On polynomial progressions in finite fields

J. Bitz, J. Echevarría Cuesta, and E. Kilgore

August 31, 2018

Abstract

We extend to a two-dimensional case the work of Peluse [1] concerning three-term polynomial progressions in finite fields. Specifically, let $A \subset \mathbb{F}_p^2$ and $P, Q : \mathbb{F}_p^2 \to \mathbb{F}_p^2$ be given by two-variate polynomials with single-variate highest-degree terms. We prove that any such A containing no non-trivial progressions of the form $x, x + P(\mu, \nu), x + Q(\mu, \nu)$ is of size $O(p^{2-\frac{1}{24}})$. We also explore cases when P and Q are instead single-variate rational functions lying in $\mathbb{F}_p(x)$, and achieve an exponential bound conditional on a conjecture requiring sophisticated algebraic geometry to verify.

1 Introduction

Intuitively, if n people attending a conference start befriending each other, eventually there must be at least three people that are friends with each other, even before everyone has become friends. The Mantel–Turán Theorem proves this intuition by stating that any graph on n vertices with at least $\lfloor n^2/4 \rfloor + 1$ edges must contain a triangle.

In the same vein, a central question in combinatorial number theory is to wonder what conditions on a subset of the integers would guarantee that it contains an arithmetic progression. Just as with the friendships above, it does not seem irrational to conjecture, as Erdős and Turan did in 1936, that all we need is for the subset to be big enough. To be able to talk about the size of a subset of the integers in any precise way, however, we need the following definition.

Definition 1.1. The **upper density** of a set $A \subset \mathbb{Z}$ is defined as

$$\limsup_{N \to \infty} \frac{|A \cap [1, N]|}{N}.$$

Remark 1.2. The lower density can be defined in an analogous fashion. We note the need to use the limit superior in the definition above as opposed to a simple limit because lower and upper densities do not always agree (for instance, the set of numbers whose decimal expansion begins with the digit 1 has lower density 1/9 and upper density 5/9). When upper and lower densities agree, the subset is said to have asymptotic density.

Erdős and Turan hence conjectured that any set of integers with positive upper density contains arbitrarily long arithmetic progressions, which started a cascade of results throughout the 20th century. In 1953, Roth proved the conjecture for progressions of lengh three

using analytic methods. Sixteen years later Szemerédi extended the result to progressions of length four using a sophisticated combinatorial argument. This was only a preview of 1975, however, the year in which Szemerédi resolved Erdős and Turan's 1936 conjecture. Although these are some highlights, there were many other intermediary results. In particular, analogous questions were raised about polynomial progressions. For instance, in 1998, Bergelson and Leibman showed that, if $P, Q \in \mathbb{Z}[y]$ with $P(0) = Q(0) = 0$, then any subset of the integers with positive upper density contains a nontrivial polynomial progression $x, x + P(y), x + Q(y)$.

At the turn of the 21st century, Green and Tao showed that the sequence of prime numbers contains arbitrarily long arithmeic progressions. Since the primes have asymptotic density zero, Erdős and Turan's conjecture did not apply directly. Nevertheless, they did use Szemerédi's Theorem, along with a transference principle that extends the theorem to subsets of the integers which are pseudorandom in a sense that they made precise. Therefore, the question still remains of whether arbitrarily long arithmetic progressions exist in the primes simply because there is enough of them or because of their pseudorandomness. In trying to answer questions along this same line, the attention has hence recently been redirected towards finite subsets of the integers. Indeed, proving bounds here is stronger than in the asymptotics and might help provide a proof of Green and Tao's theorem that only uses density arguments. Roth's original work actually showed that

$$\frac{|A \cap |1, N]|}{N} = \Omega\left(\frac{1}{\log\log N}\right)$$

is sufficient to guarantee the existence of a three-term progression in A. Since then, the best bound shown to suffice has been found by Bloom in 2016, namely,

$$\frac{|A \cap |1, N]|}{N} = \Omega\left(\frac{(\log\log N)^4}{\log N}\right).$$

In parallel, people have started to ask similar questions about polynomial progressions. However, these sort of questions are extremely hard in the integers (for instance, if a subset $A \subseteq [1, N]$ contains a sequence $x, x+y, x+y^2$, then $y \leq \sqrt{N}$, which greatly limits the values that y can take on). People have hence begun to study progressions in finite fields instead. The first ones to provide quantitative bounds for the polynomial Szemerédi Theorem in finite fields were Bourgain and Chang. In 2017, they showed that if A is a subset of $\mathbb{F}_p$ which does not contain any nontrivial progressions $x, x + y, x + y^2$, then

$$|A| = O(p^{14/15}).$$

At the end of their paper, the authors instigated further research into the topic by asking three questions. The first one asked whether similar bounds could be found for general progressions of the form $x, x + P(y), x + Q(y)$ where $P, Q \in \mathbb{Z}[y]$ are linearly independent and $P(0) = Q(0) = 0$. This was answered in the affirmative by Peluse:

Theorem 1.3 (Peluse)**.** *Let $P, Q \in \mathbb{Z}[y]$ be two linearly independent polynomials with $P(0) = Q(0) = 0$. There exists a constant $c_{P,Q} > 0$ depending only on P and Q such that if the characteristic of $\mathbb{F}_q$ is at least $c_{P,Q}$, then any $A \subset \mathbb{F}_q$ containing no nontrivial progression*

$$x, x + P(y), x + Q(y),\ y \neq 0,$$

satisfies

$$|A| = O(q^{1-1/24}).$$

Under Peluse's guidance, our objective was to answer another of Bourgain and Chang's questions which asked about similar progressions with rational functions instead. Trying to adapt Peluse's argument to this more general case, we decided instead to tackle the easier question of proving that any subset of $\mathbb{F}_p$ that lacks the configuration

$$x, x + y, x + \frac{1}{y^2}$$

has size $O(p^{1-c})$ for some $c > 0$. As described below, we got very far in this direction, eventually reducing our proof to a bound on a sum of products of sums that are hybrids between Salié and hyper-Kloosterman sums. Unfortunately, we realized that finding a bound on such object was out of our reach, so we opted to investigate what non-trivial generalizations we could make of Peluse's argument in the 2-dimensional case. This report hence explores the current work that we have done in that direction, as well as an account of some of the key material that we have had to learn along the way to tackle these questions. The main theorem that we want to prove is hence

Theorem 1.4. *Let $P, Q : \mathbb{F}_p^2 \to \mathbb{F}_p^2$ be given by $P = (P_1, P_2)$ and $Q = (Q_1, Q_2)$ where, for each $i = 1, 2$, the polynomials $P_i, Q_i \in \mathbb{F}_p[\mu, \nu]$ are linearly independent, have single-variate highest-degree terms with different degrees for a given i, where both P_i, Q_i have the same variable in their leading term for fixed i, and $P_i(0) = Q_i(0) = 0$. There exists a constant $c_{P,Q} > 0$ depending only on P and Q such that if $p \geq c_{P,Q}$, then any $A \subset \mathbb{F}_p^2$ containing no nontrivial progressions of the form*

$$x, x + P(\mu, \nu), x + Q(\mu, \nu),\ \mu, \nu \neq 0$$

satisfies

$$|A| = O\left(p^{2-\frac{1}{24}}\right).$$

2 Discrete Fourier Analysis

In order to tackle these problems we first need to develop a few tools to help us work with the objects we will consider. The first of these is Fourier analysis over finite fields. In order to understand this we first define an additive character on $\mathbb{F}_p$.

Definition 2.1. Let $\mathbb{F}$ be a field. An **additive character** of $\mathbb{F}$ is a group homomorphism from the additive group of $\mathbb{F}$ to the multiplicative group $\mathbb{C}^\times$. The set of additive characters of $\mathbb{F}$ is labeled $\hat{\mathbb{F}}$.

Of course, this is a special case of a character, which can be defined on any group.

Definition 2.2. Let G be a group. A **character** of G is a homomorphism $G \to \mathbb{C}^\times$.

To help us in describing these explicitly we will introduce a bit of notation. For any prime p, let

$$e_p(x) := e^{\frac{2\pi i x}{p}}.$$

Since $\mathbb{F}_p$ is generated by any non-zero element we can see that

Lemma 2.3. *The additive characters of $\mathbb{F}_p$ are exactly the functions*

$$\chi_n(x) = e_p(nx), \text{ where } n \in \{0, \ldots, p-1\}.$$

Proof. The additive group of $\mathbb{F}_p$ is just the cyclic group $\mathbb{Z}/p\mathbb{Z}$. Any non-zero element in this group has order p, so any additive character $\chi : \mathbb{F}_p \to \mathbb{C}^\times$ must send all elements of $\mathbb{F}_p$ to p-th roots of unity in $\mathbb{C}$. Furthermore, such a homomorphism is entirely determined by the value it takes on $1 \in \mathbb{F}_p$. Therefore, there is exactly one such homomorphism for every p-th root of unity, as desired. □

In fact, we may extend this definition to direct products of fields (since these remain additive groups) and we have the secondary result:

Corollary 2.3.1. *The additive characters over $\mathbb{F}_p^n$ are the functions*

$$\chi_{m_1,\ldots,m_n}(x) = e_p\left(\sum_{j=1}^{n} m_j x_j\right), \text{ where } m_j \in 0, \ldots, p-1, \text{ and } x = (x_1, \ldots, x_n).$$

For the sake of brevity, we have chosen not to include any of the proofs for the results on Fourier analysis that follow, but rather simply state them as facts. For proofs and more detailed discussion see [2]. For convenience, we will introduce

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

where S is some set and $f : S \to \mathbb{C}$. We can obviously interpret this as the expected value of f over the set S. We will also denote by $V(\mathbb{F}_p^n; \mathbb{C})$ the set of complex valued functions on $\mathbb{F}_p^n$.

Remark 2.4. The space $V(\mathbb{F}_p^n; \mathbb{C})$ defined above is a vector space. Moreover, it is an inner product space with inner product

$$\langle f, g \rangle = \mathbb{E}_{x \in \mathbb{F}_p^n} \overline{f(x)} g(x).$$

With this we are ready to begin looking at our Fourier transform. The most important result is

Theorem 2.5. *The additive characters $\chi \in \widehat{\mathbb{F}}_p^n$ form an orthonormal basis of $V(\mathbb{F}_p^n; \mathbb{C})$ under the inner product defined above. We denote the map from characters $\chi_{m_1,\ldots,m_n}$ to their basis coefficients corresponding to a particular f by $\hat{f} : \widehat{\mathbb{F}}_p^n \to \mathbb{C}$.*

In particular, this immediately gives us

Corollary 2.5.1 (Fourier Inversion)**.** *For any $f \in V(\mathbb{F}_p^n; \mathbb{C})$ we can write*

$$f(x) = \sum_{\chi \in \widehat{\mathbb{F}}_p^n} \hat{f}(\chi) \chi(x).$$

So finally we arrive at

Definition 2.6. Let $f \in V(\mathbb{F}_p^n; \mathbb{C})$ a function. The **Fourier transform** of f is defined as the map $\hat{f} : \widehat{\mathbb{F}}_p^n \to \mathbb{C}$ mapping a character $\chi \in \widehat{\mathbb{F}}_p^n$ to the coefficient of χ in the basis expansion above.

From this definition, and our earlier result, we immediately obtain

Lemma 2.7. *The Fourier transform is given by*

$$\hat{f}(\chi) = \mathbb{E}_{x \in \mathbb{F}_p^n} f(x)\overline{\chi}(x),$$

where $\overline{\chi}$ denotes the complex conjugate of χ.

Having defined the Fourier transform, we also will need a few basic results for later on, which will look rather familiar to those who are acquainted with Fourier transforms in an analysis context

Theorem 2.8 (Plancherel)**.** *For any $f, g \in V(\mathbb{F}_p^n; \mathbb{C})$ we have*

$$\mathbb{E}_{x \in \mathbb{F}_p^n} \overline{f(x)} g(x) = \sum_{\chi \in \widehat{\mathbb{F}}_p^n} \overline{\hat{f}(\chi)} \hat{g}(\chi).$$

which gives us

Corollary 2.8.1 (Parseval)**.** *If $f \in V(\mathbb{F}_p^n; \mathbb{C})$ then*

$$\|f\|^2 := \mathbb{E}_{x \in \mathbb{F}_p^n} \overline{f(x)} f(x) = \sum_{\chi \in \widehat{\mathbb{F}}_p^n} |\hat{f}(\chi)|^2.$$

3 Progress in Rational Case

In this section we seek to record an overview of the advances that we made in the case of progressions of the form $x, x+y, x+\frac{1}{y^2}$ in $\mathbb{F}_p$. For any $f_1, f_2 : \mathbb{F}_p \to \mathbb{C}$ such that $\|f_1\|, \|f_2\| \leq 1$, define

$$F(x) = \frac{1}{p} \sum_{y \in \mathbb{F}_p^*} f_1(x+y) f_2\left(x + \frac{1}{y^2}\right)$$

and let $\mathbb{E}[f_1] := \mathbb{E}_{x \in \mathbb{F}_p} f_1$. We want to show that

$$\|F - \mathbb{E}[f_1] \cdot \mathbb{E}[f_2]\| \leq c p^{-\delta} \|f_1\| \cdot \|f_2\|$$

for some $\delta > 0$. Expanding f_1, f_2 in Fourier sums gives

$$\begin{aligned}
F(x) &= \frac{1}{p} \sum_{y \in \mathbb{F}_p^*} \sum_{n_1, n_2 \in \mathbb{F}_p} \hat{f}_1(n_1) \hat{f}_2(n_2) e_p(n_1(x+y)) e_p\left(n_2\left(x + \frac{1}{y^2}\right)\right) \\
&= \frac{1}{p} \sum_{y \in \mathbb{F}_p^*} \sum_{n_1, n_2 \in \mathbb{F}_p} \hat{f}_1(n_1) \hat{f}_2(n_2) e_p\left(n_1 y + n_2 \frac{1}{y^2}\right) e_p((n_1+n_2)x) \\
&= \sum_{n_1, n_2 \in \mathbb{F}_p} \hat{f}_1(n_1) \hat{f}_2(n_2) c_{n_1, n_2} e_p((n_1+n_2)x)
\end{aligned}$$

with

$$c_{n_1,n_2} = \frac{1}{p}\sum_{y\in\mathbb{F}_p^*} e_p\left(n_1 y + n_2\frac{1}{y^2}\right) = \begin{cases} (p-1)/p & \text{if } n_1 = n_2 = 0 \\ -1/p & \text{if } n_1 \neq 0, n_2 = 0 \\ \frac{1}{p}\sum_{y\in\mathbb{F}_p^*} e_p\left(n_1 y + n_2\frac{1}{y^2}\right) & \text{if } n_2 \neq 0. \end{cases}$$

It follows that

$$\begin{aligned} F(x) &= \frac{p-1}{p}\hat{f}_1(0)\hat{f}_2(0) - \frac{1}{p}\sum_{n_1\in\mathbb{F}_p^*}\hat{f}_1(n_1)\hat{f}_2(0)e_p(n_1x) + \sum_{n_1\in\mathbb{F}_p}\sum_{n_2\in\mathbb{F}_p^*}\hat{f}_1(n_1)\hat{f}_2(n_2)c_{n_1,n_2}e_p((n_1+n_2)x) \\ &= \hat{f}_1(0)\hat{f}_2(0) - \frac{1}{p}\sum_{n_1\in\mathbb{F}_p}\hat{f}_1(n_1)\hat{f}_2(0)e_p(n_1x) + \sum_{n_1\in\mathbb{F}_p}\sum_{n_2\in\mathbb{F}_p^*}\hat{f}_1(n_1)\hat{f}_2(n_2)c_{n_1,n_2}e_p((n_1+n_2)x) \\ &= \mathbb{E}[f_1]\cdot\mathbb{E}[f_2] - \frac{1}{p}\mathbb{E}[f_2]f_1(x) + \sum_{n_1\in\mathbb{F}_p}\sum_{n_2\in\mathbb{F}_p^*}\hat{f}_1(n_1)\hat{f}_2(n_2)c_{n_1,n_2}e_p((n_1+n_2)x) \\ &= \mathbb{E}[f_1]\cdot\mathbb{E}[f_2] - \frac{1}{p}\mathbb{E}[f_2]f_1(x) + \sum_{s\in\mathbb{F}_p}e_p(sx)\sum_{n\in\mathbb{F}_p^*}\hat{f}_1(s-n)\hat{f}_2(n)c_{s-n,n} \end{aligned}$$

and hence

$$F(x) - \mathbb{E}[f_1]\cdot\mathbb{E}[f_2] = -\frac{1}{p}\mathbb{E}[f_2]f_1(x) + \sum_{s\in\mathbb{F}_p}e_p(sx)\sum_{n\in\mathbb{F}_p^*}\hat{f}_1(s-n)\hat{f}_2(n)c_{s-n,n}.$$

Since we have the restrictions $\|f_1\|, \|f_2\| \leq 1$, in the asymptotics, as p grows, we obtain

$$\|F(x) - \mathbb{E}[f_1]\cdot\mathbb{E}[f_2]\| = O\left(\left\|\sum_{s\in\mathbb{F}_p}e_p(sx)\sum_{n\in\mathbb{F}_p^*}\hat{f}_1(s-n)\hat{f}_2(n)c_{s-n,n}\right\|\right)$$

and by Parseval,

$$\|F(x) - \mathbb{E}[f_1]\cdot\mathbb{E}[f_2]\| = O\left(\frac{1}{\sqrt{p}}\left(\sum_{s\in\mathbb{F}_p}\left|\sum_{n\in\mathbb{F}_p}\hat{f}_1(s-n)\hat{f}_2(n)K(s-n,n)\right|^2\right)^{1/2}\right)$$

with $K(x,y) := \sqrt{p}c_{x,y}$ if $y \neq 0$ and 0 otherwise. Using quadratic Gauss sum evaluation, we notice that we can actually write

$$K(x,y) = \begin{cases} \frac{1}{\sqrt{p}}\sum_{z\in\mathbb{F}_p^*} e_p\left(xz + y\frac{1}{z^2}\right) & \text{if } x \neq 0, y \neq 0 \\ \sigma_p - \frac{1}{\sqrt{p}} & \text{if } x = 0, y \neq 0 \\ 0 & \text{if } y = 0 \end{cases}$$

where

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

Now let $x, y \neq 0$. By the change of variables $z \mapsto z/x$ we have

$$K(x,y) = \frac{1}{\sqrt{p}} \sum_{z \in \mathbb{F}_p^*} e_p\left(z + a\frac{1}{z^2}\right),$$

where $a = \frac{x^2}{y}$. We notice that this is very close to a Kloosterman sum. Moreover, by making the change of variables $z \mapsto 2z/x$ instead, we can also apply Theorem 2 of [3] to write

$$K(x,y) = \frac{1}{\sqrt{p}} \sum_{b,t \in \mathbb{F}_p^*} \left(\frac{b}{p}\right) e_p\left(b + t + a\frac{1}{bt}\right),$$

which we recognize as some sort of hybrid between a Salié sum and the high-dimensional generalization of Kloosterman sums, the hyper-Kloosterman sum (note that we have introduced the Legendre symbol). Going back to the change $z \mapsto z/x$, since $a \neq 0$, Theorem 3 of [4] thus tells us that

$$\sqrt{p} \sum_{z \in \mathbb{F}_p^*} e_p\left(z + a\frac{1}{z^2}\right) = O(p).$$

Combining this result with our previous observations, we thus have

$$K(x,y) = O(1)$$

for all $x, y \in \mathbb{F}_p$.

Having shown this, we thus find ourselves in the same territory as Bourgain and Chang's argument. As they show in [5], we have

$$\frac{1}{\sqrt{p}} \left(\sum_{s \in \mathbb{F}_p} \left| \sum_{n \in \mathbb{F}_p} \hat{f}_1(s-n) \hat{f}_2(n) K(s-n, n) \right|^2 \right)^{1/2} = O(\Omega^{1/5} \|f_1\| \cdot \|f_2\|)$$

where

$$\begin{aligned} \Omega^2 = \sum_{x,y,s,s',u} & K(x, s-x)\overline{K(x+u, s-x)}\,\overline{K(x, s'-x)}K(x+u, s'-x) \\ & \times \overline{K(y, s-y)}K(y+u, s-y)K(y, s'-y)\overline{K(s+u, s'-y)}. \end{aligned}$$

When studying progressions of the form $x, x+y, x+y^2$, their method is very specific to their case, so we cannot recover anything. On the other hand, when working on progressions of the form $x, x+y, x+\frac{1}{y}$, Bourgain and Chang rely on results from [6] to bound Ω. They are able to do this because the equivalent of our K function in their proof is a Kloosterman sum, which has been widely studied. Unfortunately, our K function as defined above is not quite a Kloosterman sum, so we cannot invoke Corollary 3.3 in [6]. On the other hand, if

we were able to show that the function $K : \mathbb{F}_p^2 \to \mathbb{C}$ satisfies the requirements of Corollary 1.6 in [6] (i.e., that it is a trace function modulo p of a bountiful sheaf), we would be done. Unfortunately, however, since the material was out of our reach, we had to stop at this stage of the proof, which hence remains incomplete.

4 Background on Varieties and their Dimension

Our approach to the problem of finding polynomial progressions relies heavily on some advanced machinery from algebraic geometry that bounds the size of sets given by the zero locus of sets of polynomials. We will not need to completely understand these results to apply them, but the one unavoidable notion is that of the dimension of a variety. The exposition below is self-contained but admittedly short — for more detail, see Chapter 9 of [7]. We begin with a definition.

Definition 4.1. Let $\mathbb{F}$ be a field, and $S = \{f_\alpha\}$ a collection of polynomials in $\mathbb{F}[X_1, \ldots, X_n]$. The **variety** corresponding to S is the subset

$$V(S) = \{x \in \mathbb{F}^n : f_\alpha(x) = 0 \text{ for all } \alpha\}.$$

Note that, if some $x \in V$ is a zero of all the f_α, it is also a zero of $gf_\alpha + hf_\beta$ for any g and h also in $\mathbb{F}[X_1, \ldots, X_n]$. Therefore, the set of all polynomials vanishing on V is in fact an ideal.

This correspondence between ideals and subsets of affine space is at the heart of algebraic geometry - it allows us to answer geometric problems using commutative algebra, and algebraic problems using geometry. In light of this, if $V \subset \mathbb{F}^n$ is a variety, we will let $I(V)$ denote the ideal given by the set of all polynomials which vanish on V. Likewise, for some ideal $I \subset F[X_1, \ldots, X_n]$, we let $V(I)$ denote the variety given by the set of points on which all polynomials in I vanish.

Remark 4.2. Beware that the operation of taking the variety corresponding to an ideal and the ideal corresponding to a variety are not always inverses. For example, let $J = \langle x^2 \rangle \subset \mathbb{R}[x, y]$. Then, $V(J)$ is the y-axis in $\mathbb{R}^2$, but $I(V(J)) = \langle x \rangle$.

However, by Hilbert's Nullstellensatz, for any polynomial f vanishing on $V(J)$, we have $f^k \in I$ for some $k \in \mathbb{N}$. Therefore, $I(V(J)) = \sqrt{J}$, and the statement $I(V(J)) = J$ if J is a radical ideal.

Varieties become useful to our argument because we may take $\mathbb{F}$ to be a finite field $\mathbb{F}_p$, and express a set of interesting points in $\mathbb{F}_p^n$ as the zeros of certain polynomials generating an ideal I. Our goal will be to determine $|V(I)|$, which is very difficult to compute directly. Instead, we will assign a notion of dimension to $V(I)$, and then utilize already-proven results which link the dimension of a variety over a finite field to the number of points it contains.

The road to a rigorous definition of the dimension of a variety is long and torturous. To begin, we set up some notation. For the rest of this section, F is a field and all polynomials lie in the ring $F[X_1, \ldots, X_n]$. For some $\alpha \in \mathbb{Z}_{\geq 0}^n$, we set

$$X^\alpha := \prod X_i^{\alpha_i}, \quad |\alpha| := \sum \alpha_i.$$

Such an α is called a **multi-index**. We may also define a partial order $\leq$ on the set $\mathbb{Z}_{\geq 0}^n$ of all multi-indices, such that $\alpha \leq \beta$ when $\alpha_i \leq \beta_i$ for all $i \leq n$. The next definition is important enough that we give it its own call-out:

Definition 4.3. A **monomial** is an element in $\mathbb{F}[X_1, \ldots, X_n]$ of the form X^α. An ideal I is said to be a **monomial ideal** if it is generated by monomials.

Example 4.4. 1) $I = \langle x^2, y \rangle \subset \mathbb{R}[x, y]$ is a monomial ideal.

2) $J = \langle x^2 - y, y \rangle$ is a monomial ideal, because it may be written as $J = \langle x^2, y \rangle$.

3) $K = \langle x^2 - y, y + x \rangle$ is not a monomial ideal, because it cannot be written with a generating set of monomials.

Because of this especially nice description, monomial ideals are particularly easy to work with. Our quality of life is improved even more by the following theorem.

Theorem 4.5 (Hilbert's Basis Theorem)**.** *Let R be a Noetherian domain. Then, $R[X]$ is also Noetherian.*

Repeated application of this theorem shows that $\mathbb{F}[X_1, \ldots, X_n]$ is Noetherian, and hence *all* of its ideals are finitely generated (not just the monomial ones). Therefore, we may write any monomial ideal as being generated by a finite number of terms $\langle X^{\alpha^{(1)}}, \ldots, X^{\alpha^{(m)}} \rangle$. Given this, we would now like to characterize $V(I)$ for monomial ideals.

Proposition 4.6. *Let I be a monomial ideal. For any $S \subset [1, \ldots, n]$, define*

$$H_S := \{z \in \mathbb{F}^n : z_i = 0 \text{ for all } i \in S\},$$

the subspace of all points in F^n where all coordinates lying in S are 0. Then, for some $S_1, \ldots, S_k$ all subsets of $[1, \ldots, n]$, we have

$$V(I) = \bigcup_{i=0}^{k} H_{S_i}.$$

Proof. Suppose that some monomial ideal I has a single generator, i.e. $I = \langle X^\alpha \rangle$. Then, we have

$$V(I) = \bigcup_{\substack{i \leq n \\ \alpha_i > 0}} H_i. \tag{1}$$

In other words, X^α will vanish so long as any one of the terms appearing with non-zero degree are zero. Now, if I has m generating monomials, $V(I)$ will be an intersection of terms resembling Equation 1, i.e.

$$V(I) = \bigcap_{i=0}^{m} \bigcup_{j \in S_i} H_j, \tag{2}$$

where S_i is the set of all variable indices appearing in generator i. Distributing the intersections over the unions and noting that $H_S \cap H_T = H_{S \cap T}$ gives us the desired result. □

So, we may express the variety corresponding to a monomial ideal as a union of subspaces H_{S_i}, each with dimension (in the standard linear-algebraic sense) $n - |S_i|$. We may then make our first partial definition of the dimension of a variety in the case of monomial ideals.

Definition 4.7. Let I be a monomial ideal. Then, with

$$V(I) = \bigcup_{i=0}^{k} H_{S_i},$$

the **dimension** of I is

$$\dim I := \max_i \{\dim H_{S_i}\} = n - \min_i \{|S_i|\}.$$

The following proposition describes an equivalent way of defining dimension in terms of the monomials *not* appearing in a given monomial ideal. This formulation is marginally less concrete, but turns out to have nicer theoretical properties that we will make use of shortly.

Definition 4.8. Let I be a monomial ideal. The **affine Hilbert function** $\mathrm{HF}_I(s)$ is defined as the number of monomial terms of total degree $\leq s$ not contained in I.

Proposition 4.9. *For any monomial ideal I, there exists a constant c_I such that, for all $s > c_I$, $HF_I(s)$ is a polynomial in s of degree* $\dim I$.

In other words, the dimension of monomial ideal corresponds to the growth rate of that ideals complement as we increase the maximum allowed degree of terms.

We would like to extend the notion of affine Hilbert functions to general ideals. The way of doing this is as follows.

Definition 4.10. Let I be an ideal. Let $F[X_1, \ldots, X_n]_{\leq s}$ be the finite dimensional $\mathbb{F}$-vector space of polynomials with degree $\leq s$, and let $I_{\leq s}$ be the $\mathbb{F}$-vector space of polynomials in I of degree $\leq s$. The **affine Hilbert function** of I is defined as

$$\mathrm{HF}_I := \dim \mathbb{F}[X_1, \ldots, X_n]_{\leq s}/I_s,$$

where the above is a vector space quotient.

Note that, if $\mathbb{F}[X_1, \ldots, X_n]_{\leq s}$ has a basis given by all monomials of total degree $\leq s$. If I is a monomial ideal, the above quotient simply removes those monomials in I from this basis, and our definition here agrees with the one specifically given for monomial ideals. It will turn out that $HF_I(s)$ is, for large enough s, always a polynomial for large enough s.

To arrive at this result, however, we first need to deal with a small issue that appears in dealing with multi-variate polynomial - ambiguity in determining a specific leading term. In the single-variate case, this is of course the term of highest degree. However, in order to choose a leading term in an expression like $xy^2 + x^2y$, we require the concept of a graded order.

Definition 4.11. Let $\mathbb{F}$ be a field and $n \in \mathbb{N}$. The **graded lexicographic order** with $y_1 > \cdots > y_n$ is the total order on monomials in $\mathbb{F}[y_1, \ldots, y_n]$ such that $y^\alpha > y^\beta$ if

1) $|\alpha| > |\beta|$, or

2) $|\alpha| = |\beta|$ and $\alpha_k > \beta_k$, where k is the smallest index i for which $\alpha_i \neq \beta_i$.

Example 4.12. If we order three variables $y_1 > y_2 > y_3$, then we have $y_1 y_2 y_3^2 < y_1 y_2^2 y_3$.

Definition 4.13. Let $P \in \mathbb{F}[X_1, \ldots, X_n]$. The **leading term** of P, denoted $\mathrm{LT}(p)$, is the monomial term in P which is largest in the graded lexicographic order described above.

Now, we may provide the final link that get us to a notion of dimension for general ideals.

Definition 4.14. Let I be an ideal. The **leading term ideal** $LT(I)$ is the ideal generated by leading terms in I, i.e.

$$LT(I) := \langle \{LT(f) : f \in I\} \rangle.$$

Proposition 4.15. *Let F be an algebraically closed field. Then, for any ideal I, we have $HF_I(s) = HF_{LT(I)}(s)$ for large enough s.*

The above proposition is significant because $LT(I)$ is a monomial ideal. We may then apply Proposition 4.9 to see that see that $HF_I(s)$ is also a polynomial for large s. We may then finally define the dimension of a general ideal.

Definition 4.16. Let I be an ideal. The **dimension** of I is the degree of $HF_I(s)$, which is a polynomial for all s large enough.

In this paper, we will be interested in showing that an ideal $I \subset \mathbb{F}[X_1, \ldots, X_n]$ has dimension zero. According to the above definition, it is sufficient to show that $LT(I)$ contains terms of the form $X_i^{k_i}$ for each i. Then, $HF_{LT(I)}(s)$ must be constant for any s large, since the set of monomials not in $LT(I)$ has bounded total degree. Therefore, the degree of $HF_{LT(I)}(s)$ as a polynomial must be zero, and $\dim I = \deg HF_I(s) = 0$ as well.

5 Finding Progressions

The problem we are interested in answering is how big (relative to p^2) a set $A \subset \mathbb{F}_p^2$ must be before we must find a progression of the form $x, x + P(\mu, \nu), x + Q(\mu, \nu)$ for two fixed polynomial functions $P, Q : \mathbb{F}_p^2 \to \mathbb{F}_p^2$ of the form

$$P(\mu, \nu) = (P_1(\mu, \nu), P_2(\mu, \nu)), Q(\mu, \nu) = (Q_1(\mu, \nu), Q_2(\mu, \nu))$$

and $x = (x, y) \in A \subset \mathbb{F}_p^2$, with $P(0, 0) = Q(0, 0) = 0$, each P_i, Q_i linearly independent. In particular we are interested in the asymptotics of this size as p gets large.

This seems a bit difficult to work with, so we will take a functional approach which has become common in additive combinatorics for dealing with this sort of problem. Namely, we will try to understand the sum

$$\Lambda_{P,Q}(f, g, h) := \mathbb{E}_{x,y,\mu,\nu \in \mathbb{F}_p} f(x, y) g(x + P_1(\mu, \nu), y + P_2(\mu, \nu)) h(x + Q_1(\mu, \nu), y + Q_2(\mu, \nu))$$

for $f, g, h \in V(\mathbb{F}_p^2; \mathbb{C})$.

Remark 5.1. We are particularly interested in this sum, as we have

$$\Lambda_{P,Q}(1_A, 1_A, 1_A) = \frac{1}{p^4} \left| \left\{ ((x, y), (x + P_1(\mu, \nu), y + P_2(\mu, \nu)), (x + Q_1(\mu, \nu), y + Q_2(\mu, \nu))) \in A^3 \right\} \right|,$$

since the product inside the expectation operator is 1 if and only if the corresponding sequence lies in $A \times A \times A$. Therefore, understanding the behavior of $\Lambda_{P,Q}$ will allows us to bound the number of progressions of the form we are interested in.

Remark 5.2. In fact, we can reduce this problem a bit further by comparing this $\Lambda_{P,Q}$ to the frequency with which we would expect a random progression to lie in $A \times A \times A$. In particular this is given simply by the cube of the density of A in $\mathbb{F}_p^2$: $\alpha^3 := \left(\frac{|A|}{p^2}\right)^3$.

It is not too hard to see that bounding this difference well enough will give us some guaranteed number of non-trivial progressions, when α (hence the size of A) is big enough. (i.e. to get an interesting result we must only show that for some $\frac{|A|}{p^2} = \alpha = o(1)$ the quantity $p^4\Lambda_{P,Q}(1_A, 1_A, 1_A) > |A|$ to guarantee some non-trivial progressions, since we have only $|A|$ trivial progressions of the form (x, x, x). For a detailed proof of this see the proof of corollary 1.2 in [5].)

To perform this comparison we will introduce one last bit of notation, define:

$$\Lambda_P(f, g) = \mathbb{E}_{x,y,\mu,\nu\in\mathbb{F}_p} f(x, y)g(x + P_1(\mu, \nu), y + P_2(\mu, \nu))$$

Then we have

Lemma 5.3. *We claim that*

$$|\Lambda_{P,Q}(1_A, 1_A, 1_A) - \alpha^3| \leq |\Lambda_{P,Q}(1_A, 1_A, 1_A - \alpha)| + \alpha|\Lambda_P(1_A, 1_A - \alpha)|$$

.

Proof. We obtain this by noting that

$$\begin{aligned}
&|\Lambda_{P,Q}(1_A, 1_A, 1_A) - \alpha^3| = |\Lambda_{P,Q}(1_A, 1_A, 1_A - \alpha) + \Lambda_{P,Q}(1_A, 1_A, \alpha) - \alpha^3| \\
&= |\Lambda_{P,Q}(1_A, 1_A, 1_A - \alpha) + \Lambda_{P,Q}(1_A, 1_A - \alpha, \alpha) + \Lambda_{P,Q}(1_A, \alpha, \alpha) - \alpha^3| \\
&= |\Lambda_{P,Q}(1_A, 1_A, 1_A - \alpha) + \alpha\Lambda_P(1_A, 1_A - \alpha)| \\
&\leq |\Lambda_{P,Q}(1_A, 1_A, 1_A - \alpha)| + \alpha|\Lambda_P(1_A, 1_A - \alpha)|.
\end{aligned}$$

□

So we have reduced our problem to bounding these two sums sufficiently well.

6 Bounding Λ_P

We will begin by bounding the simpler of these two terms, Λ_P.

Our goal is to bound this sum and to do so we will apply finite Fourier analysis, and the Weil bound.

Theorem 6.1 (Weil)**.** *Let χ be an additive character on a finite field $\mathbb{F}_q$ (of characteristic p), and $P \in \mathbb{F}_q[x]$ a polynomial. Then we have*

$$\sum_{x\in\mathbb{F}_q} \chi(P(x)) \ll_{\deg(P)} \sqrt{p}.$$

To apply this, we will show the following relation:

Lemma 6.2.

$$\Lambda_P(f, g) = \sum_{\chi_1,\chi_2\in\widehat{\mathbb{F}_p^2}} \hat{f}(\chi_2)\hat{g}(\chi_1)[\mathbb{E}_{x,y\in\mathbb{F}_p}\chi_1(x, y)\chi_2(x, y)][\mathbb{E}_{\mu,\nu\in\mathbb{F}_p}\chi_1(P_1(\mu, \nu), P_2(\mu, \nu))]$$

Proof. We use Fourier inversion.

$$\begin{aligned}\Lambda_P(f,g) &= \frac{1}{p^4}\sum_{x,y,\mu,\nu\in\mathbb{F}_p} f(x,y)g(x+P_1(\mu,\nu),y+P_2(\mu,\nu))\\ &= \frac{1}{p^4}\sum_{x,y,\mu,\nu\in\mathbb{F}_p} f(x)\sum_{\chi_1\in\widehat{\mathbb{F}_p^2}}\hat{g}(\chi_1)\chi_1(x+P_1(\mu,\nu),y+P_2(\mu,\nu))\\ &= \frac{1}{p^4}\sum_{x,y,\mu,\nu\in\mathbb{F}_p}\sum_{\chi_1,\chi_2\in\widehat{\mathbb{F}_p^2}}\hat{f}(\chi_2)\hat{g}(\chi_1)\chi_1(x,y)\chi_2(x,y)\chi_1(P_1(\mu,\nu),P_2(\mu,\nu))\\ &= \sum_{\chi_1,\chi_2\in\widehat{\mathbb{F}_p^2}}\hat{f}(\chi_2)\hat{g}(\chi_1)[\mathbb{E}_{x,y\in\mathbb{F}_p}\chi_1(x,y)\chi_2(x,y)][\mathbb{E}_{\mu,\nu\in\mathbb{F}_p}\chi_1(P_1(\mu,\nu),P_2(\mu,\nu))]\end{aligned}$$

since χ an additive character. □

In order for the sum over x,y to be non-zero, by orthogonality of characters, we must have $\chi_1=\overline{\chi_2}$, and in this case the sum is 1. Thus we obtain

$$\sum_{\chi\in\widehat{\mathbb{F}_p^2}}\hat{f}(\overline{\chi})\hat{g}(\chi)[\mathbb{E}_{\mu,\nu\in\mathbb{F}_p}\chi(P_1(\mu,\nu),P_2(\mu,\nu))]$$

We may apply the Weil bound to the sum over ν with μ fixed, and the trivial bound ($|e^{ix}|=1$) to obtain:

$$\Lambda_P(f,g)\ll_{P_1}\frac{1}{\sqrt{p}}\sum_{\chi\in\widehat{\mathbb{F}_p^2}}\hat{f}(\chi)\hat{g}(\overline{\chi})=\frac{1}{\sqrt{p}}\sum_{\chi\in\widehat{\mathbb{F}_p^2}}\hat{f}(\chi)\overline{\hat{g}(\chi)}\leq\frac{1}{\sqrt{p}}\|f\|\|g\|$$

by Plancherel and Cauchy–Schwarz. Thus in the case of $f=g=1_A$ we have

$$\Lambda_P(1_A,1_A)\ll_{P_1}\frac{\alpha}{\sqrt{p}}$$

where, as above, $\alpha=\frac{|A|}{p^2}$.

7 Bounding $\Lambda_{P,Q}$

We now turn to bounding the remaining term. This one is not so straightforward in how we may apply our Fourier theory, and requires some significant preparation through the application of Cauchy–Schwarz, and algebraic manipulation. In the application of this "Cauchy–Schwarz Method" we follow exactly the method of Peluse in [1], generalized to some extra dimensions.

So we have

Lemma 7.1. *Let $f,g,h:\mathbb{F}_p^2\to\mathbb{R}$, then*

$$\Lambda_{P,Q}(f,g,h)\leq\frac{|V_{P,Q}|^{1/8}}{p}\|f\|\cdot\|f_1\|\cdot\|f_2\|^{3/4}|\Lambda'_{P,Q}(h,h)|^{1/8}.$$

Where $V_{P,Q}$ is an affine algebraic variety over $\mathbb{F}_p^2$ determined by the polynomials P, Q and

$$\Lambda'_{P,Q}(f,g) := \mathbb{E}_{x,y\in\mathbb{F}_p,(\mu,\nu)\in V_{P,Q}} f(x,y)g(x+\Gamma_1(\mu,\nu), y+\Gamma_2(\mu,\nu))$$

where Γ_i some polynomials over $\mathbb{Z}[\mu_1,\dots,\mu_8,\mu_1,\dots,\nu_8]$ determined by P,Q to be given explicitly later.

Proof. By Cauchy–Schwarz we may first bound $|\Lambda_{P,Q}(f,g,h)|^2$ by

$$\|f\|^2\mathbb{E}_{x,y,\mu_1,\nu_1,\mu_2,\nu_2\in\mathbb{F}_p} g\begin{pmatrix}x+P_1(\mu_1,\nu_1)\\ y+P_2(\mu_1,\nu_1)\end{pmatrix} g\begin{pmatrix}x+P_1(\mu_2,\nu_2)\\ y+P_2(\mu_2,\nu_2)\end{pmatrix} h\begin{pmatrix}x+Q_1(\mu_1,\nu_1)\\ y+Q_2(\mu_1,\nu_1)\end{pmatrix} h\begin{pmatrix}x+Q_1(\mu_2,\nu_2)\\ y+Q_2(\mu_2,\nu_2)\end{pmatrix}.$$

Changing variables: $x \mapsto x - P_1(\mu_1,\nu_1), y \mapsto y - P_2(\mu_1,\nu_1)$ (and dividing through by our constant) we then have

$$\mathbb{E}_{\substack{x,y,\mu_i,\nu_i\in\mathbb{F}_p\\ i=1,2}} g\begin{pmatrix}x\\ y\end{pmatrix} g\begin{pmatrix}x-P_1(\mu_1,\nu_1)+P_1(\mu_2,\nu_2)\\ y-P_2(\mu_1,\nu_1)+P_2(\mu_2,\nu_2)\end{pmatrix} h\begin{pmatrix}x-P_1(\mu_1,\nu_1)+Q_1(\mu_1,\nu_1)\\ y-P_2(\mu_1,\nu_1)+Q_2(\mu_1,\nu_1)\end{pmatrix}$$
$$\times h\begin{pmatrix}x-P_1(\mu_1,\nu_1)+Q_1(\mu_2,\nu_2)\\ y-P_2(\mu_1,\nu_1)+Q_2(\mu_2,\nu_2)\end{pmatrix}.$$

We can then collect 4-tuples $(\mu_1,\mu_2,\nu_1,\nu_2)$ in fibers of $T_1(\mu_1,\mu_2,\nu_1,\nu_2), T_2(\mu_1,\mu_2,\nu_1,\nu_2)$ where

$$T_1(\mu_1,\mu_2,\nu_1,\nu_2) := P_1(\mu_2,\nu_2) - P_1(\mu_1,\nu_1), T_2(\mu_1,\mu_2,\nu_1,\nu_2) := P_2(\mu_2,\nu_2) - P_1(\mu_1,\nu_1).$$

Doing so, we obtain

$$\frac{1}{p^4}\sum_{x,y,z,w\in\mathbb{F}_p} g\begin{pmatrix}x\\ y\end{pmatrix} g\begin{pmatrix}x+z\\ y+w\end{pmatrix} \frac{1}{p^2}\sum_{\substack{\mu_i,\nu_i\in\mathbb{F}_p\\ i=1,2\\ T_1(\mu_1,\mu_2,\nu_1,\nu_2)=z\\ T_2(\mu_1,\mu_2,\nu_1,\nu_2)=w}} h\begin{pmatrix}x-P_1(\mu_1,\nu_1)+Q_1(\mu_1,\nu_1)\\ y-P_2(\mu_1,\nu_1)+Q_2(\mu_1,\nu_1)\end{pmatrix} h\begin{pmatrix}x-P_1(\mu_1,\nu_1)+Q_1(\mu_2,\nu_2)\\ y-P_2(\mu_1,\nu_1)+Q_2(\mu_2,\nu_2)\end{pmatrix}.$$

Then again taking a modulus squared of our quantities so far, and applying Cauchy–Schwarz to the terms depending only on x,y,z,w (thus obtaining 4 powers of $\|g\|$) we have that $|\Lambda_{P,Q}(f,g,h)|^4/\|f\|^4\|g\|^4$ is bounded by

$$\frac{1}{p^8}\sum_{\substack{x,y,z,w\in\mathbb{F}_p\\ T_1(\mu_{1+2i},\mu_{2+2i},\nu_{1+2i},\nu_{2+2i})=z\\ T_2(\mu_{1+2i},\mu_{2+2i},\nu_{1+2i},\nu_{2+2i})=w\\ i=1,2}} \sum_{\mu,\nu\in\mathbb{F}_{p^4}} \prod_{j=1,2} h\begin{pmatrix}x-P_1(\mu_{1+2j},\nu_{1+2j})+Q_1(\mu_{1+2j},\nu_{1+2j})\\ y-P_2(\mu_{1+2j},\nu_{1+2j})+Q_2(\mu_{1+2j},\nu_{1+2j})\end{pmatrix}$$
$$\times h\begin{pmatrix}x-P_1(\mu_{1+2j},\nu_{1+2j})+Q_1(\mu_{2+2j},\nu_{2+2j})\\ y-P_2(\mu_{1+2j},\nu_{1+2j})+Q_2(\mu_{2+2j},\nu_{2+2j})\end{pmatrix}$$

by Cauchy–Schwarz. We may sum the inner sum over z,w to obtain

$$\frac{1}{p^8}\sum_{\substack{x,y\in\mathbb{F}_p\\ (\mu,\nu)\in X}} \prod_{j=1,2} h\begin{pmatrix}x-P_1(\mu_{1+2j},\nu_{1+2j})+Q_1(\mu_{1+2j},\nu_{1+2j})\\ y-P_2(\mu_{1+2j},\nu_{1+2j})+Q_2(\mu_{1+2j},\nu_{1+2j})\end{pmatrix} h\begin{pmatrix}x-P_1(\mu_{1+2j},\nu_{1+2j})+Q_1(\mu_{2+2j},\nu_{2+2j})\\ y-P_2(\mu_{1+2j},\nu_{1+2j})+Q_2(\mu_{2+2j},\nu_{2+2j})\end{pmatrix},$$

where we denote by X the set of (μ,ν) given by summing over the fibers of T_1, T_2 we had previously.

We may again change variables, this time with $x \mapsto x - Q_1(\mu_1,\nu_1) + P_1(\mu_1,\nu_1), y \mapsto y - Q_2(\mu_1,\nu_1) + P_2(\mu_1,\nu_1)$. Then collecting terms in fibers of $T_3(\mu,\nu) := H_1(\mu_3,\nu_3) - H_1(\mu_1,\nu_1)$, $T_4(\mu,\nu) := H_2(\mu_3,\nu_3) - H_2(\mu_1,\nu_1)$, where $H_i(\mu,\nu) := Q_i(\mu,\nu) - P_i(\mu,\nu)$, as well as $T_5(\mu,\nu) := Q_1(\mu_2,\nu_2) - Q_1(\mu_1,\nu_1)$ and $T_6(\mu,\nu) := Q_2(\mu_2,\nu_2) - Q_2(\mu_1,\nu_1)$ we have

$$\frac{1}{p^8} \sum_{x,y,z,w,z',w' \in \mathbb{F}_p} h\begin{pmatrix} x \\ y \end{pmatrix} h\begin{pmatrix} x+z \\ y+w \end{pmatrix} h\begin{pmatrix} x+z' \\ y+w' \end{pmatrix} \sum_{\substack{(\mu,\nu)\in X \\ T_3(\mu,\nu)=z \\ T_4(\mu,\nu)=w \\ T_5(\mu,\nu)=z' \\ T_6(\mu,\nu)=w'}} h\begin{pmatrix} x+z+Q_1(\mu_4,\nu_4)-Q_2(\mu_3,\nu_3) \\ y+w+Q_2(\mu_4,\nu_4)-Q_2(\mu_3,\nu_3) \end{pmatrix}.$$

Then taking the modulus squared again, and applying Cauchy–Schwarz to the outer three terms depending only on x, y, z, w, z', w' (for 6 copies of $\|h\|$), we see that

$$\frac{|\Lambda_{P,Q}(f,g,h)|^8}{\|f\|^8 \cdot \|g\|^8 \cdot \|h\|^6}$$

is bounded by

$$\frac{1}{p^{10}} \sum_{x,y,z,w,z',w' \in \mathbb{F}_p} \sum_{\substack{(\mu,\nu)\in X\times X \\ T_3(\mu_{1,\dots,4},\nu_{1,\dots,4})=z \\ T_3(\mu_{5,\dots,8},\nu_{5,\dots,8})=z \\ T_4(\mu_{1,\dots,4},\nu_{1,\dots,4})=w \\ T_4(\mu_{5,\dots,8},\nu_{5,\dots,8})=w \\ T_5(\mu_{1,\dots,4},\nu_{1,\dots,4})=z' \\ T_5(\mu_{5,\dots,8},\nu_{5,\dots,8})=z' \\ T_6(\mu_{1,\dots,4},\nu_{1,\dots,4})=w' \\ T_6(\mu_{5,\dots,8},\nu_{5,\dots,8})=w'}} h\begin{pmatrix} x+z+Q_1(\mu_4,\nu_4)-Q_1(\mu_3,\nu_3) \\ y+w+Q_2(\mu_4,\nu_4)-Q_2(\mu_3,\nu_3) \end{pmatrix} h\begin{pmatrix} x+z+Q_1(\mu_8,\nu_8)-Q_1(\mu_7,\nu_7) \\ y+w+Q_2(\mu_8,\nu_8)-Q_2(\mu_7,\nu_7) \end{pmatrix}.$$

Changing variables a final time to $x \mapsto x - z - Q_1(\mu_4,\nu_4) + Q_1(\mu_3,\nu_3), y \mapsto y - w - Q_2(\mu_4,\nu_4) + Q_2(\mu_3,\nu_3)$ and summing over $z, w, z', w' \in \mathbb{F}_p$ we then have

$$\frac{1}{p^{10}} \sum_{x,y\in\mathbb{F}_p} \sum_{(\mu,\nu)\in V_{P,Q}} h\begin{pmatrix} x \\ y \end{pmatrix} h\begin{pmatrix} x+\Gamma_1(\mu,\nu) \\ y+\Gamma_2(\mu,\nu) \end{pmatrix}$$

where $\Gamma_i(\mu,\nu) := Q_i(\mu_8,\nu_8) - Q_i(\mu_7,\nu_7) - Q_i(\mu_4,\nu_4) + Q_i(\mu_3,\nu_3)$, and $V_{P,Q}$ the variety defined by common zeros of the polynomials

$$\begin{aligned}
R^{(1,i)}(\mu,\nu) &:= P_i(\mu_4,\nu_4) - P_i(\mu_3,\nu_3) - P_i(\mu_2,\nu_2) + P_i(\mu_1,\nu_1) \\
R^{(2,i)}(\mu,\nu) &:= P_i(\mu_8,\nu_8) - P_i(\mu_7,\nu_7) - P_i(\mu_6,\nu_6) + P_i(\mu_5,\nu_5) \\
R^{(3,i)}(\mu,\nu) &:= Q_i(\mu_6,\nu_6) - Q_i(\mu_5,\nu_5) - Q_i(\mu_2,\nu_2) + Q_i(\mu_1,\nu_1) \\
R^{(4,i)}(\mu,\nu) &:= H_i(\mu_7,\nu_7) - H_i(\mu_5,\nu_5) - H_i(\mu_3,\nu_3) + H_i(\mu_1,\nu_1).
\end{aligned}$$

So then we may rewrite this final sum as

$$\frac{|V_{P,Q}|}{p^8} \Lambda'_{P,Q}(h,h)$$

and we have our bound. □

It thus remains only to bound $|V_{P,Q}|$ and $\Lambda'_{P,Q}(h,h)$ to obtain our result. It turns out that a bound on this size of $V_{P,Q}$ will be necessary to bound the other term, so we will focus only on this.

In order to make this sum more tractable we again use Fourier inversion to obtain a character sum:

$$\Lambda'_{P,Q}(h,h) = \sum_{\chi_1,\chi_2 \in \widehat{F_p^2}} \hat{f}(\chi_2)\hat{g}(\chi_1)[\mathbb{E}_{x,y\in\mathbb{F}_p}\chi_1(x,y)\chi_2(x,y)][\mathbb{E}_{(\mu,\nu)\in V_{P,Q}}\chi_1(\Gamma_1(\mu,\nu),\Gamma_2(\mu,\nu))],$$

so by orthogonality of characters this is

$$\sum_{\chi\in\widehat{\mathbb{F}_p^2}} \hat{f}(\overline{\chi})\hat{g}(\chi)[\mathbb{E}_{(\mu,\nu)\in V_{P,Q}}\chi(\Gamma_1(\mu,\nu),\Gamma_2(\mu,\nu))].$$

It hence suffices to bound

$$\mathbb{E}_{(\mu,\nu)\in V_{P,Q}}\chi(\Gamma_1(\mu,\nu),\Gamma_2(\mu,\nu))$$

where we may safely ignore the case where χ is the trivial character in our case, since $f = g = h$ which has $\mathbb{E}_{x,y}h = 0$, and thus has 0 leading Fourier coefficient.

To bound this sum in other cases, we apply a theorem of Kowalski [8]

Theorem 7.2 (Kowalski)**.** *Let χ be an additive character, V an affine variety, $f \in \mathbb{F}_p[x_1,\dots,x_n]$. So long as f is not constant on some section of V of size $O(|V|)$, we have*

$$\left|\sum_{x\in V}\chi(f(x))\right| \ll_{f,V} \frac{|V|}{\sqrt{p}}.$$

So it remains only to bound the size of $V_{P,Q}$, and the fibers of $n\Gamma_1(\mu,\nu) + m\Gamma_2(\mu,\nu)$, which themselves clearly give an affine variety, which we will call $W_{P,Q}^{n,m}$.

$$W_{P,Q}^{n,m} = \left\{(\mu,\nu,\mu',\nu') \in V_{P,Q}\times V_{P,Q} \,\middle|\, \Gamma^{(n,m)}(\mu,\mu',\nu,\nu') = 0\right\}$$

where

$$\Gamma^{(n,m)} := n(\Gamma_1(\mu,\nu) - \Gamma_1(\mu',\nu')) + m(\Gamma_2(\mu,\nu) - \Gamma_2(\mu',\nu')).$$

This is difficult in general, but we can accomplish this thanks to a theorem of Lang and Weil.

Theorem 7.3 (Lang–Weil)**.** *Let $V(I)$ be an affine variety over $\mathbb{F}_p^n$. Then*

$$|V(I)| \ll_I p^{\dim(V)}.$$

So we need only bound the dimensions of these varieties sufficiently well to obtain our result.

Observe that obtaining dimension bounds for V and W of 8 and 15 respectively (so we may apply the bound of Kowalski, gives us the bound

$$|\Lambda'_{P,Q}(h,h)| \le \|h\|^2 \frac{p^{15/2}}{|V_{P,Q}|}$$

so then our inequality of 7.1 becomes

$$\Lambda_{P,Q}(f,g,h) \leq \frac{\|f\| \cdot \|g\| \cdot \|h\|}{p^{1/16}}$$

From this, and our bound 5.3, we have

$$|\Lambda_{P,Q}(1_A, 1_A, 1_A) - \alpha^3| \ll \frac{\|1_A\|^2 \|1_A - \alpha\|}{p^{1/16}} + \frac{\alpha^2}{\sqrt{p}}$$

where as previously $\alpha = \frac{|A|}{p^2}$.

Since this is only an order of magnitude estimate we may drop the lower order term from Λ_P and obtain a worst case bound on $\Lambda_{P,Q}(1_A, 1_A, 1_A)$ of

$$\Lambda_{P,Q}(1_A, 1_A, 1_A) \geq \alpha^3 - O\left(\frac{\|1_A\|^2 \|1_A - \alpha\|}{p^{1/16}}\right) = \alpha^3 - O\left(\frac{\alpha^{3/2}}{p^{1/16}}\right)$$

so multiplying through by p^4 to obtain a count on progressions we find that this is bounded below by.

$$\frac{|A|^3}{p^2} - O\left(|A|^{3/2} p^{1-1/16}\right)$$

so we see that for $|A| \gg p^{2-1/24}$ the left hand term will dominate our error term, and will trivially exceed p^2 so we are guaranteed non-trivial progressions obtaining our result.

8 Bounding Variety Dimensions

The Cauchy–Schwarz based argument in the previous sections proves our desired result, so long as we are able to bound the dimension of the varieties

$$V_{P,Q} = V(\langle R^{(1,1)}, R^{(1,2)}, R^{(2,1)}, R^{(2,2)}, R^{(3,1)}, R^{(3,2)}, R^{(4,1)}, R^{(4,2)} \rangle) \subset \mathbb{F}_p^{16}$$

and

$$\begin{aligned} W_{P,Q}^{n,m} = V(\langle & R^{(1,1)}(\mu_1, \ldots, \nu_8), R^{(1,2)}(\mu_1, \ldots, \nu_8), \\ & R^{(2,1)}(\mu_1, \ldots, \nu_8), R^{(2,2)}(\mu_1, \ldots, \nu_8), \\ & R^{(3,1)}(\mu_1, \ldots, \nu_8), R^{(3,2)}(\mu_1, \ldots, \nu_8), \\ & R^{(4,1)}(\mu_1, \ldots, \nu_8), R^{(4,2)}(\mu_1, \ldots, \nu_8), \\ & R^{(1,1)}(\mu_9, \ldots, \nu_{16}), R^{(1,2)}(\mu_9, \ldots, \nu_{16}), \\ & R^{(2,1)}(\mu_9, \ldots, \nu_{16}), R^{(2,2)}(\mu_9, \ldots, \nu_{16}), \\ & R^{(3,1)}(\mu_9, \ldots, \nu_{16}), R^{(3,2)}(\mu_9, \ldots, \nu_{16}), \\ & R^{(4,1)}(\mu_9, \ldots, \nu_{16}), R^{(4,2)}(\mu_9, \ldots, \nu_{16}), \\ & \Gamma^{(n,m)}(\mu_1, \ldots, \nu_{16}) \rangle) \subset \mathbb{F}_p^{32} \end{aligned}$$

Essentially, we must prove that each condition given by one of the R polynomials is independent, i.e. that $\dim V_{P,Q} \leq 8$ and $\dim W_{P,Q}^{n,m} \leq 15$ for each n, m. To this end, we follow the methods used by Peluse in [1]. It is sufficient to intersect $V_{P,Q}$ with eight hyperplanes and

get an object whose dimension as a variety is 0. For technical reasons beyond the scope of this paper (see again [1]) , it is also sufficient to calculate the dimension of $V_{P,Q}$ and $W_{P,Q}^{n,m}$ as varieties over $\overline{\mathbb{Q}}$. Working over an algebraically closed field, note that Proposition 4.15 gives us $\dim I = \dim \mathrm{LT}(I)$ for any ideal I. Therefore, our strategy will be to add equations representing eight hyperplanes to the ideal giving $V_{P,Q}$, or 15 to $W_{P,Q}^{n,m}$. Then, we will verify that for each μ_i and ν_j, a term of the form μ_i^k or ν_j^k appears as a leading term in this new ideal. This means that the affine Hilbert function $HF(s)$ counting the number of leading terms of degree less than s is bounded, and therefore has degree zero.

For the general case, we have yet to find a set of eight hyperplanes which, when added to the ideal for $V_{P,Q}$, can be combined with the R and S polynomials to produce all of the necessary leading terms. Our progress so far amounts to the following.

Proposition 8.1. *Let $V_{P,Q}$ be the variety defined above. Then,* $\dim V_{P,Q} \leq 12$.

Proof. If $\alpha_i = \deg P_i \in \mathbb{Z}^2_{\geq 0}$ and $\beta_i = \deg Q_i$, we may write

$$\begin{aligned}
P_1(\mu,\nu) &= \sum_{i+j=|\alpha_1|} c_{i,j}^{(1)} \mu^i \nu^j + \text{lower order terms} \\
P_2(\mu,\nu) &= \sum_{i+j=|\alpha_2|} c_{i,j}^{(2)} \mu^i \nu^j + \text{lower order terms} \\
P_2(\mu,\nu) &= \sum_{i+j=|\beta_1|} d_{i,j}^{(1)} \mu^i \nu^j + \text{lower order terms} \\
P_2(\mu,\nu) &= \sum_{i+j=|\beta_2|} d_{i,j}^{(2)} \mu^i \nu^j + \text{lower order terms.}
\end{aligned}$$

Now, choose a point $a = (x_1, \ldots x_8, y_1, \ldots y_8) \in \mathbb{F}_p^{16}$ such that a is contained in the top-dimensional component of $V_{P,Q}$. [1] Then, for each $i \leq 8$, we intersect with the hyperplane

$$\mu_i - x_1 = 0.$$

This gives us each μ_i as a possible leading term. Furthermore, we may extract another four ν_j terms as follows. Consider

$$R^{(1,1)} = P_1(\mu_1,\nu_1) - P_1(\mu_2,\nu_2) - P_1(\mu_3,\nu_3) + P_1(\mu_3,\nu_4).$$

Up to now, we have not specifically chosen a monomial order on the μ_i and ν_j, which we specify now as a graded lexicographic order with

$$\mu_8 < \nu_8 < \mu_6 < \nu_6 < \mu_7 < \nu_7 < \mu_1 < \nu_1 < \mu_2 < \nu_2 < \mu_3 < \nu_3 < \mu_4 < \nu_4 < \mu_5 < \nu_5.$$

(This order was chosen to make the first set of variables in the definition of each $R^{(i,j)}$ the leading terms of that particular polynomial). In this case, the leading term of R_1 will be of the form $c\mu_1^i\nu_1^j$ for some $i + j = |\alpha_1|$. If $i = 0$, we have a leading term of the form ν_1^j, as desired. Otherwise, we note that

$$R^{(1,1)} - c\mu_1^{i-1}\nu_1^j(\mu_1 - x_1)$$

[1] For technical reasons, since $V_{P,Q}$ can be made up of several disjoint components with different dimensions, we need to choose our planes to intersect the highest-dimensional component.

also lies in this ideal. Subtracting off this multiple of the linear term $\mu_1 - x_1$ removes the leading term $c\mu_1^i \nu_1^j$, and adds several lower-order terms, each of total degree $|\alpha| - 1$ (recall that x_1 is a scalar constant). We may repeat this process for all terms of total degree $|\alpha|$ in μ_1 and ν_1, with the exception of the term $c\nu_1^{|\alpha|}$, if it exists. If so, after we remove all other terms of degree $|\alpha|$, we see that $c\nu_1^{|\alpha|}$ is the leading term in our ideal. Otherwise, we may remove all degree $|\alpha|$ terms in $R^{(1,1)}$ (note that we must do this in the other three sets of μ and ν as well, but since the four copies of P_1 are all in independent variables, the procedure is exactly the same).

Repeating this process, we may continue decreasing the degree until either we arrive at a leading term of the form ν_1^k, or the resulting polynomial has total degree 1. In this second case, if there is no term of the form $c\nu_1$, then every term of P_1 must have been only in the variable μ. So, to get both μ_1 and ν_1 as leading terms, we may take the hyperplane $\nu_1 - y_1 = 0$ rather than $\mu_1 - x_1 = 0$, and the μ_1 leading term instead comes from $R^{(1,1)}$ itself.

We can repeat this process for $R^{(2,1)}$, $R^{(3,1)}$, and $R^{(4,1)}$ to get three more leading terms in ν_8, ν_6, and ν_7. Adding an additional four hyperplanes of the form $\nu_i - y_i$ for $i = 2, 3, 4, 5$ shows that way may exhibit all sixteen variables as leading terms of some polynomial in the relevant ideal, i.e. its corresponding variety has dimension zero. We have done this with a total of 12 hyperplanes, so $\dim V_{P,Q} \leq 12$. □

Conjecture 8.2. $\dim V_{P,Q} \leq 8$.

In order to reach this bound, would need to exhibit the remaining four ν_i as leading terms without adding them directly as was done at the end of the above argument. To do this, we must necessarily use information coming from the relationship between the $R^{(i,1)}$'s and the $R^{(i,2)}$'s, since further manipulation of the $R^{(i,1)}$'s alone only gives us information about the variety corresponding to the ideal $\langle R^{(1,1)}, R^{(2,1)}, R^{(3,1)}, R^{(4,1)} \rangle$, which has dimension greater than or equal to twelve (since it is based on four constraints in a 16-dimensional space).

For rather trivial cases, we may bypass this conjecture and the need to bound variety dimensions entirely. For example, suppose P_1 and Q_1 are functions only of μ, and P_2 and Q_2 are functions only of ν. Then, the search for a sequence $x, x + P(\mu, \nu), x + Q(\mu, \nu)\}$ in A can be done via a pigeonhole argument using already-known results for the one-dimensional case.

We have also verified the following special case.

Proposition 8.3. *Suppose that*

$$P(\mu, \nu) = \begin{pmatrix} \mu^a + \textit{lower order terms} \\ \nu^b + \textit{lower order terms} \end{pmatrix}$$
$$Q(\mu, \nu) = \begin{pmatrix} \mu^c + \textit{lower order terms} \\ \nu^d + \textit{lower order terms} \end{pmatrix}.$$

with $a \neq c$ and $b \neq d$. Then, $\dim V_{P,Q} \leq 8$ and $\dim W_{P,Q}^{n,m} \leq 15$ for all n and m.

Proof. Examining the definition of the $R^{(i,j)}$ polynomials which define $V_{P,Q}$, we see that we immediately have as leading terms $\mu_4, \nu_4, \mu_8, \nu_8, \mu_6, \nu_6, \mu_7$, and ν_7. Note that the conditions $a \neq c$ and $b \neq d$ guarantee that no cancellations occur in $R^{(4,i)}$, in which we subtract P_i from Q_i. This allows us to extract the μ_7 and ν_7 leading terms. Then, by adding eight additional hyperplanes $\mu_i - x_i$ and $\nu_j - y_j$ as before, we have all possible leading terms, and $\dim V_{P,Q} \leq 8$.

To bound $\dim W_{P,Q}^{n,m}$, the same logic as above gives us sixteen leading terms, one from each of the first sixteen polynomials determining $W_{P,Q}^{n,m}$ - eight ν_i's and eight ν_j's. With an additional eight hyperplanes $\nu_j - y_j$, we have all possible ν_j's as leading terms. Then, by adding seven more hyperplanes, we must recover the remaining eight μ_i's as leading terms. This is exactly the work of Peluse in [1] - although our polynomials are in two variables, Peluse's argument only deals with the terms of highest degree in each polynomial. By assumption, the polynomials $R^{(i,1)}$ have leading terms μ_j^a or μ_j^b, so the exact same logic carries through, completing the proof. □

One interpretation of this result is that, so long as the only highest-order term of the P_i and Q_i is not mixed in both μ and ν, then P and Q behave enough like single-variable polynomials that the varieties $V_{P,Q}$ and $W_{P,Q}^{n,m}$ are also very similar to the single-variate case.

9 Further Work

Since massaging the above 16-variable equations (and later on, 32-variable equations when dealing with $W_{P,Q}^{n,m}$) into the correct form has proven difficult, our work focused on a specific family of functions P and Q. Outside this case, however, nothing is currently known, so other results of any form would be new and interesting.

References

[1] S Peluse. Three-term polynomial progressions in subsets of finite fields. (1), 2017.

[2] L. Babai. The Fourier Transform and Equations over Finite Abelian Groups An introduction to the method of trigonometric sums. Technical report, 1989.

[3] Y. Ye. Hyper-Kloosterman sums and estimation of exponential sums of polynomials of higher degrees. 3, 1998.

[4] L.J. Mordell. On exponential sums related to Kloosterman sums. 1972.

[5] J. Bourgain and M. C. Chang. Nonlinear Roth type theorems in finite fields. *Israel Journal of Mathematics*, 221(2):853–867, 2017.

[6] É. Fouvry, E. Kowalski, and Philippe M. A study in sums of products. 2015.

[7] D. Cox, J. Little, and D. O'Shea. *Ideals, Varieties, and Algorithms*. Springer, New York, 3rd edition, 2007.

[8] E. Kowalski. Exponential sums over definable subsets of finite fields. *Israel Journal of Mathematics*, 160:219–251, 2007.