SETS IN FINITE FIELDS AVOIDING POLYNOMIAL PROGRESSIONS OF LENGTH 2 OR 3

TRISTAN SHIN

Abstract. In this expository paper, we give an account of the quantitative bounds for two cases of the polynomial Szemerédi theorem over a finite field $\mathbb{F}_q$. For two-term progressions, we prove nontrivial upper and lower bounds on the maximum size of a set avoiding progressions of the form $x, x+y^2$. For three-term progressions, we discuss and mostly prove the upper bound of $O(q^{23/24})$, due to Peluse, on the size of a set avoiding progressions of the form $x, x+P(y), x+Q(y)$, where $P, Q \in \mathbb{Z}[y]$ are linearly independent and have constant term 0.

1. Introduction

Let $r_3(N)$ denote the maximum size of a subset of $\{1, \dots, N\}$ containing no nontrivial three-term arithmetic progressions, i.e. progressions of the form $x, x+y, x+2y$ with $y \neq 0$. Roth's theorem [Rot52] gives a qualitative bound on $r_3(N)$, namely that[1] $r_3(N) = o(N)$. There are many natural generalizations to ask. One important generalization to extend this to progressions of length k for $k \geq 3$. In this setting, Szemerédi's theorem [Sze75] gives the qualitative bound that the maximum size of a subset of $\{1, \dots, N\}$ containing no progression of the form $x, x+y, x+2y, \dots, x+(k-1)y$ with $y \neq 0$ is $o(N)$.

A further extension is if we replace the sequence $y, 2y, \dots, (k-1)y$ with arbitrary polynomials $P_1(y), \dots, P_{k-1}(y)$. For $P_1, \dots, P_{k-1} \in \mathbb{Z}[y]$, let $r_{P_1,\dots,P_{k-1}}(N)$ denote the maximum size of a subset of $\{1, \dots, N\}$ containing no nontrivial progression of the form $x, x+P_1(y), \dots, x+P_{k-1}(y)$. Determining bounds on $r_{P_1,\dots,P_{k-1}}(N)$ is the **polynomial Szemerédi problem**.

Then even the $k = 2$ case is interesting. For example, is there a nontrivial asymptotic bound on the size of a set containing no progressions of the form $x, x+P(y)$ with $P(y) \neq 0$ for any polynomial P with $\deg P \geq 2$? Unfortunately, we cannot get an $o(N)$ bound in general. For example, there are dense subsets of $\{1, \dots, N\}$ with no progressions of the form $x, x+y^2+1$. Indeed, the set of multiples of 3 has density $1/3$, and cannot contain both x and $x+y^2+1$ because $(x+y^2+1) - x = y^2+1 \not\equiv 0 \pmod 3$. More generally, if P has no root in some $\mathbb{Z}/m\mathbb{Z}$, then $m\,\mathbb{Z} \cap \{1, \dots, N\}$ has no progressions of the form $x, x+P(y)$.

If we had $P(y) = y^2$ instead, there are no obvious dense constructions, and it is not immediately obvious if there are nontrivial bounds or not. The first result on this problem came independently from Furstenberg [Fur77] and Sárközy [So78], showing that $r_{y^2}(N) = o(N)$ (known as the Furstenberg–Sárközy theorem). Furstenberg's proof used topological methods and thus did not provide any quantitative bounds. Sárközy used Fourier analytic methods similar to those used in Roth's theorem, which led to quantitative bounds of $r_{y^2}(N) \lesssim N(\log\log N)^{2/3}(\log N)^{-1/3}$. The current best upper bound is a recent result of Green and Sawhney [GS24] that $r_{y^2}(N) \lesssim N/\exp(\Omega(\sqrt{\log N}))$. In the other direction, the current best lower bound is $r_{y^2}(N) \gtrsim N^\gamma$ for $\gamma \approx 0.7334$, due to Beigel and Gasarch [BG08].

Department of Mathematics, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0112, USA

E-mail address: trshin@ucsd.edu.

Date: 11 Feb 2025.

[1] See Section 2.1.1 for asymptotic notation.

For general P of degree, the construction above suggests that we impose a congruence condition. We say that P is an **intersective polynomial** if for all $m \in \mathbb{N}$, P has a zero over $\mathbb{Z}/m\mathbb{Z}$. As part of a more general result mentioned in the next paragraph, Bergelson, Leibman, and Lesigne [BLL08] showed that $r_P(N) = o(N)$ if and only if P is intersective. The current best upper bound of $r_P(N) \leq N/(\log N)^{\Omega(\log\log\log\log N)}$ is due to Rice [Ric19]. A lower bound can be constructed greedily. If $d = \deg P$, then in order for $1 \leq x, x + P(y) \leq N$, we must have $1 \leq x \leq N$ and $|y| \lesssim N^{1/d}$. This allows us to construct a subset $A \subseteq \{1, \dots, N\}$ of size $|A| \gtrsim N^{1-1/d}$ avoiding $x, x + P(y)$. So $r_P(N) \gtrsim N^{1-1/d}$.

Returning to the general k case, we can generalize the congruence condition on the polynomials. We say that the P_i are **jointly intersective** if for all $m \in \mathbb{N}$, the P_i have a common zero over $\mathbb{Z}/m\mathbb{Z}$. Then Bergelson, Leibman, and Lesigne [BLL08] showed the polynomial Szemerédi theorem: we have that $r_{P_1,\dots,P_{k-1}}(N) = o(N)$ if and only if the P_i are jointly intersective. A lower bound can be obtained by looking at the subprogression $x, x + P_i(y)$, where P_i has maximal degree. If $d = \max \deg P_i$, then the bound from the $k = 2$ case gives that $r_{P_1,\dots,P_{k-1}}(N) \gtrsim N^{1-1/d}$.

To simplify the intersective condition a bit, it is certainly the case that the P_i have a common zero over $\mathbb{Z}/m\mathbb{Z}$ for all m if they have a common zero over $\mathbb{Z}$. By shifting the polynomials, we can assume that this common zero is 0, i.e. all of the P_i have constant term 0. In this case, the upper bound of $r_{P_1,\dots,P_{k-1}}(N) = o(N)$ was proven earlier by Bergelson and Leibman [BL96]. Quantitative bounds for $k \geq 3$, even in special cases, did not appear until much later. Prendiville [Pre17] showed that $r_{P_1,\dots,P_{k-1}}(N) \lesssim N/(\log\log N)^c$ for some constant c when $P_i = c_i y^d$ for some constants c_i. Peluse and Prendiville [PP24] then showed the same bound in the case of the *nonlinear Roth configuration*, i.e. $k = 3$ and $(P_1(y), P_2(y)) = (y, y^2)$. Peluse [Pel20] later generalized this bound to all k when the P_i have distinct degrees and constant term 0.

To tackle problems in additive combinatorics over the integers, it is often useful to formulate the problem over a finite field and solve that problem first. For example, Meshulam proved a similar result as Roth's theorem but over $\mathbb{F}_p^n$ where p is an odd prime and n is a positive integer. This proof was an adaptation of Roth's original Fourier-analytic proof, demonstrating the simplicity with which one can show corresponding results in the finite field setting. This is the basis behind the **finite field model**, which is detailed in the surveys of Green [Gre05], Wolf [Wol15], and Peluse [Pel24].

With that in mind, we turn our attention to the main subject of this paper, namely the polynomial Szemerédi problem over a finite field $\mathbb{F}_q$. For $P_1, \dots, P_{k-1} \in \mathbb{Z}[y]$, let $r_{P_1,\dots,P_{k-1}}(\mathbb{F}_q)$ denote the maximum size of a subset of $\mathbb{F}_q$ containing no nontrivial progression of the form $x, x + P_1(y), \dots, x + P_{k-1}(y)$. We focus on two cases of interest in this paper:

- The case $k = 2$ and $P(y) = y^2$, i.e. the Furstenberg–Sárközy problem over $\mathbb{F}_q$.
- The case $k = 3$, where we avoid progressions of the form $x, x + P(y), x + Q(y)$ for some linearly independent $P, Q \in \mathbb{Z}[y]$ both with constant term 0.

For Furstenberg–Sárközy over $\mathbb{F}_q$, we first note that the problem is only nontrivial if $q \equiv 1 \pmod 4$. Indeed, suppose $q \equiv 3 \pmod 4$. Then -1 is not a square in $\mathbb{F}_q$, so for any distinct $a_1, a_2 \in A \subseteq \mathbb{F}_q$, either $a_1 - a_2$ or $a_2 - a_1$ is a square. So $r_{y^2}(\mathbb{F}_q) = 1$ when $q \equiv 3 \pmod 4$.

When $q \equiv 1 \pmod 4$, we have nontrivial bounds on both sides, due to a combination of classical results and various authors.

Theorem 1.1. *Let $q \equiv 1 \pmod 4$. For all $\epsilon > 0$, we have that*

$$\frac{1-\epsilon}{2\log 2}\log q - o_\epsilon(1) \leq r_{y^2}(\mathbb{F}_q) \leq q^{1/2}.$$

For the general $k = 3$ case over $\mathbb{F}_q$, the first nontrivial result was given by Peluse [Pel18].

Theorem 1.2 (Peluse)**.** *If $\operatorname{char} \mathbb{F}_q$ is large enough (in terms of P and Q), then $r_{P,Q}(\mathbb{F}_q) \lesssim q^{23/24}$.*

Peluse [Pel19] later generalized this result to progressions of arbitrary length. This work in the finite field model was crucial for the integer setting, as the results of Peluse and Prendiville mentioned above [PP24, Pel20] adapted several tools and methods from the finite field results of Peluse.

In this paper, we focus on the two numbered theorems above. In Section 2, we provide preliminaries for the rest of the paper. In Section 3, we prove and give context to Theorem 1.1. In Section 4, we discuss the proof of Theorem 1.2, focusing on the analytic and combinatorial aspects and leaving as a black box several results from algebraic geometry.

1.1. **Acknowledgements.** The author would like to thank Mehtaab Sawhney for pointing out some useful references. This paper was written as part of Shachar Lovett's course on Additive Combinatorics and its Applications (UCSD CSE 291E, Fall 2024).

2. Preliminaries

2.1. **Notation.** Throughout this paper, q is a prime power and $\mathbb{F}_q$ is the finite field of order q.

Let $A \subseteq \mathbb{F}_q$. The **indicator function** of A is defined by

$$\mathbb{1}_A(x) \coloneqq \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{otherwise.} \end{cases}$$

The **density** of A is $|A|/q$. The **balanced indicator function** of A is $f_A \coloneqq \mathbb{1}_A - |A|/q$.

For any function f whose domain contains a finite S, the average of f over S is

$$\mathop{\mathbb{E}}_{x \in S} f(x) = \frac{1}{|S|} \sum_{x \in S} f(x).$$

For shorthand, let $\mathbb{E}\, f$ denote the average of f over all of $\mathbb{F}_q$. For $f \colon \mathbb{F}_q \to \mathbb{C}$, let

$$\|f\|_2 \coloneqq \left(\mathop{\mathbb{E}}_{x \in \mathbb{F}_q} |f(x)|^2 \right)^{\frac{1}{2}}.$$

2.1.1. *Asymptotic notation.* We use $f \lesssim g$, $f = O(g)$, and $g = \Omega(f)$ to denote $|f| \leq Cg$ for some constant $C > 0$. We use $f = o(g)$ to denote that $f/g \to 0$ as the argument tends to ∞. Subscripts denote that the hidden constant may depend on these parameters. For example, $f = O_\epsilon(g)$ means that for every ϵ, we have that $|f| \leq C_\epsilon g$ for some constant $C_\epsilon > 0$. In a stray from convention, whenever a result or bound is stated in terms of some polynomials $P_1, \ldots, P_{k-1}$, we always allow implicit constants to depend on $P_1, \ldots, P_{k-1}$ (including parameters deduced from these polynomials, such as k or $\deg P_i$).

2.2. **Characters.** There are two types of characters that we will use.

2.2.1. *Dual group and Fourier transform.* The **(Pontryagin) dual group** of $\mathbb{F}_q$, viewed as an abelian group with addition, is

$$\widehat{\mathbb{F}_q} \coloneqq \mathrm{Hom}(\mathbb{F}_q, \mathbb{T}) = \{\text{homomorphisms from } \mathbb{F}_q \text{ to } \mathbb{T}\}$$

where $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\} \cong \mathbb{R}/\mathbb{Z}$. The characters $\psi \in \widehat{\mathbb{F}_q}$ are the **additive characters** of $\mathbb{F}_q$. Note that $\widehat{\mathbb{F}_q}$ has multiplication as its group operation. We will write ψ_0 for the **trivial character** which evaluates to 1 on all of G. Then ψ_0 is the identity element of $\widehat{G}$.

We give the dual group the unnormalized counting measure, so the 2-norm of $g \colon \widehat{\mathbb{F}_q} \to \mathbb{C}$ is

$$\|g\|_{\ell^2} \coloneqq \left(\sum_{\psi \in \widehat{\mathbb{F}_q}} |g(\psi)|^2 \right)^{\frac{1}{2}}.$$

We have two versions of orthogonality for additive characters. The first is orthogonality along characters:

$$\mathop{\mathbb{E}}_{x\in\mathbb{F}_q}\psi(x)=\begin{cases}0 & \text{if } \psi \text{ nontrivial}\\ 1 & \text{if } \psi=\psi_0.\end{cases}$$

The second is orthogonality along the group:

$$\sum_{\psi\in\widehat{\mathbb{F}_q}}\psi(x)=\begin{cases}0 & \text{if } x\neq 0\\ q & \text{if } x=0.\end{cases}$$

For $f\colon \mathbb{F}_q\to\mathbb{C}$, its **Fourier transform** $\hat{f}\colon \widehat{\mathbb{F}_q}\to\mathbb{C}$ is given by

$$\hat{f}(\psi)\coloneqq \mathop{\mathbb{E}}_{x\in\mathbb{F}_q} f(x)\psi(-x).$$

We have the following fundamental facts about the Fourier transform, which can be proven using orthogonality:

- (Fourier inversion) $f=\sum_{\psi\in\widehat{\mathbb{F}_q}}\hat{f}(\psi)\psi$
- (Parseval) $\|f\|_2=\|\hat{f}\|_{\ell^2}$

One can also check that $\hat{f}(\psi^{-1})=\overline{\widehat{\overline{f}}(\psi)}$. In particular, if $f\colon \mathbb{F}_q\to\mathbb{R}$, then $\hat{f}(\psi^{-1})=\overline{\hat{f}(\psi)}$.

Let $A\subseteq\mathbb{F}_q$ have density α. Then we have the following:

- $\widehat{\mathbb{1}_A}(\psi_0)=\alpha$
- $\widehat{f_A}(\psi_0)=0$
- $\widehat{\mathbb{1}_A}(\psi)=\widehat{f_A}(\psi)$ for $\psi\neq\psi_0$
- $\|\mathbb{1}_A\|_2=\|\widehat{\mathbb{1}_A}\|_{\ell^2}=\alpha^{1/2}$
- $\|f_A\|_2=\|\widehat{f_A}\|_{\ell^2}=(\alpha-\alpha^2)^{1/2}$

2.2.2. *Multiplicative characters.* The **multiplicative characters** of $\mathbb{F}_q$ are the homomorphisms $\chi\in\operatorname{Hom}(\mathbb{F}_q^\times,\mathbb{T})$, extended to $\mathbb{F}_q$ by setting $\chi(0)=0$. The order of χ, denoted $\operatorname{ord}\chi$, is the least positive integer m such that $\chi(x)^m=1$ for all $x\in\mathbb{F}_q^\times$. The order exists because $\chi(x)^{q-1}=\chi(x^{q-1})=\chi(1)=1$ for all $x\in\mathbb{F}_q^\times$.

One such character that we will use is the **Legendre symbol**, i.e. the quadratic symbol, defined by

$$\left(\frac{x}{\mathbb{F}_q}\right)\coloneqq\begin{cases}1 & \text{if } x \text{ is a nonzero square in } \mathbb{F}_q\\ -1 & \text{if } x \text{ is a nonsquare in } \mathbb{F}_q\\ 0 & \text{if } x=0.\end{cases}$$

For q odd, some algebra implies that $\left(\frac{x}{\mathbb{F}_q}\right)=x^{(q-1)/2}$ as an equality over $\mathbb{F}_q$, which in particular implies that the Legendre symbol is a multiplicative character of $\mathbb{F}_q$.

2.2.3. *Weil bounds.* For both types of characters, there is a nontrivial bound on character sums provided by the Weil bounds.

Proposition 2.1 (Weil bound for additive characters)**.** *Let ψ be a nontrivial additive character of $\mathbb{F}_q$, and let $f\in\mathbb{F}_q[x]$. Suppose that* $\operatorname{char}\mathbb{F}_q\nmid\deg f$*. Then*

$$\left|\sum_{x\in\mathbb{F}_q}\psi(f(x))\right|\leq(\deg f-1)q^{1/2}.$$

Proposition 2.2 (Weil bound for multiplicative characters)**.** *Let χ be a multiplicative character of $\mathbb{F}_q$, and let $f \in \mathbb{F}_q[x]$. Suppose that $f \neq cg^{\operatorname{ord}\chi}$ for any $c \in \mathbb{F}_q$ and $g \in \mathbb{F}_q[x]$. Then*

$$\left|\sum_{x\in\mathbb{F}_q} \chi(f(x))\right| \leq (\deg f - 1)q^{1/2}.$$

3. Furstenberg–Sárközy over $\mathbb{F}_q$

In this section, we prove Theorem 1.1. Throughout this section, let $q \equiv 1 \pmod 4$.

3.1. **Upper bound.** We prove a power-saving upper bound on subsets of $\mathbb{F}_q$ that do not contain a progression of the form $x, x+y^2$ with $y \neq 0$.

Proposition 3.1. *We have that $r_{y^2}(\mathbb{F}_q) \leq q^{1/2} + O(1)$.*

We provide two proofs of this result. This first is more analytic in flavor and uses the following result, which is folklore and at least as old as Vinogradov [Vin54, Chapter 5].

Lemma 3.2. *Let χ be a nontrivial multiplicative character of $\mathbb{F}_q$. Then for any subsets $A, B \subseteq \mathbb{F}_q$, we have that*

$$\left|\sum_{a\in A}\sum_{b\in B} \chi(a+b)\right| \leq \sqrt{q|A||B|}.$$

Proof. We have that

$$\begin{aligned}\left|\sum_{a\in A}\sum_{b\in B} \chi(a+b)\right|^2 &\overset{\text{C–S}}{\leq} \left(\sum_{a\in A} 1^2\right)\left(\sum_{a\in A}\left|\sum_{b\in B}\chi(a+b)\right|^2\right)\\ &\leq |A|\cdot\sum_{a\in\mathbb{F}_q}\left|\sum_{b\in B}\chi(a+b)\right|^2\\ &= |A|\cdot\sum_{b\in B}\sum_{b'\in B}\sum_{a\in\mathbb{F}_q}\chi(a+b)\overline{\chi(a+b')}.\end{aligned}$$

If $b = b'$, then this inner sum is q. If $b \neq b'$, then the sum is

$$\sum_{a\in\mathbb{F}_q\setminus\{-b'\}} \chi\left(\frac{a+b}{a+b'}\right) \overset{c=\frac{a+b}{a+b'}}{=} \sum_{c\in\mathbb{F}_q\setminus\{1\}}\chi(c) = -1.$$

Thus we have

$$\left|\sum_{a\in A}\sum_{b\in B}\chi(a+b)\right|^2 \leq |A|\cdot|B|\cdot q - |A|\cdot(|B|^2-|B|)\cdot 1 \leq q|A||B|$$

as desired. □

Proof of Proposition 3.1. Let $A \subseteq \mathbb{F}_q$ have no nontrivial progressions of the form $x, x+y^2$. In Lemma 3.2, take χ to be the Legendre symbol over $\mathbb{F}_q$ and $B = -A$. For $a_1, a_2 \in A$, we have that

$$\chi(a_1 - a_2) = \begin{cases} -1 & \text{if } a_1 \neq a_2\\ 0 & \text{if } a_1 = a_2.\end{cases}$$

So the result is that

$$|A|^2 - |A| \leq \sqrt{q}|A|,$$

so $|A| \leq \sqrt{q}+1$ as desired. □

The second proof of the upper bound is a simple combinatorial argument, given in a survey of Croot and Lev [CL07] but possibly older. This gives the precise bound in Theorem 1.1.

Second proof of Proposition 3.1. Let $A \subseteq \mathbb{F}_q$ have no nontrivial progressions of the form $x, x+y^2$, and suppose $|A| > \sqrt{q}$. Fix some nonsquare $r \in \mathbb{F}_q^\times$. Consider $ra_1 - a_2$ for $a_1, a_2 \in A$. By pigeonhole, two of them must be equal, i.e. $ra_1 - a_2 = ra_3 - a_4$ for some $a_1, a_2, a_3, a_4 \in A$ with $(a_1, a_2) \neq (a_3, a_4)$. Note that $a_1 \neq a_3$, as $a_1 = a_3$ implies $a_2 = a_4$ also. So $r = \frac{a_2 - a_4}{a_1 - a_3}$. But $a_2 - a_4$ and $a_1 - a_3$ are both nonsquares, and the quotient of any two nonsquares is a square, so r is a square, contradiction. Thus $|A| \leq \sqrt{q}$. □

For the main order, this is the best upper bound known on $r_{y^2}(\mathbb{F}_q)$. The constant in front can be improved in some cases. For example, when $q = p$ for some prime $p \equiv 1 \pmod 4$, an argument of Hanson and Petridis [HP21] (using Stepanov's method of auxiliary polynomials) gives that $r_{y^2}(\mathbb{F}_p) \leq \frac{1}{\sqrt{2}} p^{1/2} + O(1)$. Their argument does not generalize to arbitrary $\mathbb{F}_q$ because it uses the nonvanishing of binomial coefficients when the first argument is less than the order of the field, which is false over $\mathbb{F}_q$ when q is not prime.

3.2. **Lower bound.** The lower bound roughly follows an argument by Cohen [Coh88], later simplified by Fabrykowski [Fab93], for a closely related problem about the clique number of the Paley graph.

The idea is to iteratively construct A. Run the following procedure:

(1) Initialize $A_0 = \mathbb{F}_q$ and $k = 1$.
(2) Choose any $a_k \in A_{k-1}$.
(3) Let $A_k = \{x \in \mathbb{F}_q : x - a_i \text{ is a nonsquare for } i = 1, \ldots, k\} \subseteq A_{k-1}$.
(4) If A_k is empty, set $\ell = k$ and **STOP**.
(5) Otherwise, increment k and go back to step (2).

Note that at step (3), we do not need conditions about $a_i - x$ because $x - a_i$ is a square if and only if $a_i - x$ is a square (using the fact that $q \equiv 1 \pmod 4$). Note that if A_k is nonempty, then $r_{y^2}(\mathbb{F}_q) \geq k + 1$. So we would like to lower bound $|A_k|$.

Let χ be the Legendre symbol over $\mathbb{F}_q$. Then

$$|A_k| = \sum_{\substack{x \in \mathbb{F}_q \\ x \notin \{a_1, \ldots, a_k\}}} \prod_{i=1}^{k} \frac{1 - \chi(x - a_i)}{2} = \sum_{x \in \mathbb{F}_q} \prod_{i=1}^{k} \frac{1 - \chi(x - a_i)}{2} - k \cdot \frac{1}{2}.$$

Let us estimate this sum. We have that

$$\sum_{x \in \mathbb{F}_q} \prod_{i=1}^{k} (1 - \chi(x - a_i)) = \sum_{x \in \mathbb{F}_q} \sum_{I \subseteq \{1, \ldots, k\}} (-1)^{|I|} \prod_{i \in I} \chi(x - a_i) = \sum_{I \subseteq \{1, \ldots, k\}} (-1)^{|I|} \sum_{x \in \mathbb{F}_q} \chi\left(\prod_{i \in I} (x - a_i) \right).$$

The index set $I = \emptyset$ contributes q to this sum. To bound the remaining terms, we use the Weil bound for multiplicative characters (Proposition 2.2). The Weil bound can be applied to the Legendre symbol χ of order 2 and $f(x) = \prod (x - a_i)$ to get that

$$\left| \sum_{x \in \mathbb{F}_q} \chi\left(\prod_{i \in I} (x - a_i) \right) \right| \leq (|I| - 1) q^{1/2}$$

whenever $|I| \geq 1$. So

$$\left| 2^k |A_k| + k 2^{k-1} - q \right| \leq \sum_{\substack{I \subseteq \{1, \ldots, k\} \\ I \neq \emptyset}} (|I| - 1) q^{1/2} = q^{1/2} \sum_{m=1}^{k} \binom{k}{m} (m - 1) = q^{1/2} \cdot (k 2^{k-1} - 2^k + 1),$$

so

$$|A_k| \geq q2^{-k} - q^{1/2} \cdot (k/2 - 1 + 2^{-k}) - k/2.$$

Let $\epsilon > 0$, and set $k = \lfloor \frac{1-\epsilon}{2 \log 2} \log q \rfloor$. Then

$$|A_k| \geq q^{(1+\epsilon)/2} - \frac{1-\epsilon}{4 \log 2} q^{1/2} \log q + q^{1/2} - 2q^{\epsilon/2} - \frac{1-\epsilon}{4 \log 2} \log q > 0$$

for q large enough (in terms of ϵ). This gives the following result.

Proposition 3.3. *For all $\epsilon > 0$, we have that $r_{y^2}(\mathbb{F}_q) \geq \frac{1-\epsilon}{2 \log 2} \log q - o_\epsilon(1)$.*

4. The case $k = 3$

We now turn our attention to the $k = 3$ case. Let $P, Q \in \mathbb{Z}[y]$ be linearly independent and have constant term 0, and let $A \subseteq \mathbb{F}_q$. Define

$$C_{P,Q}(A) \coloneqq \#\{\text{progressions } x, x + P(y), x + Q(y) \text{ in } A\}.$$

In order to prove the bound on $r_{P,Q}(\mathbb{F}_q)$, it suffices to prove the following counting theorem.

Theorem 4.1. *If $\operatorname{char} \mathbb{F}_q$ is large enough (in terms of P and Q), then for all $A \subseteq \mathbb{F}_q$, we have that*

$$C_{P,Q}(A) = |A|^3 q^{-1} + O(|A|^{3/2} q^{7/16}).$$

We provide a quick derivation of the bound using this counting theorem.

Proof of Theorem 1.2 using Theorem 4.1. Let $A \subseteq \mathbb{F}_q$ have no nontrivial progressions of the form $x, x + P(y), x + Q(y)$. Then $C_{P,Q}(A) = |A|$, as each of the pairs (x, y) with $x \in A$ and $y = 0$ gets counted once. So Theorem 4.1 tells us that

$$|A|^{3/2} - |A|^{-1/2} q = O(q^{23/16}).$$

If $|A| \leq 2q^{1/2}$, then we are done, so assume $|A| > 2q^{1/2}$. Then $|A|^{3/2} = O(q^{23/16})$, which is the desired bound. □

We provide most of the proof of Theorem 4.1 in this section. In particular, we fully discuss all the steps that primarily involve combinatorics and analysis, with the omitted parts being entirely algebrogeometric.

4.1. A trilinear average.

We would like to estimate $C_{P,Q}$. The first step is to connect $C_{P,Q}$ with a trilinear average. For $f_1, f_2, f_3 \colon \mathbb{F}_q \to \mathbb{C}$, define

$$\Lambda_{P,Q}(f_1, f_2, f_3) \coloneqq \mathop{\mathbb{E}}_{x,y \in \mathbb{F}_q} f_1(x) f_2(x + P(y)) f_3(x + Q(y)).$$

Let $A \subseteq \mathbb{F}_q$ have density α. Then we have that

$$C_{P,Q}(A) = q^2 \Lambda_{P,Q}(\mathbb{1}_A, \mathbb{1}_A, \mathbb{1}_A).$$

So the counting theorem we wish to prove is that

$$\Lambda_{P,Q}(\mathbb{1}_A, \mathbb{1}_A, \mathbb{1}_A) = \alpha^3 + O(\alpha^{3/2} q^{-1/16}).$$

The main term of this result is at least plausible—if A is a random set of density α, the expected value of $\Lambda_{P,Q}(\mathbb{1}_A, \mathbb{1}_A, \mathbb{1}_A)$ is precisely α^3. So we wish to show that the count does not deviate too much from its average.

It will be easier to work with these averages when one function is balanced, so we replace the last $\mathbb{1}_A$ with f_A. Using linearity, we have that

$$\Lambda_{P,Q}(\mathbb{1}_A, \mathbb{1}_A, \mathbb{1}_A) = \Lambda_{P,Q}(\mathbb{1}_A, \mathbb{1}_A, f_A) + \alpha \Lambda_{P,Q}(\mathbb{1}_A, \mathbb{1}_A, 1).$$

But

$$\Lambda_{P,Q}(f_1, f_2, 1) = \mathop{\mathbb{E}}_{x,y\in\mathbb{F}_q} f_1(x)f_2(x + P(y)).$$

This only depends on P, f_1, and f_2, so we can call it $\Lambda_P(f_1, f_2)$. So

$$\Lambda_{P,Q}(\mathbb{1}_A, \mathbb{1}_A, \mathbb{1}_A) = \Lambda_{P,Q}(\mathbb{1}_A, \mathbb{1}_A, f_A) + \alpha\Lambda_P(\mathbb{1}_A, \mathbb{1}_A).$$

Once again, we would like this last average to have a balanced function, so we replace the last $\mathbb{1}_A$ with f_A. We have that

$$\Lambda_P(\mathbb{1}_A, \mathbb{1}_A) = \Lambda_P(\mathbb{1}_A, f_A) + \alpha\Lambda_P(\mathbb{1}_A, 1).$$

But now

$$\Lambda_P(\mathbb{1}_A, 1) = \mathop{\mathbb{E}}_{x,y\in\mathbb{F}_q} \mathbb{1}_A(x) = \alpha,$$

so

$$\Lambda_{P,Q}(\mathbb{1}_A, \mathbb{1}_A, \mathbb{1}_A) = \Lambda_{P,Q}(\mathbb{1}_A, \mathbb{1}_A, f_A) + \alpha\Lambda_P(\mathbb{1}_A, f_A) + \alpha^3.$$

Thus we have the following bound.

Proposition 4.2. *We have that*

$$|\Lambda_{P,Q}(\mathbb{1}_A, \mathbb{1}_A, \mathbb{1}_A) - \alpha^3| \le |\Lambda_{P,Q}(\mathbb{1}_A, \mathbb{1}_A, f_A)| + \alpha|\Lambda_P(\mathbb{1}_A, f_A)|.$$

So to show Theorem 4.1, it suffices to bound $\Lambda_{P,Q}(\mathbb{1}_A, \mathbb{1}_A, f_A)$ by $\alpha^{3/2}q^{-1/16}$ and $\Lambda_P(\mathbb{1}_A, f_A)$ by $\alpha^{1/2}q^{-1/16}$.

4.2. **Bounding Λ_P.** Let us first tackle $\Lambda_P(\mathbb{1}_A, f_A)$. We prove the following upper bound for the bilinear average.

Lemma 4.3. *We have that* $|\Lambda_P(\mathbb{1}_A, f_A)| \le (\deg P)\alpha q^{-1/2}$.

Proof. Expand out the average, then expand the functions using Fourier coefficients:

$$\begin{aligned}
\Lambda_P(\mathbb{1}_A, f_A) &= \mathop{\mathbb{E}}_{x,y\in\mathbb{F}_q} \mathbb{1}_A(x)f_A(x + P(y)) \\
&= \mathop{\mathbb{E}}_{x,y\in\mathbb{F}_q} \left(\sum_{\gamma\in\widehat{\mathbb{F}_q}} \widehat{\mathbb{1}_A}(\gamma)\gamma(x)\right)\left(\sum_{\psi\in\widehat{\mathbb{F}_q}} \widehat{f_A}(\psi)\psi(x + P(y))\right) \\
&= \sum_{\gamma,\psi\in\widehat{\mathbb{F}_q}} \widehat{\mathbb{1}_A}(\gamma)\widehat{f_A}(\psi)\left(\mathop{\mathbb{E}}_{x\in\mathbb{F}_q} \gamma(x)\psi(x)\right)\left(\mathop{\mathbb{E}}_{y\in\mathbb{F}_q} \psi(P(y))\right).
\end{aligned}$$

The sum over x is 1 if $\gamma = \psi^{-1}$ and 0 otherwise, so

$$\Lambda_P(\mathbb{1}_A, f_A) = \sum_{\psi\in\widehat{\mathbb{F}_q}} \widehat{\mathbb{1}_A}(\psi^{-1})\widehat{f_A}(\psi)\left(\mathop{\mathbb{E}}_{y\in\mathbb{F}_q} \psi(P(y))\right).$$

First, the summand at ψ_0 is 0 because $\widehat{f_A}(\psi_0) = 0$. So we are left with the contributions from $\psi \ne \psi_0$. Then $\widehat{\mathbb{1}_A}(\psi^{-1}) = \overline{\widehat{\mathbb{1}_A}(\psi)}$ and $\widehat{f_A}(\psi) = \widehat{\mathbb{1}_A}(\psi)$. By the Weil bound for additive characters (Proposition 2.1), assuming $\operatorname{char}\mathbb{F}_q > \deg P$, we have that

$$\left|\mathop{\mathbb{E}}_{y\in\mathbb{F}_q} \psi(P(y))\right| < (\deg P)q^{-1/2}.$$

So we have that

$$|\Lambda_P(\mathbb{1}_A, f_A)| \le (\deg P)q^{-1/2} \sum_{\psi \in \widehat{\mathbb{F}_q}\setminus\{\psi_0\}} \left|\widehat{\mathbb{1}_A}(\psi)\right|^2 = (\deg P)q^{-1/2} \cdot (\alpha - \alpha^2)$$

as desired. □

This is tiny compared to the error we are allowed to tolerate because $\alpha \le 1$ and $q > 1$.

4.3. **Bounding $\Lambda_{P,Q}$ in terms of another bilinear average.** Now, we bound $\Lambda_{P,Q}$ in terms of another bilinear average. To do so, we repeatedly apply the Cauchy–Schwarz inequality in a specific manner. Let us record the formulation in a statement.

Lemma 4.4 (Cauchy–Schwarz lemma)**.** *Let S be a set. Suppose we have functions $f_i\colon \mathbb{F}_q \to \mathbb{R}$ and $\alpha_i\colon S \to \mathbb{F}_q$ for $i = 1, \dots, k$, as well as $g_j\colon \mathbb{F}_q \to \mathbb{R}$ and $\beta_j\colon S \to \mathbb{F}_q$ for $j = 1, \dots, \ell$. Define*

$$S' \coloneqq \{(y, \widetilde{y}) \in S \times S \,:\, \alpha_i(y) - \alpha_1(y) = \alpha_i(\widetilde{y}) - \alpha_1(\widetilde{y}) \text{ for } i = 1, \dots, k\}.$$

Then

$$\begin{aligned}&\mathop{\mathbb{E}}_{x\in\mathbb{F}_q} \mathop{\mathbb{E}}_{y\in S} \prod_{i=1}^{k} f_i(x + \alpha_i(y)) \prod_{j=1}^{\ell} g_j(x + \beta_j(y)) \\ &\le \frac{q^{(k-1)/2}|S'|^{1/2}}{|S|} \left(\prod_{i=1}^{k} \|f_i\|_2\right) \left(\mathop{\mathbb{E}}_{x\in\mathbb{F}_q} \mathop{\mathbb{E}}_{(y,\widetilde{y})\in S'} \prod_{j=1}^{\ell} g_j(x + \beta_j(y) - \alpha_1(y)) g_j(x + \beta_j(\widetilde{y}) - \alpha_1(\widetilde{y}))\right)^{1/2}.\end{aligned}$$

Proof. For $\mathbf{z} = (z_1, \dots, z_k) \in \mathbb{F}_q^k$, let

$$S_{\mathbf{z}} \coloneqq \{y \in S \,:\, \alpha_i(y) - \alpha_1(y) = z_i \text{ for } i = 1, \dots, k\},$$

so that

$$S' = \bigsqcup_{\substack{\mathbf{z}\in\mathbb{F}_q^k \\ z_1 = 0}} S_{\mathbf{z}} \times S_{\mathbf{z}}.$$

View the average as a normalized sum over the variables $x' = x + \alpha_1(y)$ and $z_i = \alpha_i(y) - \alpha_1(y)$ for $i = 1, \dots, k$. Then the average in question is equal to

$$\frac{1}{q|S|} \sum_{x'\in\mathbb{F}_q} \sum_{\substack{\mathbf{z}\in\mathbb{F}_q^k \\ z_1 = 0}} \left(\prod_{i=1}^{k} f_i(x' + z_i)\right) \left(\sum_{y\in S_{\mathbf{z}}} \prod_{j=1}^{\ell} g_j(x' + \beta_j(y) - \alpha_1(y))\right).$$

Apply Cauchy–Schwarz to this, summing over the first two sums and taking each of the large parentheses as the components of the two vectors. This gives that the average is at most

$$\frac{1}{q|S|} \left(\sum_{x'\in\mathbb{F}_q} \sum_{\substack{\mathbf{z}\in\mathbb{F}_q^k \\ z_1 = 0}} \left(\prod_{i=1}^{k} f_i(x' + z_i)\right)^2\right)^{1/2} \left(\sum_{x'\in\mathbb{F}_q} \sum_{\substack{\mathbf{z}\in\mathbb{F}_q^k \\ z_1 = 0}} \left(\sum_{y\in S_{\mathbf{z}}} \prod_{j=1}^{\ell} g_j(x' + \beta_j(y) - \alpha_1(y))\right)^2\right)^{1/2}.$$

The first double sum is straightforward. We can factor to see that it is

$$\sum_{x'\in\mathbb{F}_q} f_1(x')^2 \prod_{i=2}^{k} \sum_{z_i\in\mathbb{F}_q} f_i(x' + z_i)^2 = \sum_{x'\in\mathbb{F}_q} f_1(x')^2 \prod_{i=2}^{k} (q\|f_i\|_2^2) = q^k \prod_{i=1}^{k} \|f_i\|_2^2.$$

For the second sum, we expand out the square by duplicating the variable y into $\widetilde{y}$. It becomes

$$\sum_{x' \in \mathbb{F}_q} \sum_{\substack{\mathbf{z} \in \mathbb{F}_q^k \\ z_1 = 0}} \sum_{y, \widetilde{y} \in S_{\mathbf{z}}} \prod_{j=1}^{\ell} g_j(x' + \beta_j(y) - \alpha_1(y)) g_j(x' + \beta_j(\widetilde{y}) - \alpha_1(\widetilde{y})).$$

The second and third sum combine to a single sum over $(y, \widetilde{y}) \in S'$. After normalizing the two remaining sums into averages, this becomes

$$q|S'| \mathop{\mathbb{E}}_{x' \in \mathbb{F}_q} \mathop{\mathbb{E}}_{(y,\widetilde{y}) \in S'} \prod_{j=1}^{\ell} g_j(x' + \beta_j(y) - \alpha_1(y)) g_j(x' + \beta_j(\widetilde{y}) - \alpha_1(\widetilde{y})).$$

These two computations show that the upper bound obtained is indeed the desired result, upon changing x' to x. □

Remark 4.5. Though we provided the lemma in full generality, we will only ever use it when all the f_i are the same function f. Then the product of the norms of the f_i is replaced by $\|f\|_2^k$.

Note that the average inside the square root in the bound is of the form of the original average, so this lemma is ripe for repeated application. Indeed, let us do that with $\Lambda_{P,Q}(\mathbb{1}_A, \mathbb{1}_A, f_A)$. We have that

$$\Lambda_{P,Q}(\mathbb{1}_A, \mathbb{1}_A, f_A) = \mathop{\mathbb{E}}_{x \in \mathbb{F}_q} \mathop{\mathbb{E}}_{y \in \mathbb{F}_q} \mathbb{1}_A(x)\, \mathbb{1}_A(x + P(y)) f_A(x + Q(y)).$$

Applying the Cauchy–Schwarz lemma with $S = \mathbb{F}_q$ and $k = 1$ so that $S' = \mathbb{F}_q \times \mathbb{F}_q$, we have that $\Lambda_{P,Q}(\mathbb{1}_A, \mathbb{1}_A, f_A)^2$ is at most α times

$$\mathop{\mathbb{E}}_{x \in \mathbb{F}_q} \mathop{\mathbb{E}}_{(y_1,y_2) \in S'} \mathbb{1}_A(x + P(y_1))\, \mathbb{1}_A(x + P(y_2)) f_A(x + Q(y_1)) f_A(x + Q(y_2)).$$

Another application of the Cauchy–Schwarz lemma with S' and $k = 2$ so that

$$S'' = \{\mathbf{y} \in \mathbb{F}_q^4 \,:\, P(y_2) - P(y_1) = P(y_4) - P(y_3)\}$$

gives that $\Lambda_{P,Q}(\mathbb{1}_A, \mathbb{1}_A, f_A)^4$ is at most $\alpha^2 \cdot \frac{|S''|}{q^3} \alpha^2$ times

$$\mathop{\mathbb{E}}_{x \in \mathbb{F}_q} \mathop{\mathbb{E}}_{\mathbf{y} \in S''} f_A(x + Q(y_1) - P(y_1)) f_A(x + Q(y_3) - P(y_3)) f_A(x + Q(y_2) - P(y_1)) f_A(x + Q(y_4) - P(y_3)).$$

One final application of the Cauchy–Schwarz lemma with S'' and $k = 3$ so that

$$S''' = \left\{ \mathbf{y} \in \mathbb{F}_q^8 \,:\, \begin{array}{c} P(y_2) - P(y_1) = P(y_4) - P(y_3) \\ P(y_6) - P(y_5) = P(y_8) - P(y_7) \\ Q(y_3) - P(y_3) - Q(y_1) + P(y_1) = Q(y_7) - P(y_7) - Q(y_5) + P(y_5) \\ Q(y_2) - Q(y_1) = Q(y_6) - Q(y_5) \end{array} \right\}$$

gives that $\Lambda_{P,Q}(\mathbb{1}_A, \mathbb{1}_A, f_A)^8$ is at most $\alpha^4 \cdot \frac{|S''|^2}{q^6} \alpha^4 \cdot \frac{q^2 |S'''|}{|S''|^2} (\alpha - \alpha^2)^3$ times

$$\mathop{\mathbb{E}}_{x \in \mathbb{F}_q} \mathop{\mathbb{E}}_{\mathbf{y} \in S'''} f_A(x + Q(y_4) - P(y_3) - Q(y_1) + P(y_1)) f_A(x + Q(y_8) - P(y_7) - Q(y_5) + P(y_5)). \quad (\dagger)$$

Define a new bilinear average

$$\Lambda'_{P,Q}(f_1, f_2) \coloneqq \mathop{\mathbb{E}}_{x \in \mathbb{F}_q} \mathop{\mathbb{E}}_{\mathbf{y} \in S'''} f_1(x) f_2(x + R_{P,Q}(\mathbf{y})),$$

where

$$R_{P,Q}(\mathbf{y}) = Q(y_8) - Q(y_7) - Q(y_4) + Q(y_3).$$

Then the average in (†) is equal to $\Lambda'_{P,Q}(f_A, f_A)$. To see why, first let $x' = x + Q(y_4) - P(y_3) - Q(y_1) + P(y_1)$ so that

$$x+Q(y_8)-P(y_7)-Q(y_5)+P(y_5) = x'+Q(y_8)-P(y_7)-Q(y_5)+P(y_5)-Q(y_4)+P(y_3)+Q(y_1)-P(y_1).$$

But for $\mathbf{y} \in S'''$, we have that

$$\begin{aligned}
&Q(y_8) - P(y_7) - Q(y_5) + P(y_5) - Q(y_4) + P(y_3) + Q(y_1) - P(y_1)\\
&= Q(y_8) - Q(y_7) - Q(y_5) + Q(y_5) - Q(y_4) + Q(y_3) + Q(y_1) - Q(y_1)\\
&= Q(y_8) - Q(y_7) - Q(y_4) + Q(y_3)\\
&= R_{P,Q}(\mathbf{y})
\end{aligned}$$

by using the third condition of S'''. So the average in (†) is the same as the average of $f_A(x')f_A(x' + R_{P,Q}(\mathbf{y}))$, which is precisely $\Lambda'_{P,Q}(f_A, f_A)$.

The result of this repeated application of the Cauchy–Schwarz lemma is summarized in the following bound.

Proposition 4.6. *We have that* $\Lambda_{P,Q}(\mathbb{1}_A, \mathbb{1}_A, f_A) \leq \alpha^{11/8} q^{-1/2} |S'''|^{1/8} \Lambda'_{P,Q}(f_A, f_A)^{1/8}$.

Remark 4.7. There are many places where we could have made different choices about how to apply the Cauchy–Schwarz lemma, even if we restrict ourselves to only taking out the same function in the first factor. For example, at the very first application, we can naively take $k = 2$ as the first two functions in the trilinear average are both $\mathbb{1}_A$. This might seem like an appealing thing to do, because taking higher values of k results in less terms in the product. If we do this, then the bound reduces to a power-saving bound on the bilinear average

$$\mathop{\mathbb{E}}_{x\in\mathbb{F}_q} \mathop{\mathbb{E}}_{\substack{y_1,y_2\in\mathbb{F}_q\\ P(y_1)=P(y_2)}} f_A(x)f_A(x + Q(y_2) - Q(y_1)).$$

By a similar calculation as was done to prove Lemma 4.3, this is equal to

$$\sum_{\psi\in\widehat{\mathbb{F}_q}\setminus\{\psi_0\}} |\widehat{\mathbb{1}_A}(\psi)|^2 \mathop{\mathbb{E}}_{\substack{y_1,y_2\in\mathbb{F}_q\\ P(y_1)=P(y_2)}} \psi(Q(y_2) - Q(y_1)).$$

But we have no hope of proving strong bounds for this quantity in general. If Q is a function of P (for example in the nonlinear Roth configuration of $P(y) = y$ and $Q(y) = y^2$), then this inner average is always 1, so this is exactly equal to $\alpha - \alpha^2$.

Another consideration is if we take all functions in the first factor, i.e. use $\ell = 0$. As an example, one can try this in the last application of the Cauchy–Schwarz lemma, taking $k = 4$ instead. The resulting variety $\widetilde{S'''}$ in $\mathbb{F}_q^8$ we need to consider has 5 polynomial constraints, so its dimension is at least 3. But the upper bound on $\Lambda_{P,Q}$ that we get is $\alpha^{5/4} q^{-3/8} |\widetilde{S'''}|^{1/8}$, so we cannot get a power-saving bound in general.

Now to show Theorem 4.1, it suffices to show bounds on $|S'''|$ and $\Lambda'_{P,Q}(f_A, f_A)$.

4.4. **Bounds from algebraic geometry.** The following two bounds, when plugged in to Proposition 4.6, give the error estimate needed to show Theorem 4.1:

- $|S'''| \lesssim q^4$
- $\Lambda'_{P,Q}(f_A, f_A) \lesssim \alpha q^{-1/2}$

Both of these bounds are consequences of an argument using algebraic geometry. Let us discuss what results we will need.

For any field $\mathbb{F}$, define the two varieties

$$V_{P,Q}(\mathbb{F}) \coloneqq \left\{ \mathbf{y} \in \mathbb{F}^8 \,:\, \begin{array}{c} P(y_2) - P(y_1) = P(y_4) - P(y_3) \\ P(y_6) - P(y_5) = P(y_8) - P(y_7) \\ Q(y_3) - P(y_3) - Q(y_1) + P(y_1) = Q(y_7) - P(y_7) - Q(y_5) + P(y_5) \\ Q(y_2) - Q(y_1) = Q(y_6) - Q(y_5) \end{array} \right\}$$

and

$$W_{P,Q}(\mathbb{F}) \coloneqq \{(\mathbf{y}_1, \mathbf{y}_2) \in V_{P,Q}(\mathbb{F}) \times V_{P,Q}(\mathbb{F}) \,:\, R_{P,Q}(\mathbf{y}_1) = R_{P,Q}(\mathbf{y}_2)\}.$$

Note that $S''' = V_{P,Q}(\mathbb{F}_q)$. The necessary bound will be a control on the dimension of these two varieties.

For technical reasons, we now assume that $\deg P \le \deg Q$, and if $\deg P = \deg Q$ then the leading coefficients of P and Q are different. The former condition is possible by swapping P and Q if necessary. For the latter, we have a correspondence between progressions $x, x+P(y), x+Q(y)$ and progressions $\widetilde{x}, \widetilde{x}+P(y)-Q(y), \widetilde{x}-Q(y)$ by letting $\widetilde{x} = x+Q(y)$. So if the leading terms of P and Q are the same, we can replace (P, Q) by $(P-Q, -Q)$ which now has unequal degrees.

Proposition 4.8. *If* $\operatorname{char} \mathbb{F}_q$ *is large enough (in terms of* P *and* Q*), then* $\dim V_{P,Q}(\overline{\mathbb{F}_q}) \le 4$ *and* $\dim W_{P,Q}(\overline{\mathbb{F}_q}) \le 7$.

Peluse proves this via a lengthy argument using the **graded lexicographical order**. The proof of the bound for $V_{P,Q}$ is much easier than the proof of the bound for $W_{P,Q}$. For the proof, see [Pel18, Section 3].

We can feed this dimension bound into a form of the Lang–Weil bound from [LW54].

Proposition 4.9 (Schwartz–Zippel type Lang–Weil bound)**.** *If* V *is an affine variety over* $\overline{\mathbb{F}_q}$ *of complexity*[2] M*, then* $|V(\mathbb{F}_q)| \lesssim_M q^{\dim V}$.

This immediately implies the claimed bound on $|S'''| = |V_{P,Q}(\mathbb{F}_q)|$.

Corollary 4.10. *We have that* $|S'''| \lesssim q^4$.

The Lang–Weil bound also gives that $|W_{P,Q}(\mathbb{F}_q)| \lesssim q^7$. To deduce the bound on $\Lambda'_{P,Q}(f_A, f_A)$, we can use a special case of a result of Kowalski [Kow07], as stated in [Pel18, Proposition 3.1].

Proposition 4.11. *Let* $V \subseteq \mathbb{Z}^n$ *be an affine subscheme, and* ψ *be a nontrivial additive character of* $\mathbb{F}_q$*, and* $F \in \mathbb{Z}[V]$ *be a regular function on* V*. If* $\operatorname{char} \mathbb{F}_q$ *is large enough (in terms of* V *and* $\deg F$*) and* $|F^{-1}(a)|/|V(\mathbb{F}_q)|$ *for* $a \in \mathbb{F}_q$ *is bounded in terms of* V*, then*

$$\sum_{\mathbf{y} \in V(\mathbb{F}_q)} \psi(F(\mathbf{y})) \lesssim_{V, \deg F} q^{\dim V(\overline{\mathbb{F}_q}) - 1/2}.$$

Using this, we can prove a character sum bound.

Lemma 4.12. *Let* ψ *be a nontrivial additive character of* $\mathbb{F}_q$*. If* $\operatorname{char} \mathbb{F}_q$ *is large enough (in terms of* P *and* Q*), then*

$$\mathop{\mathbb{E}}_{\mathbf{y} \in S'''} \psi(R_{P,Q}(\mathbf{y})) \lesssim q^{-1/2}.$$

Before we jump into the proof, let us record a counting version of the Cauchy–Schwarz lemma (Lemma 4.4), which is actually just a special case of the original lemma.

Lemma 4.13 (Cauchy–Schwarz lemma, counting version)**.** *For* S *and* S' *as in the setting of the Cauchy–Schwarz lemma, we have that* $|S'| \ge q^{-(k-1)}|S|^2$.

[2] If $V = \{\mathbf{y} \in \overline{\mathbb{F}_q}^n \,:\, P_1(\mathbf{y}) = \cdots = P_k(\mathbf{y}) = 0\}$ for some polynomials $P_1, \dots, P_k$, we say that the **complexity** of V is $\max\{n, k, \deg P_1, \dots, \deg P_k\}$.

Proof. Let all of the functions in the Cauchy–Schwarz lemma be the constant 1 function. □

Proof of Lemma 4.12. We wish to apply Proposition 4.11 with $V = V_{P,Q}(\overline{\mathbb{F}_q})$ and $F = R_{P,Q}$. It suffices to check that $|R_{P,Q}^{-1}(a)|/|S'''|$ is bounded.

First, we bound $|R_{P,Q}^{-1}(a)|$. By looking at fibres, we have that

$$|W_{P,Q}(\mathbb{F}_q)| = \sum_{a \in \mathbb{F}_q} |R_{P,Q}^{-1}(a)|^2,$$

so this sum is asymptotically at most q^7. It follows that $\dim R_{P,Q}^{-1}(a) \leq 3$ for all $a \in \mathbb{F}_q$. So by the Lang–Weil bound, we have that $|R_{P,Q}^{-1}(a)| \lesssim q^3$.

Now, we show that S''' is not too small. Our set S''' was constructed by repeatedly applying the Cauchy–Schwarz lemma. Recalling the proof of Proposition 4.6 and tracking the values of k used, we can use the counting version of the Cauchy–Schwarz lemma to deduce that

$$|S'''| \geq q^{-2}|S''|^2 \geq q^{-2} \cdot q^{-2}|\mathbb{F}_q \times \mathbb{F}_q|^4 = q^4.$$

So we get that $|R_{P,Q}^{-1}(a)|/|S'''| \lesssim q^{-1} \lesssim 1$. Then we can apply Proposition 4.11 to deduce that

$$\sum_{\mathbf{y} \in S'''} \psi(R_{P,Q}(\mathbf{y})) \lesssim q^{7/2}.$$

Dividing by $|S'''| \geq q^4$ gives the desired bound. □

Now we can prove the desired bound on $\Lambda'_{P,Q}(f_A, f_A)$.

Proposition 4.14. *We have that* $\Lambda'_{P,Q}(f_A, f_A) \lesssim \alpha q^{-1/2}$.

Proof. Expand f_A using Fourier coefficients. By a similar calculation as was done to prove Lemma 4.3, we have that

$$\Lambda'_{P,Q}(f_A, f_A) = \sum_{\psi \in \widehat{\mathbb{F}_q} \setminus \{\psi_0\}} |\widehat{\mathbb{1}_A}(\psi)|^2 \mathop{\mathbb{E}}_{\mathbf{y} \in S'''} \psi(R_{P,Q}(\mathbf{y})).$$

By Lemma 4.12, we can bound each of these averages asymptotically by $q^{-1/2}$, so

$$|\Lambda'_{P,Q}(f_A, f_A)| \lesssim q^{-1/2} \sum_{\psi \in \widehat{\mathbb{F}_q} \setminus \{\psi_0\}} |\widehat{\mathbb{1}_A}(\psi)|^2 = q^{-1/2}(\alpha - \alpha^2)$$

as desired. □

With that, we can plug in Corollary 4.10 and Proposition 4.14 to Proposition 4.6 to deduce that $\Lambda_{P,Q}(\mathbb{1}_A, \mathbb{1}_A, f_A) \lesssim \alpha^{3/2} q^{-1/16}$. Combined with Lemma 4.3 and Proposition 4.2, we have that

$$\Lambda_{P,Q}(\mathbb{1}_A, \mathbb{1}_A, \mathbb{1}_A) = \alpha^3 + O(\alpha^{3/2} q^{-1/16}),$$

which implies Theorem 4.1 and thus Theorem 1.2, the main result of Peluse for three-term polynomial progressions over a finite field.

References

[BG08] Richard Beigel and William Gasarch, *Square-Difference-Free Sets of Size* $\Omega(n^{0.7334\cdots})$, arXiv preprint arXiv:0804.4892 (2008).

[BL96] V. Bergelson and A. Leibman, *Polynomial extensions of van der Waerden's and Szemerédi's theorems*, J. Amer. Math. Soc. **9** (1996), no. 3, 725–753.

[BLL08] V. Bergelson, A. Leibman, and E. Lesigne, *Intersective polynomials and the polynomial Szemerédi theorem*, Adv. Math. **219** (2008), no. 1, 369–388.

[CL07] Ernest S. Croot, III and Vsevolod F. Lev, *Open problems in additive combinatorics*, Additive combinatorics, CRM Proc. Lecture Notes, vol. 43, Amer. Math. Soc., Providence, RI, 2007, pp. 207–233.

[Coh88] Stephen D. Cohen, *Clique numbers of Paley graphs*, Quaestiones Math. **11** (1988), no. 2, 225–231.

[Fab93] J. Fabrykowski, *On maximal residue difference sets modulo p*, Canad. Math. Bull. **36** (1993), no. 2, 144–146.

[Fur77] Harry Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math. **31** (1977), 204–256.

[Gre05] Ben Green, *Finite field models in additive combinatorics*, Surveys in combinatorics 2005, London Math. Soc. Lecture Note Ser., vol. 327, Cambridge Univ. Press, Cambridge, 2005, pp. 1–27.

[GS24] Ben Green and Mehtaab Sawhney, *New bounds for the Furstenberg–Sárközy theorem*, arXiv preprint arXiv:2411.17448 (2024).

[HP21] Brandon Hanson and Giorgis Petridis, *Refined estimates concerning sumsets contained in the roots of unity*, Proc. Lond. Math. Soc. (3) **122** (2021), no. 3, 353–358.

[Kow07] E. Kowalski, *Exponential sums over definable subsets of finite fields*, Israel J. Math. **160** (2007), 219–251.

[LW54] Serge Lang and André Weil, *Number of points of varieties in finite fields*, Amer. J. Math. **76** (1954), 819–827.

[Pel18] Sarah Peluse, *Three-term polynomial progressions in subsets of finite fields*, Israel J. Math. **228** (2018), no. 1, 379–405.

[Pel19] ———, *On the polynomial Szemerédi theorem in finite fields*, Duke Math. J. **168** (2019), no. 5, 749–774.

[Pel20] ———, *Bounds for sets with no polynomial progressions*, Forum Math. Pi **8** (2020), e16, 55.

[Pel24] ———, *Finite field models in arithmetic combinatorics—twenty years on*, Surveys in combinatorics 2024, London Math. Soc. Lecture Note Ser., vol. 493, Cambridge Univ. Press, Cambridge, 2024, pp. 159–199.

[PP24] Sarah Peluse and Sean Prendiville, *Quantitative bounds in the nonlinear Roth theorem*, Invent. Math. **238** (2024), no. 3, 865–903.

[Pre17] Sean Prendiville, *Quantitative bounds in the polynomial Szemerédi theorem: the homogeneous case*, Discrete Anal. (2017), Paper No. 5, 34.

[Ric19] Alex Rice, *A maximal extension of the best-known bounds for the Furstenberg-Sárközy theorem*, Acta Arith. **187** (2019), no. 1, 1–41.

[Rot52] Klaus Roth, *Sur quelques ensembles d'entiers*, C. R. Acad. Sci. Paris **234** (1952), 388–390.

[So78] A. Sárk″ ozy, *On difference sets of sequences of integers. I*, Acta Math. Acad. Sci. Hungar. **31** (1978), no. 1-2, 125–149.

[Sze75] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245.

[Vin54] I. M. Vinogradov, *Elements of number theory*, Dover Publications, Inc., New York, 1954, Translated by S. Kravetz.

[Wol15] J. Wolf, *Finite field models in arithmetic combinatorics—ten years on*, Finite Fields Appl. **32** (2015), 233–274.