x+4) = x^3 + 5x^2 + 4x$ and the elliptic curve $C_2$ defined by $y^2 = x(x-1)(x-9) = x^3 - 10x^2 + 9x$ are isogenous over $\mathbb{Q}$. Therefore, for any prime number $p > 3$, the elliptic curves $C_1$ and $C_2$ are isogenous over $\mathbb{F}_p$, and hence

$$\#C_1(\mathbb{F}_p) = \#C_2(\mathbb{F}_p)$$

by Tate's isogeny theorem [44].

For the rest of the proof, let $p \equiv 1 \pmod 4$ so that $\left(\frac{-1}{p}\right) = 1$. Next, we will obtain that $E_1(\mathbb{F}_p) - p - 1 = E_3(\mathbb{F}_p) - p - 1$. Now, for any prime $p > 3$

$$\begin{aligned}\#E_1(\mathbb{F}_p) - p - 1 &= \sum_{x=0}^{p-1} \left(\frac{x(x+3)(x+4)}{p}\right) = \sum_{x=0}^{p-1} \left(\frac{-x(-x+3)(-x+4)}{p}\right) \\ &= \sum_{x=0}^{p-1} \left(\frac{x(x-3)(x-4)}{p}\right) = \sum_{x=0}^{p-1} \left(\frac{x(x+1)(x+4)}{p}\right) \\ &= \#C_1(\mathbb{F}_p) - p - 1 = \#C_2(\mathbb{F}_p) - p - 1 \\ &= \sum_{x=0}^{p-1} \left(\frac{x(x-1)(x-9)}{p}\right) = \sum_{x=0}^{p-1} \left(\frac{x(x+8)(x+9)}{p}\right) \\ &= \#E_3(\mathbb{F}_p) - p - 1.\end{aligned}$$

In other words, the corresponding subsums of (55) and (57) coming from the elliptic curves are equal. Recall the elliptic curve

$$E_1 : y^2 = x(x+3)(x+4)$$

has no complex multiplication and its $j$-invariant is $35152/9 \notin \mathbb{Z}$. As in the previous theorem, set

$$2\cos\theta_p = \frac{\#E_1(\mathbb{F}_p) - p - 1}{\sqrt{p}},$$

where $\theta_p \in [0, \pi]$. Then by the proof of the Sato-Tate conjecture [3,24], one immediately gets that

$$\lim_{N\to\infty} \frac{|\{p \le N : 0 \le \theta_p \le \pi/6\}|}{|\{p \le N : p \equiv 1 \pmod 4\}|} = \frac{2}{\varphi(4)\pi} \int\limits_0^{\pi/6} \sin^2\theta \, d\theta = \frac{1}{48}(2\pi - 3\sqrt{3}) \approx 0.022646,$$

and $\cos(\pi/6) = \sqrt{3}/2$. Express the number of 5-APs in $S_p$ as

$$\frac{p^2}{32} + R_p,$$

where $R_p = O(p^{\frac{3}{2}})$. By the consequence of the Sato-Tate conjecture, for infinitely many primes $p \equiv 1 \pmod 4$, Equations (55), (57) and (61) will bring an error term $T_p$ and

$$T_p \ge \frac{1}{32p}\left(p^2 - 1\right) 6\cos(\pi/6)\sqrt{p} = \frac{1}{32p}\left(p^2 - 1\right) 3\sqrt{3}\sqrt{p}. \tag{84}$$

The subsums in (56) and (62), namely

$$\frac{1}{32p}\left(p^2 - 1\right)\left(\frac{-1}{p}\right)\left(\#E_2(\mathbb{F}_p) - p - 1\right)$$

and

$$\frac{1}{32p}\left(p^2 - 1\right)\left(\frac{-1}{p}\right)\left(\#E_4(\mathbb{F}_p) - p - 1\right)$$

can cancel out at most

$$\frac{1}{32p}\left(p^2 - 1\right) 4\sqrt{p}$$

of the term in (84), by Hasse's estimate. Hence, there are two positive absolute constants $c_1$ and $c_2$ such that the inequality

$$c_1 p^{\frac{3}{2}} \le |R_p| \le c_2 p^{\frac{3}{2}}$$

holds for infinitely many prime numbers $p$. This yields that the error term $O(p^{\frac{3}{2}})$ is best possible. □

## 7. Kummer sums and 3-APs

Let $p$ be an odd prime with $p \equiv 1 \pmod 3$. Then, one has that $|C_p| = 1 + \frac{p-1}{3} = \frac{p+2}{3}$ where $C_p = \{t^3 : t \in \mathbb{F}_p\}$. Let

$$K(p) = \sum_{x=0}^{p-1} e_p(x^3)$$

be the Kummer sum. The Kummer sum is related to a cubic Gauss sum which we define next. Let $g$ be a primitive root modulo $p$ and $w = e^{2\pi i/3}$. Define the multiplicative cubic character

$$\chi : \mathbb{F}_p \to \{0, 1, w, w^2\}$$

as follows: $\chi(0) = 0$ and $\chi(g) = w$. Thus, $\chi(g^m) = w^r$, where $r$ is the remainder when $m$ is divided by 3. Note also that $\chi$ extends to $\mathbb{N}$ as a Dirichlet character. Let

$$\tau_p = \sum_{x=0}^{p-1} \chi(x) e_p(x)$$

be the cubic Gauss sum. One can observe that

$$\overline{\tau_p} = \sum_{x=0}^{p-1} \overline{\chi}(x) e_p(x)$$

as $\chi(-1) = 1$. We have that $\tau_p$ has norm $\sqrt{p}$, from [12, Chapter 3]. Thus, $\tau_p = \sqrt{p} e^{i\theta_p}$ and $\overline{\tau_p} = \sqrt{p} e^{-i\theta_p}$ for some angle $\theta_p$. Unlike the quadratic Gauss sum, there is no specific formula for $\theta_p$ and in fact there is an equidistribution result by Heath-Brown and Patterson in [26] as $p$ varies, and this refuted Kummer's guess.

For $x \neq 0$, note that $1+\chi(x)+\overline{\chi}(x) = 3$ if $x$ is in $C_p$ and otherwise $1+\chi(x)+\overline{\chi}(x) = 0$. This yields that

$$K(p) = \sum_{x=0}^{p-1} (1 + \chi(x) + \overline{\chi}(x)) e_p(x) = \tau_p + \overline{\tau_p} = 2\sqrt{p}\cos(\theta_p).$$

Similarly, for $a \neq 0$,

$$\begin{aligned} K(a,p) = \sum_{x=0}^{p-1} e_p(ax^3) &= \sum_{x=0}^{p-1} (1 + \chi(x) + \overline{\chi}(x)) e_p(ax) \\ &= \sum_{x=0}^{p-1} \chi(x) e_p(ax) + \sum_{x=0}^{p-1} \overline{\chi}(x) e_p(ax) \\ &= \overline{\chi}(a)\tau_p + \chi(a)\overline{\tau_p}. \end{aligned}$$

The following lemma yields the Fourier transform of the set $C_p$.

**Lemma 7.1.** *For any nonzero integer $m$ modulo $p$,*

$$\widehat{C_p}(m) = \frac{1}{3p}(\overline{\chi}(m)\tau_p + \chi(m)\overline{\tau_p} + 2).$$

**Proof.** Let $m$ be a nonzero integer modulo $p$. Then

$$\begin{aligned}\widehat{C_p}(m) &= \frac{1}{p}\sum_{x=0}^{p-1} e_p(-mx)C_p(x) = \frac{1}{p}\sum_{x\in C_p} e_p(-mx)\\ &= \frac{1}{p}\left(1 + \sum_{x\in C_p-\{0\}} e_p(-mx)\right) = \frac{1}{p}\left(1 + \frac{1}{3}\sum_{x=1}^{p-1} e_p(-mx^3)\right)\\ &= \frac{1}{3p}\left(2 + \sum_{x=0}^{p-1} e_p(-mx^3)\right) = \frac{1}{3p}(\overline{\chi}(m)\tau_p + \chi(m)\overline{\tau_p} + 2),\end{aligned}$$

as $\chi(m) = \chi(-m)$, and we also apply the previous observation above. □

Now, we are ready to count the number of non-trivial 3-term arithmetic progressions in $C_p$.

**Proof of Theorem 1.5.** As we did in the proof of Lemma 2.4,

$$\begin{aligned}Q_p &= p^2(\widehat{C_p}(0))^3 + p^2\sum_{m=1}^{p-1}\widehat{C_p}(m)\widehat{C_p}(m)\widehat{C_p}(-2m) - \frac{p+2}{3}\\ &= \frac{(p+2)^3}{27p} + p^2\sum_{m=1}^{p-1}\widehat{C_p}(m)\widehat{C_p}(m)\widehat{C_p}(-2m) - \frac{p+2}{3}.\end{aligned} \tag{85}$$

Notice that $\tau_p\overline{\tau_p} = p$, one has $\widehat{C_p}(m) = \widehat{C_p}(-m)$, $\overline{\chi}(m)\chi(m) = 1$ and $\chi(m^2) = \chi(m)^2 = \overline{\chi}(m)$ for any nonzero $m$ modulo $p$. Then, using the observations above, for any nonzero $m$ modulo $p$, one sees that

$$\begin{aligned}&\widehat{C_p}(m)\widehat{C_p}(m)\widehat{C_p}(2m)\\ &= \frac{1}{27p^3}(\overline{\chi}(m)\tau_p + \chi(m)\overline{\tau_p} + 2)^2(\overline{\chi}(2m)\tau_p + \chi(2m)\overline{\tau_p} + 2)\\ &= \frac{1}{27p^3}(\overline{\chi}(m^2)\tau_p^2 + \chi(m^2)\overline{\tau_p}^2 + 4 + 2p + 4\overline{\chi}(m)\tau_p\\ &+ 4\chi(m)\overline{\tau_p})\cdot(\overline{\chi}(2m)\tau_p + \chi(2m)\overline{\tau_p} + 2)\\ &= \frac{1}{27p^3}\Big(\overline{\chi}(2)\tau_p^3 + \overline{\chi}(m)\chi(2)\tau_p^2\overline{\tau_p} + 2\chi(m)\tau_p^2 + \chi(m)\chi(2)\overline{\tau_p}^2\tau_p + \chi(2)\overline{\tau_p}^3\end{aligned}$$

$$+ 2\overline{\chi}(m)\overline{\tau_p}^2 + (2p+4)\overline{\chi}(2m)\tau_p + (2p+4)\chi(2m)\overline{\tau_p} + 4p + 8 + 4\overline{\chi}(2)\chi(m)\tau_p^2$$
$$\left. + 4\chi(2)p + 8\overline{\chi}(m)\tau_p + 4\overline{\chi}(2)p + 4\chi(2)\overline{\chi}(m)\overline{\tau_p}^2 + 8\chi(m)\overline{\tau_p}\right).$$

By orthogonality, for any non-principal Dirichlet character $h$ modulo $p$, we have that

$$\sum_{m=1}^{p-1} h(m) = 0.$$

By the previous calculations and orthogonality of $\chi$ and $\overline{\chi}$, Equation (85) becomes

$$Q_p = \frac{(p+2)^3}{27p} + \frac{p-1}{27p}(\overline{\chi}(2)\tau_p^3 + \chi(2)\overline{\tau_p}^3 + 4p + 8 + 4\chi(2)p + 4\overline{\chi}(2)p) - \frac{p+2}{3}. \quad (86)$$

As $\tau_p = \sqrt{p}e^{i\theta_p}$, we have the first part of the theorem. Note that $\chi(2)+\overline{\chi}(2) = 2$ if 2 is a cubic residue, and otherwise $\chi(2)+\overline{\chi}(2) = -1$. Also $\tau_p^3 + \overline{\tau_p}^3 = 2p^{3/2}\cos(3\theta_p)$. By [12, Chapter 3], one has that

$$\cos(3\theta_p) = \frac{a}{2\sqrt{p}},$$

where $4p = a^2 + 27b^2$ and $a \equiv 1 \pmod p$. By [12, Chapter 3], we know that

$$\tau_p^3 = p\sum_{t=1}^{p-1}\chi(t(1+t)) = p(A+Bw),$$

for some integers $A$ and $B$. Thus, $\overline{\tau_p}^3 = p(A+Bw^2)$. Moreover $p = A^2 - AB + B^2$ and $4p = (2A-B)^2 + 3B^2$. This yields that

$$\cos(3\theta_p) = \frac{2A-B}{2\sqrt{p}}$$

and

$$\sin(3\theta_p) = \frac{B\sqrt{3}}{2\sqrt{p}}.$$

Next, we compute the sum

$$z_p = \overline{\chi}(2)\tau_p^3 + \chi(2)\overline{\tau_p}^3.$$

If $\chi(2) = 1 = \overline{\chi}(2)$, we already computed the sum. Now, suppose $\overline{\chi}(2) = w = e^{2\pi i/3}$. Then,

$$z_p = 2p^{3/2}\cos(3\theta_p + 2\pi/3) = 2p^{3/2}\left(-\frac{\cos(3\theta_p)}{2} - \frac{\sin(3\theta_p)\sqrt{3}}{2}\right) \tag{87}$$

$$= 2p^{3/2}\left(-\frac{2A-B}{4\sqrt{p}} - \frac{3B}{4\sqrt{p}}\right) = -p(A+B). \tag{88}$$

Similarly, if $\overline{\chi}(2) = w^2 = e^{-2\pi i/3}$, then

$$z_p = 2p^{3/2}\cos(3\theta_p - 2\pi/3) = 2p^{3/2}\left(-\frac{2A-B}{4\sqrt{p}} + \frac{3B}{4\sqrt{p}}\right) = -p(A-2B). \tag{89}$$

By assembling (86), (87), (89) and the value of $\chi(2) + \overline{\chi}(2)$, which is an element of the set $\{2, -1\}$, we deduce that

$$Q_p = \frac{(p+2)^3}{27p} + \frac{p-1}{27p}(pc_p + 4p + 8) - \frac{p+2}{3},$$

where $c_p \in \mathbb{Z}$ is a computable constant which depends on $p$.

Lastly, suppose that $p$ is of the form $u^2 + 27v^2$ for some integers $u$ and $v$ with $u \equiv 2 \pmod 3$. Thus $4p = (2u)^2 + 27(2v)^2$. Then by [10, Theorem 4.15], we know that 2 is a cubic residue, in other words $\chi(2) = 1 = \overline{\chi}(2)$. Then, by (86) we conclude that

$$\begin{aligned}
Q_p &= \frac{(p+2)^3}{27p} + \frac{p-1}{27p}(\tau_p^3 + \overline{\tau_p}^3 + 12p + 8) - \frac{p+2}{3} \\
&= \frac{(p+2)^3}{27p} + \frac{p-1}{27p}(2p^{3/2}\cos(3\theta_p) + 12p + 8) - \frac{p+2}{3} \\
&= \frac{(p+2)^3}{27p} + \frac{p-1}{27p}(2p^{3/2}\frac{2u}{2\sqrt{p}} + 12p + 8) - \frac{p+2}{3} \\
&= \frac{(p+2)^3}{27p} + \frac{p-1}{27p}(2up + 12p + 8) - \frac{p+2}{3}. \quad \square
\end{aligned}$$

**Example 7.2.** Let $p = 31$. Then $p = u^2 + 27v^2$ where $u = 2$ and $v = 1$. Applying our theorem, we deduce that there are 50 many non-trivial 3-term arithmetic progressions in $C_{31}$. Let $p = 43$. Then $p = u^2 + 27v^2$ where $u = -4$ and $v = 1$. Applying our theorem, we deduce that there are 70 many non-trivial 3-term arithmetic progressions in $C_{43}$.

## 8. Arithmetic progressions and the Sárközy problem in Salem sets

From now on, we will work with Salem families and under some Fourier coefficient bounds, we will show the existence of long arithmetic progressions in Salem families. Besides, we also deal with the generalized Sárközy problem in Salem families by making use of the Weil estimates on exponential sums. Note that the following result does not give Theorem 1.1 if $k \geq 4$.

**Proposition 8.1.** *Let $k \geq 3$ be a positive integer and $\{A_N\}_{N\in\mathcal{B}}$ be an $\alpha$-Salem family with the $\alpha$-Salem constant $C$ such that $\alpha \leq \frac{1}{k-1}$. Suppose that $|A_N| \geq 16(C+1)^2 N^{\frac{k-1}{k}}$ for all sufficiently large $N \in \mathcal{B}$. Then $A_N$ contains a non-trivial $k$-term arithmetic progression for sufficiently large $N$.*

**Proof.** For any fixed $\alpha \in \left(0, \frac{1}{k-1}\right]$, let $\{A_N\}_{N\in\mathcal{B}}$ be an $\alpha$-Salem family with the $\alpha$-Salem constant $C$ and $k \geq 3$ be a positive integer. By Lemma 2.4, we know that the number of $k$-APs in $A_N$ is

$$\frac{N^2|A_N|^k}{N^k} + R,$$

where

$$\begin{aligned} R = & N^2 \sum_{(x_1,x_2,\dots,x_{k-2})\neq 0} \widehat{A}_N(x_1)\cdots\widehat{A}_N(x_{k-2})\widehat{A}_N(x_1+2x_2+\cdots+(k-2)x_{k-2}) \\ & \cdot \widehat{A}_N(-2x_1-3x_2-\cdots-(k-1)x_{k-2}). \end{aligned}$$

Notice that there are $k$-terms in $R$ related with the Fourier transform at the values $x_1, x_2, \dots, x_{k-2},\ x_1+2x_2+\cdots+(k-2)x_{k-2}$ and $-2x_1-3x_2-\cdots-(k-1)x_{k-2}$. Set $y_i = x_i$ for $i \in \{1,\dots,k-2\}$, $y_{k-1} = x_1+2x_2+\cdots+(k-2)x_{k-2}$ and $y_k = -2x_1-3x_2-\cdots-(k-1)x_{k-2}$. We will look for an upper bound for $R$. There are two critical situations of the terms $y_1,\dots,y_k$ to be aware of: the former is terms being zero, and the latter is terms being nonzero. First, we will prove that at the same time, at most $k-2$ of the terms can be 0. Recall that from the definition of $R$, we don't let the case $x_1 = x_2 = \cdots = x_{k-2} = 0$. Clearly, if one of $y_s \neq 0$ and the remaining $y_i$'s are all 0, then both $y_{k-1} = x_1+2x_2+\cdots+(k-2)x_{k-2} = sx_s$ and $y_k = -2x_1-3x_2-\cdots-(k-1)x_{k-2} = -(s+1)x_s$ are zero, and this yields that $y_s = x_s = 0$ as well and this is impossible. Hence, at most $k-2$ of the terms can be 0. From elementary number theory, the following congruence

$$az \equiv b \pmod{N} \tag{90}$$

is solvable if and only if $\gcd(a,N) \mid b$, and if it is solvable, then it has $\gcd(a,N)$ many incongruent solutions. Next, we analyze the case where $k-2$ of the terms are zero. For this purpose, we assume $k-3$ of $x_1,\dots,x_{k-2}$ and one of

$$y_{k-1} = x_1+2x_2+\cdots+(k-2)x_{k-2} \text{ and } y_k = -2x_1-3x_2-\cdots-(k-1)x_{k-2}$$

is zero. Say $y_i = x_i \neq 0$ and $y_{k-1} = x_1+2x_2+\cdots+(k-2)x_{k-2} = 0$. Then since all remaining terms are zero, we must have $ix_i = 0$ and by (90), we have at most $i \leq k$ many choices for $x_i$ (a similar case holds for $y_k = 2x_1+\cdots+(k-1)x_{k-2} = 0$).

Now suppose $k-4$ of $x_1,\dots,x_{k-2}$ are 0 and $y_{k-1}=y_k=0$. Without loss of generality assume that $x_m, x_s \neq 0$ with $1 \leq m < s \leq k-2$. Then, we have the following system of equations:

$$(m+1)x_m + (s+1)x_s = 0$$
$$mx_m + sx_s = 0,$$

and it gives that $x_s = -x_m$ and $(m-s)x_m = 0$. Once again, we have at most $k$ many options for $x_m$. As a result, the contribution of this case to $R$ is at most

$$k\binom{k}{2}\frac{|A_N|^{k-2}}{N^{k-4}}C^2N^{-2}|A_N|^{2\alpha} = k\binom{k}{2}C^2\frac{|A_N|^{k-2+2\alpha}}{N^{k-2}}. \tag{91}$$

If $k-3$ of them are zero and say the terms $y_{i_1}, y_{i_2}, y_{i_3}$ are nonzero, we get the sum of at most $\binom{k}{3}$ many expressions in the following form

$$N^2\frac{|A_N|^{k-3}}{N^{k-3}}\widehat{A}_N(y_{i_1})\widehat{A}_N(y_{i_2})\widehat{A}_N(y_{i_3}). \tag{92}$$

There are basically two cases for the choices of $y_{i_1}, y_{i_2}, y_{i_3}$. The first option is they are of the form $x_i, x_j, y_\ell$ where $i, j$ are distinct and $\ell \in \{k-1, k\}$, and the second option is they are of the form $x_i, x_j, x_\ell$ where $i, j$ and $\ell$ are distinct. For the first option, we may assume that $\ell = k-1$. Thus $y_k = -(i+1)x_i - (j+1)x_j = 0$ and $x_j$ depends on $x_i$ and there are at most $k$ choices for $x_j$ when $x_i$ is fixed by (90). For the second option, we have $y_{k-1} = y_k = 0$, in other words, $ix_i + jx_j + \ell x_\ell = -(i+1)x_i - (j+1)x_j - (\ell+1)x_\ell = 0$. From this, we deduce that $x_\ell = -x_i - x_j$ and $(i-\ell)x_i + (j-\ell)x_j = 0$. Again, if $x_i$ is fixed then there are at most $k$ many choices for other variables. So the above sum (92) is actually a single variable sum up to at most $k$ many choices. It follows from the $\alpha$-Salem condition, the triangle inequality and the above observation that the contribution of this case to $R$ is at most

$$k\binom{k}{3}\frac{|A_N|^{k-3}}{N^{k-5}}C^3N^{-3}|A_N|^{3\alpha}N = k\binom{k}{3}C^3\frac{|A_N|^{k-3+3\alpha}}{N^{k-3}}. \tag{93}$$

If none of $y_{i_1}, y_{i_2}, \dots, y_{i_j}$ is zero and the other equations are zero, we have $\binom{k}{j}$ many expressions in the following form

$$N^2\frac{|A_N|^{k-j}}{N^{k-j}}\widehat{A}_N(y_{i_1})\widehat{A}_N(y_{i_2})\cdots\widehat{A}_N(y_{i_j}),$$

and similar to the above discussion, this brings the error term

$$k\binom{k}{j}N^2\frac{|A_N|^{k-j}}{N^{k-j}}C^j\frac{|A_N|^{j\alpha}}{N^j}N^{j-2} = k\binom{k}{j}C^j\frac{|A_N|^{k-j+j\alpha}}{N^{k-j}}.$$

If none of $y_1, y_2, \ldots, y_{k-2}, y_{k-1}$ and $y_k$ is zero, this yields the following most dominant error term:

$$\begin{aligned}&\left| N^2 \sum_{(y_1,y_2,\ldots,y_k)\neq 0} \widehat{A}_N(y_1)\cdots\widehat{A}_N(y_{k-2})\widehat{A}_N(y_{k-1})\widehat{A}_N(y_k)\right| \\ &\quad \leq N^2 C^k \frac{|A_N|^{k\alpha}}{N^k} N^{k-2} = C^k |A_N|^{k\alpha}.\end{aligned}$$

As $C^j \leq (C+1)^k$ for any $j \in \{1, \ldots, k\}$ and for any non-negative real number $C$, for the error term $R$, we have the following estimate

$$|R| \leq k2^k (C+1)^k |A_N|^{k\alpha}.$$

Next, we will see that

$$k2^k(C+1)^k|A_N|^{k\alpha} \leq \frac{1}{2N^{k-2}}|A_N|^k,$$

in other words

$$|A_N|^{k(1-\alpha)} \geq k2^{k+1}(C+1)^k N^{k-2}$$

when $N$ is large enough. Recall that $|A_N| \geq 16(C+1)^2 N^{\frac{k-1}{k}}$ if $N$ is sufficiently large and $\alpha \leq \frac{1}{k-1}$. Therefore, $1-\alpha \geq 1 - \frac{1}{k-1}$. Then,

$$|A_N|^{k(1-\alpha)} \geq 16^{k(1-\frac{1}{k-1})}(C+1)^{2k(1-\frac{1}{k-1})}N^{k-2} = 2^{4k(1-\frac{1}{k-1})}(C+1)^{2k(1-\frac{1}{k-1})}N^{k-2}. \tag{94}$$

Observe that $2(1-\frac{1}{k-1}) \geq 1$, $4k(1-\frac{1}{k-1}) \geq 2k$ and $2^{2k} \geq k2^{k+1}$ when $k \geq 3$. Hence, by (94) and by the previous inequalities, we obtain that

$$|R| \leq \frac{1}{2N^{k-2}}|A_N|^k$$

when $N$ is large enough. Therefore, when $N$ is large enough

$$\begin{aligned}\sum_{t\in\mathbb{Z}_N} Q_N(t) &= \left| \frac{|A_N|^k}{N^{k-2}} + R \right| \\ &\geq \left| \frac{|A_N|^k}{N^{k-2}} - |R| \right| \\ &\geq \frac{1}{2N^{k-2}}|A_N|^k.\end{aligned}$$

There are $|A_N|$ trivial $k$-APs in $A_N$. So, the number of non-trivial $k$-APs in $A_N$ is at least

$$\begin{aligned}\frac{1}{2N^{k-2}}|A_N|^k - |A_N| &= |A_N|\left(\frac{1}{2N^{k-2}}|A_N|^{k-1} - 1\right)\\ &\geq 16(C+1)^2\ N^{\frac{k-1}{k}}\left(\frac{1}{2N^{k-2}}16^{k-1}(C+1)^{2k-2}N^{(k-1)^2/k} - 1\right)\\ &= 16(C+1)^2\ N^{\frac{k-1}{k}}\left(\frac{16^{k-1}(C+1)^{2k-2}}{2}N^{1/k} - 1\right),\end{aligned}$$

which is positive if $N$ is sufficiently large. □

Now, let us make a convenient arrangement of Proposition 8.1 to prove a special case of Szemerédi's theorem when the Fourier coefficients are extremely small.

**Corollary 8.2.** *Let $\gamma(n) : \mathbb{Z}_{>0} \to \mathbb{R}_{>0}$ be an arithmetic function such that $\gamma(n) = \mathcal{O}_\varepsilon(n^\varepsilon)$ for every $\varepsilon > 0$. Let $A$ be any subset of positive integers and $\{A_N\}_{N\in\mathcal{B}}$, where $\mathcal{B}$ is an infinite subset of positive integers, be a family of subsets of positive integers such that $A_N = A \cap \{1, 2, \dots, N\}$. Moreover, we regard $A_N$ as a subset of $\mathbb{Z}_N$. If $\{A_N\}_{N\in\mathcal{B}}$ satisfies the following Fourier coefficient bounds and growth conditions*

$$|\widehat{A_N}(m)| \leq \frac{\gamma(|A_N|)}{N} \text{ and } |A_N| \geq \frac{N}{\gamma(N)},$$

*for each $N \in \mathcal{B}$ and $m \neq 0$, then $A$ contains arbitrarily long arithmetic progressions.*

**Proof.** Suppose that $\{A_N\}_{N\in\mathcal{B}}$ satisfies the Fourier bound condition $|\widehat{A_N}(m)| \leq \frac{\gamma(|A_N|)}{N}$ when $m$ is nonzero and $|A_N| \geq \frac{N}{\gamma(N)}$ holds for each $N \in \mathcal{B}$. Let $k \geq 3$ be given. It follows from the definition of $\gamma$ that

$$|\widehat{A_N}(m)| \leq \frac{|A_N|^{\frac{1}{k-1}}}{N},$$

when $m$ is nonzero and

$$|A_N| \geq \frac{N}{\gamma(N)} \geq 64N^{\frac{k-1}{k}}$$

holds for sufficiently large $N \in \mathcal{B}$. Thus, $A_N$ contains a non-trivial $k$-term $\mathbb{Z}_N$-arithmetic progression for sufficiently large $N \in \mathcal{B}$ by Proposition 8.1. Hence, $A$ contains arbitrarily long $\mathbb{Z}$-arithmetic progressions by Proposition 2.3. □

As we mentioned before, the sets $\{S_p\}$ constitute a $\frac{1}{2}$-Salem family. Now, we consider the generalized Sárközy problem for $\alpha$-Salem families.

**Proposition 8.3.** *Let $\{A_N\}_{N\in\mathcal{P}}$ be an $\alpha$-Salem family for some fixed $\alpha \in (0, \frac{3}{4})$ with the $\alpha$-Salem constant $C$, where $\mathcal{P}$ is an infinite subset of primes. Let $f \in \mathbb{Z}[X]$ be a non-linear polynomial. Suppose that*

$$|A_N| \geq DN^{\frac{1}{4(1-\alpha)}}$$

*for all $N \in \mathcal{P}$ large enough, where $D = (2(\deg f - 1)C^2)^{\frac{1}{2(1-\alpha)}}$. Then, for sufficiently large $N$, the set $A_N$ contains two distinct elements whose difference is $f(i)$ for some $i \in \mathbb{Z}_N$.*

**Proof.** For any fixed $\alpha \in (0, \frac{3}{4})$, let $\{A_N\}_{N\in\mathcal{P}}$ be an $\alpha$-Salem family with the $\alpha$-Salem constant C. Moreover, we may assume that $f \notin N\mathbb{Z}[X]$ by taking $N \in \mathcal{P}$ sufficiently large. Now, let us define

$$Q(t) = |\{(x, y) \in A_N \times A_N : x - y = f(t)\}|.$$

So, the number of pairs in $A_N$ whose difference is in the image of $f$ is at least

$$\frac{1}{\deg f} \sum_{t\in\mathbb{Z}_N} Q(t) = \frac{1}{\deg f} \sum_{x,t\in\mathbb{Z}_N} A_N(x)A_N(x + f(t)).$$

If we use the Fourier inversion formula, then we get that

$$\begin{aligned} \sum_{t\in\mathbb{Z}_N} Q(t) &= \sum_{x,m,s\in\mathbb{Z}_N} \widehat{A}_N(s)e_N(sx)\widehat{A}_N(m)e_N(mx) \sum_{t\in\mathbb{Z}_N} e_N(mf(t)) \\ &= \sum_{s,m\in\mathbb{Z}_N} \widehat{A}_N(s)\widehat{A}_N(m) \sum_{x\in\mathbb{Z}_N} e_N((s + m)x) \sum_{t\in\mathbb{Z}_N} e_N(mf(t)). \end{aligned} \tag{95}$$

It follows from Equation (3) that

$$\sum_{x\in\mathbb{Z}_N} e_N((s + m)x) = \begin{cases} N & \text{if } s + m = 0, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, one sees that

$$\begin{aligned} \sum_{t\in\mathbb{Z}_N} Q(t) &= N \sum_{m\in\mathbb{Z}_N} \widehat{A}_N(-m)\widehat{A}_N(m) \sum_{t\in\mathbb{Z}_N} e_N(mf(t)) \\ &= N\widehat{A}_N(0)\widehat{A}_N(0)N + N \sum_{m\neq 0} \widehat{A}_N(-m)\widehat{A}_N(m) \sum_{t\in\mathbb{Z}_N} e_N(mf(t)) \\ &= |A_N|^2 + N \sum_{m\neq 0} \widehat{A}_N(-m)\widehat{A}_N(m) \sum_{t\in\mathbb{Z}_N} e_N(mf(t)). \end{aligned}$$

We denote this by

$$\sum_{t\in\mathbb{Z}_N} Q(t) = |A_N|^2 + R,$$

where

$$R = N \sum_{m \neq 0} \widehat{A}_N(-m)\widehat{A}_N(m) \sum_{t \in \mathbb{Z}_N} e_N(mf(t)).$$

By triangle inequality, Salem condition and Theorem 2.15, we infer that

$$\begin{aligned} |R| &\leq (\deg f - 1) \cdot N^{\frac{3}{2}} \sum_{m \neq 0} |\widehat{A}_N(-m)||\widehat{A}_N(m)| \\ &\leq (\deg f - 1) \cdot N^{\frac{3}{2}} \frac{C^2}{N^2} |A_N|^{2\alpha} N \\ &= (\deg f - 1) \cdot C^2 \cdot \sqrt{N} |A_N|^{2\alpha}. \end{aligned}$$

Hence, for $N$ large enough

$$\begin{aligned} \sum_{t \in \mathbb{Z}_N} Q(t) &= \left| |A_N|^2 + R \right| \\ &\geq \left| |A_N|^2 - |R| \right| \\ &\geq |A_N|^2 - (\deg f - 1) \cdot C^2 \cdot \sqrt{N} |A_N|^{2\alpha}. \end{aligned}$$

As we want to count the distinct pairs, we must subtract those off. There are at most $|A_N|$ of them. So the number of distinct elements whose difference is $f(i)$ for some $i \in \mathbb{Z}_N$ is greater than or equal to

$$\frac{1}{\deg f}\left( |A_N|^2 - (\deg f - 1) \cdot C^2 \cdot \sqrt{N} |A_N|^{2\alpha} \right) - |A_N| \geq \frac{|A_N|^2}{2 \deg f} - |A_N|,$$

which is positive when $N$ is large enough since $\alpha \in (0, \frac{3}{4})$ and $|A_N| \geq D N^{\frac{1}{4(1-\alpha)}}$, where $D = (2(\deg f - 1)C^2)^{\frac{1}{2(1-\alpha)}}$. □

**Corollary 8.4.** *Let $f \in \mathbb{Z}[X]$ be a non-constant polynomial and $n \geq 2$. Consider the Diophantine equation*

$$x^n - y^n = f(z). \tag{96}$$

*If $p$ is a sufficiently large prime number, then (96) always admits a solution in $\mathbb{F}_p$ with $x \neq y$.*

**Proof.** Consider the family $\{\Omega_p^n\}_{p \in \mathbb{P}}$, where $\Omega_p^n = \{t^n : t \in \mathbb{F}_p\}$. We know that this family is a $\frac{1}{2}$-Salem family and $|\Omega_p^n| \geq 1 + \frac{p-1}{n}$. Then, we conclude the corollary by the previous proposition. □

## Data availability

No data was used for the research described in the article.

## Acknowledgment

This work is supported by the Scientific and Technological Research Council of Turkey (TÜBİTAK) with the project number 122F027, and it is carried out by the second author. We would like to thank Antonio Rojas-León for pointing out Theorem 2.12 to us. The authors would like to thank the referee for many valuable comments which immensely improved the quality of the manuscript.

## References

[1] N.S. Aladov, Sur la distribution des résidus quadratiques et non-quadratiques d'un nombre premier $p$ dans la suite $1, 2, ..., p - 1$, Mat. Sb. 18 (1896) 61–75.
[2] T.M. Apostol, An Introduction to Analytic Number Theory, Springer-Verlag, New York, 1976.
[3] T. Barnet-Lamb, D. Geraghty, M. Harris, R. Taylor, A family of Calabi–Yau varieties and potential automorphy II, Publ. Res. Inst. Math. Sci. 47 (2011) 29–98.
[4] V. Bergelson, A. Leibman, Polynomial extensions of van der Waerden's and Szemerédi's theorems, J. Am. Math. Soc. 9 (1996) 725–753.
[5] B. Berndt, R. Evans, The determination of Gauss sums, Bull. Am. Math. Soc. 5 (1981) 107–129.
[6] B. Berndt, R. Evans, K. Williams, Gauss and Jacobi Sums, Wiley and Sons, 1998.
[7] J. Bourgain, On arithmetic progressions in sums of sets of integers, in: A Tribute to Paul Erdös, CUP, 1990.
[8] L. Clozel, M. Harris, R. Taylor, Automorphy for some $\ell$-adic lifts of automorphic mod $\ell$ Galois representations, Publ. Math. IHES 108 (2008) 1–182.
[9] K. Conrad, Quadratic residue patterns modulo a prime, https://kconrad.math.uconn.edu/blurbs/ugradnumthy/QuadraticResiduePatterns.pdf, 2018.
[10] D.A. Cox, Primes of the form $x^2 + ny^2$, in: Pure and Applied Mathematics, in: A Wiley Series of Texts, 1989.
[11] H. Darmon, L. Merel, Winding quotients and some variants of Fermat's last theorem, J. Reine Angew. Math. 490 (1997) 81–100.
[12] H. Davenport, Multiplicative Number Theory, Springer, New York, 1980.
[13] H. Davenport, On the distribution of quadratic residues (mod $p$), J. Lond. Math. Soc. 6 (1931) 49–54.
[14] H. Davenport, On the distribution of quadratic residues (mod $p$) (second paper), J. Lond. Math. Soc. 8 (1933) 46–52.
[15] P. Deligne, Théorèmes de finitude en cohomologie l-adique, dans Cohomologie Etale, in: Séminaire de Géométrie Algébrique du Bois-Marie, SGA 4 1/2, in: Lecture Notes in Math., vol. 569, 1977, pp. 233–251.
[16] P. Deligne, La Conjecture de Weil I, IHES Publ. Math. 43 (1974) 273–308.
[17] P. Deligne, La Conjecture de Weil II, IHES Publ. Math. 52 (1980) 137–252.
[18] P. Deligne, Application de la formule des traces aux sommes trigonométriques, in: Cohomologie Etale (SGA 4 12), in: Lecture Notes in Math., vol. 569, Springer, 1977, pp. 168–232.
[19] P. Erdős, P. Turán, On some sequences of integers, J. Lond. Math. Soc. 11 (1936) 261–264.
[20] H. Furstenberg, Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions, J. Anal. Math. 31 (1977) 204–256.
[21] H. Furstenberg, Y. Katznelson, A density version of the Hales-Jewett theorem, J. Anal. Math. 57 (1991) 64–119.
[22] W.T. Gowers, A new proof of Szemerédi's Theorem for arithmetic progressions of length four, Geom. Funct. Anal. 8 (1998) 529–551.
[23] W.T. Gowers, A new proof of Szemerédi's Theorem, Geom. Funct. Anal. 11 (2001) 465–588.

[24] M. Harris, N. Shepherd-Barron, R. Taylor, A family of Calabi–Yau varieties and potential automorphy, Ann. Math. 171 (2010) 779–813.

[25] H. Hasse, Zur Theorie der abstrakten elliptischen Funktionenkörper I. II. III, J. Reine Angew. Math. 175 (1936) 55–62, 69–88, 193–208.

[26] D.R. Heath-Brown, S.J. Patterson, The distribution of Kummer sums at prime arguments, J. Reine Angew. Math. 310 (1979) 111–130.

[27] N.M. Katz, Gauss Sums, Kloosterman Sums, and Monodromy Groups, (AM-116), Princeton University Press, 1988.

[28] N.M. Katz, Estimates for nonsingular multiplicative character sums, Int. Math. Res. Not. 7 (2002) 333–349.

[29] R. Lercier, F. Morain, Counting the number of points on elliptic curves over finite fields: strategies and performances, in: L.C. Guillou, J.J. Quisquater (Eds.), Advances in Cryptology – EUROCRYPT '95, EUROCRYPT 1995, in: Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 1995.

[30] A. Lott, Roth's theorem on arithmetic progressions, https://people.math.rochester.edu/faculty/iosevich/lott17.pdf, 2017.

[31] M.R. Murty, V.K. Murty, The Sato–Tate conjecture and generalizations, in: Current Trends in Science: Platinum Jubilee Special, Indian Academy of Sciences, 2009, pp. 639–646, https://mast.queensu.ca/~murty/Sato-Tate-CurrentTrends.pdf.

[32] R.K. Pandey, On certain sums with quadratic expressions involving the Legendre symbol, J. Integer Seq. 21 (2018) 1–10.

[33] A. Rojas-León, Estimates for singular multiplicative character sums, Int. Math. Res. Not. 20 (2005) 1221–1234.

[34] A. Rojas-León, On a generalization of Jacobi sums, Finite Fields Appl. 77 (2022) 1–15.

[35] K. Roth, On certain sets of integers, J. Lond. Math. Soc. 28 (1953) 104–109.

[36] SageMath, The Sage Mathematics Software System (Version 8.3), The Sage Developers, http://www.sagemath.org, 2018.

[37] A. Sárközy, On difference sets of sequences of integers I, Acta Math. Acad. Sci. Hung. 31 (1978) 125–149.

[38] R. Schoof, Elliptic curves over finite fields and the computation of square roots mod $p$, Math. Comput. 44 (170) (1985) 483–494.

[39] R. Schoof, Counting points on elliptic curves over finite fields, J. Théor. Nr. Bordx. 7 (1995) 219–254.

[40] J.H. Silverman, Advanced Topics in the Arithmetic of Elliptic Curves, Graduate Texts in Mathematics, Springer-Verlag, 1999.

[41] E.M. Stein, R. Shakarchi, Fourier Analysis: An Introduction, Princeton University Press, 2003.

[42] E. Szemerédi, On sets of integers containing no 4 elements in arithmetic progression, Acta Math. Acad. Sci. Hung. 20 (1969) 89–104.

[43] E. Szemerédi, On sets of integers containing no $k$ elements in arithmetic progression, Acta Arith. 27 (1975) 199–245.

[44] J. Tate, Endomorphisms of abelian varieties over finite fields, Invent. Math. 2 (1966) 134–144.

[45] R. Taylor, Automorphy for some $\ell$-adic lifts of automorphic mod $\ell$ Galois representations II, Publ. Math. IHES 108 (2008) 183–239.

[46] B.L. van der Waerden, Beweis einer Baudetschen Vermutung, Nieuw Arch. Wiskd. 15 (1927) 212–216.

[47] L.T. Washington, Elliptic Curves: Number Theory and Cryptography, Chapman and Hall/CRC, 2008.

[48] A. Weil, On some exponential sums, Proc. Natl. Acad. Sci. USA 34 (1948) 204–207.