The Polynomial Method in Arithmetic Combinatorics

by

Robert Harling

MA4K9 Dissertation

Submitted to The University of Warwick

Mathematics Institute

April, 2020

Contents

1 Introduction

In recent years a new method of proof using algebraic techniques has enabled significant advances in several problems in arithmetic combinatorics. This method is the polynomial method. At its core, the polynomial method is a way of bounding sets by placing them in the zero set of a polynomial, and controlling the zero set through some measure of complexity of the polynomial, such as its degree.

This approach is one we use implicitly all the time. How many real numbers a are there such that $a^2 = 5$? These numbers are exactly the zeros of $f(x) = x^2 - 5$. As f is of degree 2, there can be at most two such a. This may seem trivial, but this simple notion can be extended and employed to prove powerful bounds in a range of problems.

Instances of the polynomial method have existed for decades [Tao13], but it is only recently that it has been considered independently as a method, and there is much to determine yet about its power and its limitations. In this report we shall explore various applications of the polynomial method, as this will demonstrate its power, its nature, and also how it might be applied in the future.

First we shall consider a simple application in providing an alternative proof of a classic result of Erdős, Ginzburg and Ziv [AD93]. Next we will consider the recent breakthroughs made in the problem of progression free sets using the polynomial method. We shall see the breakthrough made by Croot, Lev and Pach [CLP16], and the generalisation of this result by Ellenberg, Gisjwijt [EG16] and Tao [Tao]. In this we will consider the notion of slice rank developed by Tao, which we will then see applied by Naslund and Sawin [NS16] to prove new bounds on sunflower-free sets using the polynomial method. We will then turn to a classic proof of a bound on the points on a hyperelliptic curve in a finite field. We shall observe how we can use the multiplicty of zeros to obtain stronger bounds, and introduce the notion of the Hasse derivative [IKS04]. Finally we shall consider the recent progress made by Hanson and Petridis [HP19] on bounds on sumsets inside roots of unity using the polynomial method. We shall also consider how this result could be extended. Throughout we shall consider what, in our proofs, constitutes the polynomial method, and compare and contrast its different applications.

2 Finite Fields and Polynomials

While we shall consider complex and involved proofs throughout this report, there are few mathematical areas in which the reader will require pre-requisite knowledge. We will briefly discuss now some basic properties of finite fields and polynomial factorisation as these notions will be employed throughout the report.

2.1 Finite Fields

Let $\mathbb{F}$ be a field containing a finite number of elements. Consider the sequence of elements $\bar{1}, \bar{2}, \bar{3}, ..., \bar{n}, ...$ in $\mathbb{F}$, where $\bar{n}$ is just $1 + ... + 1$ n times. As $\mathbb{F}$ is finite, this sequence must repeat at some point. So suppose $\bar{n} = \bar{m}$ for some $n < m$. Then $\bar{k} = \bar{n} - \bar{m} = 0$. We can choose k to be the lowest such value. This k is called the characteristic of the field $\mathbb{F}$ and we write $\text{char}(\mathbb{F}) = k$. Suppose $k = ab$ for some $a, b \in \mathbb{Z}$. Then $\bar{a}\bar{b} = \bar{k} = 0$ in $\mathbb{F}$ and so either $\bar{a} = 0$ or $\bar{b} = 0$ in $\mathbb{F}$ as it is a field. Then either a or b is a multiple of k by the minimal choice of k, and each is at most k, so we conclude that either a or b equals k. So k must be prime. Let us then write $p = k$ for the characteristic of the field.

It can also be shown that $\mathbb{F}$ must have a prime power number of elements, $q = \tilde{p}^n$ for some prime $\tilde{p}$ and natural number n [Hana]. In fact, this $\tilde{p}$ is the same p as the characteristic of the field. Two finite fields of the same order are isomorphic, so we write $\mathbb{F}_q$ for the finite field of q elements [Hana]. We can also observe that $\mathbb{F}_q^\times = \mathbb{F}_q \setminus \{0\}$ is a multiplicative group of $q - 1$ elements, and so by Lagrange's theorem, $a^{q-1} = 1$ for all non-zero a in $\mathbb{F}_q$.

A final important observation is to compare the finite field $\mathbb{F}_q$ and the ring $\mathbb{Z}/q\mathbb{Z}$. We can consider the natural homomorphism $\phi : \mathbb{Z}/q\mathbb{Z} \mapsto \mathbb{F}_q$ where $\phi(a) = \bar{a}$. We observe that $\phi(a) = 0 \iff \bar{a} = 0 \iff p|a$. In the case $q = p$, we can therefore see that ϕ is injective and is an isomorphism. But when q is some non-trivial power of p this is not an isomorphism and we must be careful not to imagine $\mathbb{F}_q$ as $\mathbb{Z}/q\mathbb{Z}$.

2.2 Polynomial Factorisation

The ring of polynomials over a field $\mathbb{F}$ is the set

$$\mathbb{F}[x] = \{f(x) = a_n x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0 |\, a_n, ..., a_0 \in \mathbb{F}, n \geq 0\}$$

The degree of the polynomial $f(x) = a_n x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0$ where $a_n \neq 0$ is n. A zero in $\mathbb{F}$ of a polynomial f is an element $a \in \mathbb{F}$ such that $f(a) = 0$.

As $\mathbb{F}$ is a field, $\mathbb{F}[x]$ is a unique factorisation domain (UFD), and polynomials of the form $x - a$ are prime elements (but not necessarily the only prime elements). As $\mathbb{F}[x]$ is a Euclidean domain, we can show that $f(a) = 0$ if and only if $(x - a)|f(x)$. We can extend this to define a notion of multiplicity. f has a zero of multiplicity k at a if $(x - a)^k|f$ [Rot98].

Let $a_1, ..., a_r$ be zeros of f in $\mathbb{F}$, each of multiplicity $e_1, ..., e_r$. As $\mathbb{F}[x]$ is a Euclidean domain it is a unique factorisation domain and we have $e_1 + ... + e_r \leq deg(f)$. So f can only have finitely many zeros, and in fact at most $deg(f)$ zeros. One of the definitions of algebraic closure $\overline{\mathbb{F}}$ is that it is the smallest field containing $\mathbb{F}$ such that for any $f \in \mathbb{F}[x]$, for all the zeros $\overline{a_1}, ..., \overline{a_n}$ in $\overline{\mathbb{F}}$ of multiplicity $\overline{e_1}, ..., \overline{e_n}$, we have $\overline{e_1} + ... + \overline{e_n} = deg(f)$. (So f factors into linear factors completely in $\overline{\mathbb{F}}$).

3 Erdős-Ginzburg-Ziv

3.1 Zero-Sum Problems

Consider an abelian group G and a positive integer n. We can ask what is the largest subset of G such that it does not contain n elements that sum to zero. Such a problem is called a zero-sum problem. The classic result in this area is that of Erdős, Ginzburg and Ziv [AD93] who proved the best possible bound for the case $G = \mathbb{Z}/n\mathbb{Z}$.

Theorem 3.1 (**Erdős-Ginzburg-Ziv**). *Let n be an integer and $a_1, .., a_{2n-1}$ a sequence of not necessarily distinct elements in $\mathbb{Z}/n\mathbb{Z}$. Then there exists a subsequence of n elements that sum to 0.*

Indeed, $2n - 1$ is the smallest possible sequence length for which this works. To see this, consider the multiset of $n - 1$ copies of 0 and $n - 1$ copies of 1. There is no subset of cardinality n whose elements sum to 0 mod n, as we just miss being able to sum to 0 or n.

One way to extend the result is to consider the 2 dimensional case $G = \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$. The following result is conjectured:

Conjecture 3.2. *Let $a_1, ..., a_{4n-3}$ be a sequence of elements in $\mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$. Then there exists a subsequence of n elements that sum to 0.*

If true this would be the tightest possible bound, as we can see by considering the multiset of $n - 1$ copies each of $(0, 0), (0, 1), (1, 0), (1, 1)$. In general we can let $s(n, d)$ be the size of the smallest sequence in $(\mathbb{Z}/n\mathbb{Z})^d$ such that it must contain a subsequence of n elements that sum to 0. Then Theorem 3.1 says that $s(n, 1) = 2n - 1$ and Conjecture 3.2 says that $s(n, 2) = 4n - 3$. Alon and Dubiner showed that $s(n, 2) \leq 6n - 5$, and in general it can be shown that $s(n, d) \leq c_d n$ where c_d is some constant depending only on d [AD93].

One can also consider functions $f : A \to \mathbb{Z}/n\mathbb{Z}$ for some set A and ask about subsets of $f(A)$ that sum to zero. It can be shown that if $|A| = 2n - 2$ and $|f(A)| \geq 3$ then there is a set of n elements in $f(A)$ that sum to zero [FO96]. Note that the condition $|f(A)| \geq 3$ is crucial here, and prevents our example of $n - 1$ copies of 1 and $n - 1$ copies of 0 being a counterexample. There is a large amount of research into zero-sum problems. They have intrinsic links with graph theory, and can be studied with a large variety of methods

[GG06].

3.2 Proof of Erdős-Ginzburg-Ziv

This result was initially proved in 1961 and there are various proofs [AD93]. A key element of all the proofs is that it is sufficient to prove the result for prime n.

Lemma 3.3. *Suppose Theorem 3.1 holds for prime n. Then it holds for all n.*

Proof. We prove by induction on the number of prime factors of n (counting multiplicity). When n has only one prime factor, it is itself prime so we have the result by assumption. Now suppose $n = pm$ for some prime p. Consider a sequence $a_1, ..., a_{2pm-1}$ in $\mathbb{Z}/(pm)\mathbb{Z}$. By assumption we can find a subset A_1 of the sequence which has p elements which sum to 0 mod p. We can keep finding and removing such sequences A_i as long as there are $2p-1$ elements left, so we can do this $2m-1$ times.
So we have the subsets $A_1, ...A_{2m-1}$ of $\{a_1, ..., a_{2n-1}\}$ each of p elements which sum to 0 mod p.
Let $\bar{A}_i = \sum_{a \in A_i} a/p$. Then we have the sequence $\bar{A}_1, ..., \bar{A}_{2m-1}$ of $2m-1$ elements. So we may apply the induction hypothesis to find a subsequence of m elements which sum to 0 mod m. As each $\bar{A}_i$ is a sum of p elements that is 0 mod p, we get a sum of pm elements congruent to 0 mod pm which is exactly the result. □

We give a short description of the initial proof to demonstrate its difference to the proof using the polynomial method. The original proof makes use of the Cauchy-Davenport lemma [AD93]:

Lemma 3.4 (**Cauchy-Davenport**)**.** *Let p be a prime and let A,B be non empty subsets of $\mathbb{Z}/p\mathbb{Z}$. Then*

$$|A + B| \geq min\{p, |A| + |B| - 1\}$$

We can then easily apply this lemma to prove Theorem 3.1.

Classical proof of Erdős-Ginzburg-Ziv. Order the $a_1, ..., a_{2p-1}$ so that the sequence is increasing. If $a_i = a_{i+p-1}$ for some $i \leq p-1$ then $a_i + a_{i+1} + ... + a_{i+p+1} = pa_i = 0$ and we are done. Otherwise, let $A_i = \{a_i, a_{i+p-1}\}$ for $1 \leq i \leq p-1$. Repeatedly applying the Cauchy-Davenport lemma gives us that

$$|A_1 + ... + A_{p-1}| = p$$

so $A_1 + ... + A_{p-1} = \mathbb{Z}/p\mathbb{Z}$ This is because every time we add an A_i, the Cauchy-Davenport lemma shows that the size of the set must increase by at least 1. So every element of $\mathbb{Z}/p\mathbb{Z}$

is a sum of $p-1$ elements of the first $2p-2$ elements. So $-a_{2p-1}$ is a sum of such elements. So we can obtain a sequence of p elements summing to 0. □

This original proof is quite elementary and very different in nature to the proof using the polynomial method, which uses further properties of finite fields. The proof relies on the Chevalley-Warning theorem. This theorem's proof also uses techniques of the polynomial method. Roughly speaking, Chevalley-Warning states that if we have a system of polynomials that are all 'small' enough in degree, but in enough variables, then we are guaranteed a non-trivial common zero.

Theorem 3.5 (Chevalley-Warning). *Let $\mathbb{F}_q$ be a finite field of q elements and and characteristic p. Let $P_i(x_1, ..., x_m)$, $1 \leq i \leq n$, be polynomials over $\mathbb{F}_q$ of degree r_i respectively. If $\sum_{i=1}^{n} r_i < m$, then the number N of common zeros of the P_i satisfies $N \equiv 0$ mod p.*

Proof. The proof relies on two key observations that follow from the basic observation that $|\mathbb{F}_q^{\times}| = q-1$.
First, consider the sum

$$\sum_{x \in \mathbb{F}_q} x^r$$

for some $r < q-1$. We wish to show that this sum is zero. Observe that $x^r - 1 = 0$ has at most $r < q-1$ solutions so there exists a non-zero $y \in \mathbb{F}_q$ such that $y^r \neq 1$. Then observe that

$$y^r \sum_{x \in \mathbb{F}_q} x^r = \sum_{x \in \mathbb{F}_q} y^r x^r = \sum_{x \in \mathbb{F}_q} (yx)^r = \sum_{x \in \mathbb{F}_q} x^r$$

where the final equality comes because $\mathbb{F}_q$ is a field and so y is invertible. As $y^r \neq 1$ we must conclude that $\sum_{x \in \mathbb{F}_q} x^r = 0$.
Our second observation is that

$$N \equiv \sum_{x_1,...,x_m \in \mathbb{F}_q} \prod_{i=1}^{n} (1 - P_i(x_1, ..., x_m)^{q-1}) \text{ mod } p$$

This follows from our observation in Section 2 that in $\mathbb{F}_q$ we have $x^{q-1} = \begin{cases} 1 \text{ if } x \neq 0 \\ 0 \text{ if } x = 0 \end{cases}$
Expanding the expression for N on the right would give us various monomials of the form

$$x_1^{k_1} x_2^{k_2} ... x_m^{k_m}$$

with $\sum_{j=1}^{m} k_j \leq \sum_{i=1}^{n} (q-1) r_i < (q-1)m$ by assumption. So by the pigeonhole principle there is a j for which $k_j < q-1$. Then by our first observation, factoring out $x_j^{k_j}$ and summing over all x_j in $\mathbb{F}_q$ gives 0. We can do this for every monomial to get $N \equiv 0$ mod p. □

Although the actual statement refers to N modulo p, the key part is that if there is one common zero then there must be another. We now construct a system of polynomials such that their solution will give us the sequence we require.

Proof using the polynomial method. Consider the two polynomials over $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$

$$\sum_{i=1}^{2p-1} a_i x_i^{p-1} = 0$$

$$\sum_{i=1}^{2p-1} x_i^{p-1} = 0$$

Note that $x_1 = ... = x_{2p-1} = 0$ is a solution to both polynomials. Each polynomial has degree $p-1$. As $p-1+p-1 = 2p-2 < 2p-1$ we may apply the Chevalley-Warning theorem to conclude that there exists a non-trivial common zero. Let us call this non-trivial zero $(y_1, ..., y_{2p-1})$. We can see that the second polynomial implies that exactly p of the y_i are non-zero and that the first implies that the corresponding a_i (for which y_i is non-zero), sum to 0. These corresponding a_i give us our subsequence. □

This proof demonstrates the general concept of the polynomial method. Here, we construct polynomials whose common zero satisfies the properties we want. The polynomials are 'small enough' (here in terms of degree) that we can comment on the existence of the zero and hence prove the existence of our desired sequence.

We shall see that our approach in proving Chevalley-Warning is key in several polynomial method proofs. We construct a polynomial to count zeros, expand it into monomials and then use the pigeonhole principle to factor out low degree factors.

4 Cap Set Problem

4.1 History of the Cap Set Problem and Roth's Theorem

In the classic card game Set, players play with a deck of special cards. Each card has 3 defining properties: the number of shapes on it, the type of shape, and the shading of the shape. There are 3 possibilities for each property and so 81 possible cards. The aim of the game is to collect *sets* of 3 cards where in each property the 3 cards are all the same or are all different. For example the cards in Figure 1 form a *set*.

A 'cap set' is any collection of cards containing no *sets*. We can ask then what is the largest possible cap set? And what if we allowed each card to have more than 3 defining properties?

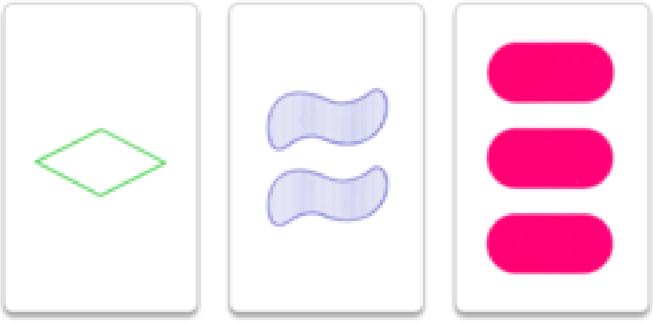

Figure 1: Set cards [Eng05]

Suppose each card has n defining properties each with 3 possibilities. Then we can easily construct a cap set by choosing 2 of the options for each property and only selecting cards that have either of those two options. Then any *set* of 3 cards can't be all different in each property and must all be the same, so in fact our 3 cards are just the same card. So there are cap sets of size 2^n.

We can view each card as an element a in $(\mathbb{Z}/3\mathbb{Z})^n$. Then 3 cards a, b, c form a *set* if and only if in each coordinate they are all different or all the same. We can then make the following observation:

Lemma 4.1. *For* 3 *elements* $\alpha, \beta, \gamma \in \mathbb{Z}/3\mathbb{Z}$*,* $\alpha + \beta + \gamma = 0$ *if and only if* $\alpha = \beta = \gamma$ *or* α, β, γ *are all different.*

This lemma can be shown by just considering all possible values for α, β, γ (there are 27 cases to check). This can naturally lead us to comparing this to the problem of finding arithmetic progression free sets.

A k-term arithmetic progression is a sequence of not all equal elements $x, x+r, x+2r, ..., x+(k-1)r$ where $r \neq 0$. We are specifically interested in 3-term arithmetic progressions where we have 3 not all equal elements $x, y, z \in \mathbb{F}$ such that $y = x+r$, $z = x+2r$ for some $r \in \mathbb{F}$. We can equivalently say that x, y, z satisfy $x + z = 2y$. Let $\mathbb{F} = (\mathbb{Z}/3\mathbb{Z})^n$. Then x, y, z are a 3-term arithmetic progression if and only if $x + y + z = 0$ as $1 = -2$ in characteristic 3.

We can then to notice the key equivalency that x, y, z forming a *set* in $\mathbb{F} = (\mathbb{Z}/3\mathbb{Z})^n$ is equivalent to x, y, z forming a 3-term arithmetic progression in $\mathbb{F} = (\mathbb{Z}/3\mathbb{Z})^n$.

$x, y, z \in (\mathbb{Z}/3\mathbb{Z})^n$ are all different or the same in each coordinate $\iff x+y+z = 0 \iff$ x, y, z form a 3-term arithmetic progression.

This problem is clearly related to finding sets in $\mathbb{Z}$ containing no 3-term arithmetic progressions. This integer case has a rich history and we consider the discussion of this history given in [Rah]. The problem of the integer case can be seen to begin with the following theorem of van der Warden:

Theorem 4.2 (**van der Warden 1927**). *Let r and k be natural numbers. There exists a natural number N such that any r-colour colouring of $1, ..., N$ contains a monochromatic k-term arithmetic progression.*

van der Warden's proof did not provide good bounds on N however. Erdős and Turan proposed the following conjecture to encourage work to find better bounds:

Conjecture 4.3 (**Erdős-Turan**). *For a given density $\delta > 0$ and positive integer k, there exists a natural number N such that any subset A of N with $|A| > \delta N$ contains a k-term arithmetic progression.*

Using Fourier analysis, Roth proved the conjecture for $k = 3$ and obtained the following bound:

Theorem 4.4 (**Roth 1953**). *For a given density $\delta > 0$, if $N > exp(exp(C\delta^{-1}))$ for an absolute constant C, then any subset A of density at least δ contains a 3-term arithmetic progression.*

Roth's bound was not very strong though and it would take a while to obtain stronger bounds. Later in the 20th century, Szeremedi managed to prove the Erdős-Turan conjecture for all k-term arithmetic progressions. His proof however relied on combinatorial techniques and did not obtain good bounds either. In 2001 Gowers finally managed to establish strong bounds:

Theorem 4.5 (**Gowers 2001**). *For a given density $\delta > 0$ and positive integer k, if $N \geq exp(exp((C\delta^{-1})^{1/c_k}))$ then any subset A of N with $|A| > \delta N$ contains a k-term arithmetic progression for some constant c_k.*

A beautiful consequence of Gowers bounds was the result of Tao and Green showing that the prime numbers contain arbitrarily long arithmetic progressions.

Several of the proofs use fourier analysis which relies on embedding $\mathbb{Z}$ into $\mathbb{Z}/N\mathbb{Z}$. We can then naturally ask about arithmetic progressions in other abelian groups. Let $r_3(G)$ be the largest subset of the abelian group G with no 3-term arithmetic progression. Our cap set problem is then the study of $r_3((\mathbb{Z}/3\mathbb{Z})^n)$. We have the trivial bound $r_3((\mathbb{Z}/3\mathbb{Z})^n) \leq 3^n$. In 1982 Brown and Buhler showed that $r_3((\mathbb{Z}/3\mathbb{Z})^n$ is $o(3^n)$, so is stronger than the trivial bound [EG16]. Until recently the best bound was $r_3((\mathbb{Z}/3\mathbb{Z})^n) \leq O(\frac{3^n}{n^{1+\epsilon}})$ for some $\epsilon > 0$. Recently however, Croot, Lev and Pach [CLP16] made a breakthrough using the polynomial method:

Theorem 4.6 (**Croot-Lev-Pach**). $r_3((\mathbb{Z}/4\mathbb{Z})^n) \leq c^n$ *for some* $c < 4$.

Soon after, Ellenberg and Gijswijt [EG16] generalised this application of the polynomial method for all $r_3(\mathbb{F}_q^n)$:

Theorem 4.7 (Ellenberg-Gijswijt). *Let $a, b, c \in \mathbb{F}_q$ such that $a + b + c = 0$. Suppose A is a subset of $\mathbb{F}_q^n$ such that for any $x, y, z \in A$, $ax + by + cz = 0$ if and only if $x = y = z$. Then $|A| = o(c^n)$ for some $c < q$. So $r_3(\mathbb{F}_q^n) \leq c^n$ for some $c < q$.*

Being 3-term arithmetic progression free is then to take $a = c = 1$ and $b = -2$ in the above. So taking $q = 3$, we have that the largest cap set is bounded by c^n for $c < 3$.

The best known lower bound on $r_3((\mathbb{Z}/3\mathbb{Z})^n)$ is $\approx 2.2^n$, given by a constructive proof [Ede04]. This lower bound is of the same exponential shape as our new upper bound. In the case of $r_3(\mathbb{Z}/N\mathbb{Z})$ no such exponential lower bound exists, i.e., $r_3(\mathbb{Z}/N\mathbb{Z})$ grows more quickly than $N^{1-\epsilon}$ for all $\epsilon > 0$ [EG16]. Historically the case of $(\mathbb{Z}/3\mathbb{Z})^n$ had been seen as a good analogue for the $\mathbb{Z}/N\mathbb{Z}$ case but this shows a disconnect between the two.

4.2 Proof using the Polynomial Method

These advances were made possible by the application of the polynomial method. In 2016 Tao offered a symmetric formulation of the results of Croot-Lev-Pach and Ellenberg-Gijswijt which we will now consider [Tao]. While Tao proved just the specific cap set case, we extend it to the more general case of Theorem 4.7.

We first introduce the concept of the slice rank of a function. Let $F : A^k \to \mathbb{F}$ be a function where A is some finite set and $\mathbb{F}$ some field. F is a 'rank one' function if we can write

$$F(x_1, ..., x_k) = f(x_i)g(x_1, ..., x_{i-1}, x_{i+1}, ..., x_k)$$

for all $(x_1, ..., x_k) \in A^k$ and for some $1 \leq i \leq k$ where $f : A \to \mathbb{F}$ and $g : A^{k-1} \to \mathbb{F}$. The slice rank of F is the smallest number of rank one functions needed to write F as a linear combination of rank one functions. Slice rank can be seen as the generalisation of rank of linear functions to higher dimensions.

The notion of slice rank was developed by Tao for this proof, and has since been employed in other proofs such as tri-coloured sum free sets in abelian groups. Tao and Sawin have also further developed the notion, finding new lower bounds on the slice rank of functions [TS].

We next observe that the slice rank of a 'diagonal hypermatrix' is analogous to the two dimensional case. Let δ_y be the standard kronecker-delta function where

$$\delta_y(x) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{if } x \neq y \end{cases}$$

Lemma 4.8 (Rank of Diagonal Hypermatrices). *Let A be a finite set, $\mathbb{F}$ a field, $k \geq 2$ and $c_a \in \mathbb{F}$ constants for each $a \in A$. Then the function $F : A^k \to \mathbb{F}$ where*

$$F(x_1, ..., x_k) = \sum_{a \in A} c_a \delta_a(x_1)...\delta_a(x_k)$$

has slice rank equal to the number of non-zero c_a.

In the case $k = 2$ this is just saying that the rank of a diagonal matrix is the number of non-zero entries it has.

Proof. We prove by induction on k.
In the case k=2, we have

$$F(x_1, x_2) = \sum_{a \in A} c_a \delta_a(x_1)\delta_a(x_2)$$

Then F is already written uniquely as a sum of rank one functions so F must have rank equal to the number of non-zero c_a.
We now assume the result holds for $k - 1$ and consider the case for k.
It is clear that the rank is at most the number of non-zero c_a as all the functions in the sum are rank one functions, so we just wish to find the lower bound. We may assume without loss of generality that all the c_a are non-zero by removing from A any a where $c_a = 0$. Then the number of non-zero c_a is just $|A|$.
Seeking a contradiction, we now assume that F has rank at most $|A| - 1$. So we may write

$$\sum_{a \in A} c_a \delta_a(x_1)...\delta_a(x_k) = \sum_{i=1}^{k} \sum_{\alpha \in I_i} f_{i,\alpha}(x_i) g_{i,\alpha}(x_1, ..., x_{i-1}, x_{i+1}, ..., x_k) \tag{1}$$

where $I_1, .., I_k$ are some index sets where $|I_1| + ... + |I_k| \leq |A| - 1$ and for functions $f_{i,\alpha} : A \to \mathbb{F}$ and $g_{i,\alpha} : A^{k-1} \to \mathbb{F}$.
We now consider the space of functions h from A to $\mathbb{F}$ that are orthogonal to all the $f_{k,\alpha}$, that is,

$$\sum_{x \in A} f_{k,\alpha}(x) h(x) = 0$$

for all $\alpha \in I_k$. Each h is specified by the $|A|$ values chosen for the elements in A, subject to the $|I_k|$ conditions imposed by the above statement. So this space has dimension d at least

$|A|-|I_k|$. We wish to find an h that is non-zero on at least $|A|-|I_k|$ elements. Consider the $d \times |A|$ matrix of the values of the basis elements of this space. Such a matrix is non-singular and so contains a non-singular $d \times d$ minor. So such an h can be constructed from the basis functions represented in this $d \times d$ minor. Now, we multiply equation 1 for F by $h(x_k)$ and sum over all x_k. On the left hand side, our summands are only non-zero when $x_k = a$. On the right hand side we essentially eliminate the variable x_k from our functions because of the orthogonality of h. We obtain

$$\sum_{a \in A} c_a h(a)\delta_a(x_1)...\delta_a(x_{k-1}) = \sum_{i=1}^{k-1}\sum_{\alpha \in I_i} f_{i,\alpha}(x_i)\tilde{g}_{i,\alpha}(x_1,...,x_{i-1},x_{i+1},...,x_{k-1})$$

where

$$\tilde{g}_{i,\alpha}(x_1,...,x_{i-1},x_{i+1},...,x_{k-1}) = \sum_{x_k \in \alpha} g_{i,\alpha}(x_1,...,x_{i-1},x_{i+1},...,x_k)h(x_k)$$

Then the right hand side has rank at most $|A|-1-|I_k|$ as it is a sum of rank one functions, whereas the left hand side has rank $|A|-|I_k|$ by the induction hypothesis. This is a contradiction. □

Let a,b,c be three elements in $\mathbb{F}_q$ such that $a+b+c=0$. Suppose $A \subset \mathbb{F}_q$ such that for all $x,y,z \in A$, $ax+by+cz=0$ if and only if $x=y=z$. Taking $a=c=1$, $b=-2$, this is to say that A contains no 3-term arithmetic progressions.
Consider the space M_n of monomials over $\mathbb{F}_q$ in n variables where each variable is of degree at most $q-1$. Then let M_n^d be the space restricted to monomials of degree at most d. So $M_n^d = \{\alpha x_1^{i_1}...x_n^{i_n} | \alpha \in \mathbb{F}_q, 0 \le i_1,...,i_n \le q-1, i_1+...+i_n = d\}$. Then let $m_d^n = \dim(M_d^n)$.

Lemma 4.9. *The function* $(x,y,z) \mapsto \delta_{0^n}(ax+by+cz)$ *on* $(\mathbb{F}_q^n)^3$ *has rank at most* $3m_{(q-1)n/3}^n$.

Proof. Our proof relies on a similar observation to that of the proof of Chevalley-Warning, that we can express characteristic functions as polynomials. Indeed,

$$\delta_{0^n}(ax+by+cz) = \prod_{i=1}^{n}(1-(ax_i+by_i+cz_i)^{q-1})$$

We can see that the right hand side of this equation expands to give a sum of monomials, each of degree at most $(q-1)n$. Each such monomial is of the form

$$\alpha x_1^{i_1}...x_n^{i_n} y_1^{j_1}...y_n^{j_n} z_1^{k_1}...z_n^{k_n}$$

where $\alpha \in \mathbb{F}_q$, $i_1,...,i_n,j_1,...,j_n,k_1,...,k_n \in \{1,..,q-1\}$ and $i_1+...+i_n+j_1+...+j_n+k_1+...+k_n \le (q-1)n$. By the pigeonhole principle, at least one of the total powers of the x_i, y_i or z_i must be at most $(q-1)n/3$. Consider all monomials where this is the case for the x_i. We can collect all such monomials in the sum

$$\sum_{f \in M^n_{(q-1)n/3}} f(x) g_f(y,z)$$

where g_f is some corresponding monomial in y and z. This is a sum of at most $m^n_{(q-1)n/3}$ rank one functions. We can similarly write the monomials where $j_1, ..., j_n \leq (q-1)n/3$ and $k_1, ..., k_n \leq (q-1)n/3$ as the sums of at most $m^n_{(q-1)n/3}$ rank one functions. We conclude that our function has rank at most $3m^n_{(q-1)n/3}$. □

In this proof we see clearly the relation to our proof of Erdos-Ginzburg-Ziv. We constructed a similar polynomial which is non-zero only at our 'special elements': in this case, 3-term arithmetic progressions. We expanded this into monomials and used the pigeonhole principle to factor out common factors of low degree. In this case however, we use this to bound the functions slice rank, not to consider it modulo a value. One key difference in this proof is that we have used slice rank as the measure of 'complexity' of our polynomial, not the degree.

Recall that for $x, y, z \in A$, $ax + by + cz = 0$ if and only if $x = y = z$. So restricting our function to A^3, we see that

$$\delta_{0^n}(ax + by + cz) = \sum_{\alpha \in A} \delta_\alpha(x)\delta_\alpha(y)\delta_\alpha(z)$$

By Lemma 4.8, this function has slice rank equal to $|A|$. By Lemma 4.9, the slice rank is bounded above by $3m^n_{(q-1)n/3}$. So it remains to obtain a bound on $m^n_{(q-1)n/3}$. We can do this by applying Cramér's theorem for large deviation problems. Let X be a random variable taking values $0, 1, ..., q-1$ uniformly (so with probability $1/q$). Then $m_{(q-1)n/3}/q^n$ is the probability that n independent copies of X have mean at most $(q-1)/3$. By Cramér's theorem [EG16], we have that

$$\lim_{n\to\infty} \frac{1}{n} log(m_{(q-1)n/3}/q^n) = -I((q-1)/3)$$

where

$$I(x) = \sup_{\theta \in \mathbb{R}} \{\ f_x(\theta) = \theta x - log((1 + e^\theta + ... + e^{(q-1)\theta})/q)\}$$

Now, $f_x(0) = 0$ and $f'_x(0) > 0$ when $x \neq (q-1)/2$, so $-I((q-1)/3) < 0$. Then our equation gives us that $m_{(q-1)n/3} = O(q^n e^{-\alpha n}) = O(c^n)$ for some $c < q$. We have therefore proved Theorem 4.7.

Corollary 4.9.1. *If A is a cap set in $(\mathbb{Z}/3\mathbb{Z})^n$, then $|A| = o(c^n)$ for $c \approx 2.756$.*

Proof. As A is a cap set we may apply Theorem 4.7 with $a = b = c = 1$ and $q = 3$. We can find that $3e^{-I(2/3)} < 2.756$ [EG16]. So by Theorem 4.7, $|A| = o(2.756^n)$. □

5 Sunflowers

5.1 Sunflower-free sets

We now turn to the closely related topic of sunflowers in collections of sets and consider the work of Naslund and Sawin [NS16]. A collection of n sets form an n-sunflower if the intersection of any two distinct sets is the same.
More formally, let U be some universal set and let $\mathcal{F} = \{A_1, ..., A_n\}$ be a collection of subsets of U. $\mathcal{F}$ is an n-sunflower if $A_i \cap A_j = A_k \cap A_l$ for all $i, j, k, l \in \{1, ..., n\}$.
In this section we're interested in sunflower-free collections. More specifically, we're interested in collections that contain no 3-sunflowers, i.e. no 3 sets A_1, A_2, A_3 such that $A_1 \cap A_2 = A_2 \cap A_3 = A_3 \cap A_1$. This is a condition on $\mathcal{F}$ and we can ask how big can an $\mathcal{F}$ with this property be? Historically, such a problem can be formulated in 2 ways.

Firstly, we can consider a 3-sunflower-free family of sets $\mathcal{F}$ where each set is of size n and ask how big can such $\mathcal{F}$ be with respect to n?
Alternatively, we can consider $\mathcal{F}$ a collection of subsets of $\{1, ..., n\}$ and ask how big can $\mathcal{F}$ be with respect to n?
Let us consider this second case. There are 2^n possible subsets of $\{1, .., n\}$ so trivially $|\mathcal{F}| \leq 2^n$. It was a conjecture of Erdős and Szemerédi that the exponent could be reduced. [NS16]

Conjecture 5.1 (Erdős-Szemerédi). *Let $k \geq 3$. If $\mathcal{F}$ is a k-sunflower-free collection of subsets of $\{1, ..., n\}$ then*

$$|\mathcal{F}| < c_k^n$$

for some constant $c_k < 2$ depending on k.

Erdős and Szemerédi were able to obtain a bound of $2^n exp(-c\sqrt{n})$ for some constant c, which is stronger than the trivial bound but is not the conjectured reduction in the exponent.
In the case $k = 3$, Alon et al. were able to show that the result of Ellenberg and Gijswijt proved the conjecture [NS16].
We frame the problem as follows. Let $F_k(n)$ be the size of the largest k-sunflower-free collection of subsets of $\{1, ..., n\}$. Then let

$$\mu_k = \limsup_{n\to\infty} F_k(n)$$

Then the above conjecture states that $\mu_k < 2$ for $k \geq 3$.

Theorem 5.2. *$\mu_3 \leq \sqrt{1 + c}$ where $c \approx 2.756$ is the constant from Corollary 4.9.1.*

Proof. Let $\mathcal{F}$ be a collection of subsets of $\{1, ..., n\}$. We wish to show that if $\mathcal{F}$ contains a sunflower, we can construct a cap set with related size. We can then apply the cap set bound from the previous section.

Given a set $S \subset \{1, ..., 2n\}$, we can consider the corresponding set $S' \in \{0,1\}^{2n}$ (where $S'_i = \begin{cases} 1 \text{ if } i \in S \\ 0 \text{ if } i \notin S \end{cases}$). We then wish to encode S' into an element of $\tilde{S} \in \{0,1,2,3\}^n$. We do this by considering the four vectors: $u_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $u_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $u_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $u_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Each pair of points in S' can be expressed as u_1, u_2, u_3 or u_4, so by taking the subscript value we obtain our $\tilde{S} \in \{0,1,2,3\}^n$. Let $\tilde{\mathcal{F}}$ be the set of vectors in $\{0,1,2,3\}^n$ corresponding to the sets in $\mathcal{F}$.

Now take a point $x \in \{0,1\}^n$ and consider the set $\tilde{\mathcal{F}}_x = \{\tilde{S} \in \tilde{\mathcal{F}} : \tilde{S}_i = 3 \text{ if and only if } x_i = 1\}$. Let $|x| = \sum_{i=1}^n x_i$. Then for any element $\tilde{S}$ in $\tilde{\mathcal{F}}_x$ we may ignore co-ordinates with $\tilde{S}_i = 3$ and consider $\tilde{S}$ as an element in $\{0,1,2\}^{n-|x|} = \mathbb{F}_3^{n-|x|}$. Suppose three elements in $\tilde{\mathcal{F}} \subset \mathbb{F}_3^{n-|x|}$ form a non-trivial 3-term arithmetic progression. Then from our consideration of the cap set problem we know that in each coordinate the elements are exactly $0,1,2$ in some order or are all the same, and in at least one coordinate are $0,1,2$. As u_0, u_1, u_2 correspond to sets forming a sunflower (they each have empty intersection with each other), we can conclude that the 3 sets in $\mathcal{F}$ corresponding to our arithmetic progression form a sunflower.

So $\mathcal{F}$ 3-sunflower-free implies that $\tilde{\mathcal{F}}_x$ is a cap set and hence by Theorem 4.7, $|\tilde{\mathcal{F}}_x| \leq c^{n-|x|}$. So

$$|\mathcal{F}| \leq \sum_{x \in \{0,1\}^n} c^{n-|x|} = \sum_{j=0}^{n} \binom{n}{j} c^{n-j} = (1+c)^n$$

And so $\mu_3^s \leq \sqrt{1+c} \leq \sqrt{1+2.756} \approx 1.938$. □

This is an impressive application of the cap set result to prove Erdős-Szemerédi's conjecture. However, Naslund and Sawin were able to obtain an even stronger result by applying the polynomial method directly [NS16].

We begin showing this result by making some observations. As we used in the previous proof, given a set $A \subset \{1, ..., n\}$ there is a corresponding element $A' = (A_0, \ldots, A_n) \in \{0,1\}^n$ (where $A_i = \begin{cases} 1 \text{ if } i \in A \\ 0 \text{ if } i \notin A \end{cases}$).

Suppose $\mathcal{F}$ is collection of sunflower-free subsets of $\{1, ..., n\}$ and let $\mathcal{F}'$ be the corresponding collection of vectors in $\{0,1\}^n$. Then for any three distinct vectors $\underline{x}, \underline{y}, \underline{z} \in \mathcal{F}'$ there exists an integer $1 \leq i \leq n$ such that $x_i, y_i, z_i = \{0,1,1\}$ (as otherwise the corresponding elements in $\mathcal{F}$ would form a sunflower).

Suppose we make a further requirement on $\mathcal{F}$ that it contains no two sets where one is a proper subset of another. Then we can make a stronger statement that for any $\underline{x}, \underline{y}, \underline{z} \in \mathcal{F}'$ not all equal, there exists an integer $1 \leq i \leq n$ such that $\{x_i, y_i, z_i\} = \{0, 1, 1\}$.

We also reformulate Lemma 4.8.

Lemma 5.3. *Let A be a finite set, $\mathbb{F}$ a field and $T : A^3 \to \mathbb{F}$ such that $T(x, y, z) \neq 0$ if and only if $x = y = z$. Then the slice rank of T is $|A|$.*

Proof. Let $c_a = T(a, a, a)$ for each $a \in A$. Then

$$T(x, y, z) = \sum_{a \in A} c_a \delta_a(x) \delta_a(y) \delta_a(z)$$

and by Lemma 4.8 T has slice rank $|A|$ as each c_a is non-zero. □

Theorem 5.4. *Let $\mathcal{F}$ be a 3-sunflower-free collection of subsets of $\{1, ..., n\}$. Then*

$$|S| \leq \left(\frac{3}{2^{2/3}} \right)^{n(1+o(1))}$$

Proof. Suppose $\mathcal{F}' \subset \{0, 1\}^n$ is 3-sunflower-free. Let $\mathcal{F}'_l$ be the elements of $\mathcal{F}'$ with exactly l entries that are 1. So $\mathcal{F}'_l$ is sunflower-free and contains no two sets where one is a proper subset of another. Take $x, y, z \in \mathcal{F}'_l$. Then

$$\begin{aligned} x + y + z \in \{0, 1, 3\}^n &\iff \text{there doesn't exist } i \text{ such that } \{x_i, y_i, z_i\} = \{0, 1, 1\} \\ &\iff x = y = z \end{aligned}$$

So consider the function $T : (\{0, 1\}^n)^3 \to$ where

$$T(x, y, z) = \prod_{i=1}^{n} (2 - (x_i + y_i + z_i))$$

T is non-vanishing if and only $x + y + z \in \{0, 1, 3\}$. So if we restrict T to $\mathcal{F}'_l$ then T is non-vanishing if and only if $x = y = z$. So by Lemma 5.3, T has slice rank at least $|\mathcal{F}'_l|$. Now just as we did in the proof of the cap set bound, we expand the right hand side of our product into monomials to find a bound on the function's rank. Expanding the right hand side gives monomials of the form

$$x_1^{i_1} ... x_n^{i_n} y_1^{j_1} ... y_n^{j_n} z_1^{k_1} ... z_n^{k_n}$$

where $i_1, ..., i_n, j_1, ..., j_n, k_1, ..., k_n \in \{0, 1\}$ and $i_1 + ... + i_n + j_1 + ... + j_n + k_1 + ... + k_n \leq n$. So by the pigeon hole principle at least one of the sums $i_1 + ... + i_n, j_1 + ... + j_n, k_1 + ... + k_n$ is at most $n/3$. Consider all monomials where this is the case for the x_i. We can collect all such monomials in the sum

$$\sum_f f(x)g_f(y,z)$$

There are $\sum_{k\leq n/3}\binom{n}{k}$ such possible f. Considering the cases for the sums $j_1+...+j_n$ and $k_1+...+k_n$ we find that T has rank at most $3\sum_{k\leq n/3}\binom{n}{k}$. So $|\mathcal{F}'_l|\leq 3\sum_{k\leq n/3}\binom{n}{k}$ and so

$$|\mathcal{F}'|\leq\sum_{l=0}^{n}|\mathcal{F}'_l|\leq 3(n+1)\sum_{k\leq n/3}\binom{n}{k}\leq\left(\frac{3}{2^{2/3}}\right)^{n(1+o(1))}$$

□

Does this approach give us a stronger bound? Well $\frac{3}{2^{2/3}}\approx 1.8899$. So by applying the polynomial method directly, we have obtained a bound which is better than the ≈ 1.938 we obtained by using the cap set bound.

This proof is clearly very similar to our proof of the cap set bound. We created a polynomial that was non-zero at our 'special elements'. We expanded into monomials and factored out low degree elements using the pigeonhole principle. We then argued combinatorially for the bound.

5.2 Sunflowers in $(\mathbb{Z}/D\mathbb{Z})^n$

We now turn to the problem of sunflowers in $(\mathbb{Z}/D\mathbb{Z})^n$ for $D>2$. Our proof here will use a variation on the polynomial method that uses multiplicative characters instead of polynomials. This may seem to be a large deviation from our standard approach, but we will see clearly that at the core this is still the polynomial method.

We define a k-sunflower in $(\mathbb{Z}/D\mathbb{Z})^n$ (where $k\leq d$) to be a collection of k vectors such that in each coordinate they are all different or are all the same. In the case $k=D=3$, this is in fact equivalent to the collection not containing any 3-term arithmetic progressions. Then our result from the cap set section implies a bound on 3-sunflower-free collections $A\subset(\mathbb{Z}/D\mathbb{Z})^n$ of $|A|\leq c_D^n$ when D is prime for some constant c_D. There is however a conjecture of even stronger bounds on k-sunflower-free sets in $(\mathbb{Z}/D\mathbb{Z})^n$:

Conjecture 5.5. *Let $k\leq D$ and let A be a subset of $(\mathbb{Z}/D\mathbb{Z})^n$ that is k sunflower-free. Then*

$$|A|\leq b_k^n$$

for some constant b_k that depends only on k.

The key here is that the constant b_k depends only on k and not on D. Unfortunately we will not be able to prove such a bound. But we will apply our methods directly again to obtain a bound on 3-sunflower-free sets in $(\mathbb{Z}/D\mathbb{Z})^n$.

Theorem 5.6. *Let $D \geq 3$ and let A be a subset of $(\mathbb{Z}/D\mathbb{Z})^n$ that is 3-sunflower-free. Then*

$$|A| \leq c_D^n$$

where $c_D = \frac{3}{2^{2/3}}(D-1)^{2/3}$.

Before we prove this theorem, it is useful to properly introduce the notion of a character. For our purposes, a character is a group homomorphism

$$\chi : \mathbb{Z}/D\mathbb{Z} \to \mathbb{C}^\times$$

The first key comment is that here we consider $\mathbb{Z}/D\mathbb{Z}$ as an additive group, so 0 is our identity element and 1 is a generator of the group. This distinction is important, as we are considering $\mathbb{C}^\times$ as a multiplicative group. So $\chi(a+b) = \chi(a)\chi(b)$. From this we can make some easy observations about the image of χ. For any $x \in \mathbb{Z}/D\mathbb{Z}$, $Dx = 0$, so $\chi(a)^D = \chi(Da) = \chi(0) = 1$ and so every point in the image of χ must be a D^{th} root of unity. A consequence of this is that $|\chi(x)| = 1$ for all $x \in \mathbb{Z}/D\mathbb{Z}$. Now we may observe that $\chi(a)\overline{\chi(a)} = |\chi(a)|^2 = 1$ and conclude that $\chi(a)^{-1} = \overline{\chi(a)}$.

In the sense of characters of group representations, our characters are characters of 1 dimensional representations. So by the orthogonality relations we have the following [JL01]:

Lemma 5.7.

$$\frac{1}{|D|}\sum_{\chi}\chi(a-b) = \begin{cases} 1 \text{ } \textit{if } a = b \\ 0 \text{ } \textit{otherwise} \end{cases}$$

This then allows us to construct the following function on $(\mathbb{Z}/D\mathbb{Z})^3$:

$$\frac{1}{|D|}\sum_{\chi}(\chi(a)\overline{\chi(b)} + \chi(b)\overline{\chi(c)} + \chi(a)\overline{\chi(c)}) = \begin{cases} 0 \text{ if } a,b,c \text{ are distinct} \\ 1 \text{ if exactly two of } a,b,c \text{ are equal} \\ 3 \text{ if } a = b = c \end{cases}$$

where we remember that $\chi(a-b) = \chi(a)\overline{\chi(b)}$. We can now prove Theorem 5.6:

Proof of Theorem 5.6. Define the function $T : (\mathbb{Z}/D\mathbb{Z})^n) \times (\mathbb{Z}/D\mathbb{Z})^n) \times (\mathbb{Z}/D\mathbb{Z})^n) \to \mathbb{C}^\times$ by

$$T(x,y,z) = \prod_{j=1}^{n}(\frac{1}{|D|}\sum_{\chi}(\chi(x_j)\overline{\chi(y_j)} + \chi(y_j)\overline{\chi(z_j)} + \chi(x_j)\overline{\chi(z_j)}) - 1)$$

Then T is non-zero if and only if the x, y, z form a 3-sunflower or are all equal. If we restrict T to $A \times A \times A$ then T is non-zero if and only if $x = y = z$. So by Lemma 5.3, the slice rank of T is at least $|A|$. This approach is clearly analogous to the proof of Theorem 5.4 and the next step is clear. Instead of considering expanding the right hand side into monomials, we expand it into a sum of products of characters of the form:

$$\chi_1(x_i)...\chi_n(x_n)\psi_1(y_1)...\psi_n(y_n)\zeta_1(z_1)...\zeta_n(z_n)$$

where at most $2n$ of the characters are non-trivial. So by the pigeonhole principle for at least one of the χ_i, ψ_i or ζ_i, there are at most $2n/3$ which are non-trivial. Suppose it is true for the χ_i. We can then factor out from all appropriate terms the product of the χ_i to obtain a rank one function. There are $\sum_{k \leq 2n/3} \binom{n}{k}(D-1)^k$ possibilities for our χ_i as there are $D-1$ distinct non-trivial characters $\chi : \mathbb{Z}/D\mathbb{Z} \to \mathbb{C}^\times$. Doing the same thing for the cases when the ψ_i or ζ_i have at most $2n/3$ non-trivial characters, we can bound the slice rank of T and find that

$$|A| \leq 3 \sum_{k \leq \frac{2n}{3}} \binom{n}{k} (D-1)^k$$

We then require just some minor combinatorial manipulation to obtain the bound in the form we would like. We note that as $D > 2$, $\frac{D-1}{2} \geq 1$.

$$\begin{aligned}
\sum_{k \leq \frac{2n}{3}} \binom{n}{k}(D-1)^k &\leq \sum_{k \leq \frac{2n}{3}} \binom{n}{k}(D-1)^k(\frac{D-1}{2})^{2n/3-k} \\
&= (\frac{D-1}{2})^{-n/3} \sum_{k \leq \frac{2n}{3}} \binom{n}{k}(D-1)^k(\frac{D-1}{2})^{n-k} \\
&\leq (\frac{D-1}{2})^{-n/3} \sum_{k \leq n} \binom{n}{k}(D-1)^k(\frac{D-1}{2})^{n-k} \\
&= (\frac{D-1}{2})^{-n/3}(D-1+\frac{D-1}{2})^n = \left(\frac{3}{2^{2/3}}(D-1)^{\frac{2}{3}}\right)^n
\end{aligned}$$

We can then let $c_D = \frac{3}{2^{2/3}}(D-1)^{\frac{2}{3}}$. We now just need to use a small trick to remove the factor of 3 in our bound $|A| \leq 3c_D^n$. Observe that if A is a sunflower-free set in $(\mathbb{Z}/D\mathbb{Z})^n$ then A^k is a sunflower-free set in $(\mathbb{Z}/D\mathbb{Z})^{kn}$, so $|A| \leq (3c_D)^{nk} = 3^{1/k}c_D^n$. We can then just take $k \to \infty$ to obtain our bound. □

So although we have used characters here, we can see clearly the links to our other polynomial method proofs. We constructed a function out of characters that was non-zero only at our special elements. We used slice rank as a measure of complexity. We expanded the function into monomials of characters, and used the pigeonhole principle to factor out small common factors. This then enabled us to argue combinatorially for a bound on the function's slice rank.

6 Points on a Hyperelliptic Curve in a Finite Field

6.1 Hyperelliptic curves

We now turn to the study of hyperelliptic curves and the points lying on them, following the arguments in [IKS04]. One can also compare to the discussion in [Hanb] and [Mas16]. A hyperelliptic curve C_f is given by a polynomial $f \in \mathbb{F}[x]$ of degree $m \geq 3$ for some field $\mathbb{F}$ and is the set of points (x, y) satisfying

$$y^2 = f(x)$$

In the case f is of degree 3, the curve is an elliptic curve which is a much studied object in number theory. For our purposes, we shall assume that f is not a square in $\mathbb{F}[x]$.
Finding points on an elliptic curve is not simple. One must first choose what field to look for points in. Naturally, as f is defined over $\mathbb{F}$ we may look for points in $\mathbb{F}^2$, but we could also consider points over the algebraic closure $\overline{\mathbb{F}}$ of $\mathbb{F}$. In this section we are interested in the case where $\mathbb{F} = \mathbb{F}_q$ is a finite field of q elements and characteristic p. Let $C_f(\mathbb{F}_q)$ be the set of points in $\mathbb{F}_q$ on C_f and let $N = |C_f(\mathbb{F}_q)|$ be the number of such points. We want to know if we can obtain good bounds on N. We can easily establish some trivial bounds on N. As $C_f(\mathbb{F}_q) \subset \mathbb{F}_q^2$, we must have $|N| \leq q^2$. But we can also note that for every $x \in \mathbb{F}_q$, $y^2 = f(x)$ has at most two solutions, so in fact $|N| \leq 2q$. N must be non-negative also, so we can write

$$|N - q| \leq q$$

We can also make some predictions about what sort of bound we expect. What we're interested in is whether $f(x)$ is a quadratic residue or not for each x. As half of all elements of $\mathbb{F}_q$ are quadratic residues, we might expect that 'on average' $y^2 = f(x)$ has one solution for each x. So we could expect that N would be close to q. In fact, we will prove the following bound

$$|N - q| \leq 8m\sqrt{q}$$

This bound is only an improvement when $\sqrt{q} \geq 8m$ so we shall assume this throughout. To prove this bound we will use the following observation:

Lemma 6.1. *For $w \in \mathbb{F}_q$,*

$$w^{\frac{q-1}{2}} = \begin{cases} 0 \textit{ if } w = 0 \\ 1 \textit{ if } w \textit{ is a non-zero square} \\ -1 \textit{ if } w \textit{ is a non-square} \end{cases}$$

Proof. In general $(w^{\frac{q-1}{2}})^2 = 1$ for any non-zero w, so $w^{\frac{q-1}{2}}$ must be $0, 1$ or -1.
The case when $w = 0$ is obvious.
When $w = z^2$ for some non-zero z in $\mathbb{F}_q$, $w^{\frac{q-1}{2}} = z^{q-1} = 1$.
The polynomial $x^{\frac{q-1}{2}} - 1 = 0$ can have at most $\frac{q-1}{2}$ solutions. Consider $x = m^2$ for $m = 1, 2, ..., (q-1)/2$. Suppose $m_1^2 = m_2^2$, then we get $(m_1 - m_2)(m_1 + m_2) = 0$. As $0 < m_1, m_2 \leq (q-1)/2$, we have $m_1 + m_2 \neq 0$ and $m_1 = m_2$, or $m_1 = m_2 = (q-1)/2$ and we have equality also. So we have $(q-1)/2$ distinct squares and so exactly $(q-1)/2$ solutions of $x^{\frac{q-1}{2}} - 1 = 0$.
All our remaining w in $\mathbb{F}_q$ must then give us $w^{\frac{q-1}{2}} = -1$, and these remaining w are exactly the non-squares in $\mathbb{F}_q$. □

This useful lemma gives us a way to relate points on the curve and zeros of a polynomial. We will consider the polynomial

$$g(x) = f(x)^c$$

where $c = (q-1)/2$. By our lemma we have that

$$g(x) = \begin{cases} 0 \text{ if } f(x) = 0 \iff y^2 = f(x) \text{ has one solution in } \mathbb{F}_q \\ 1 \text{ if } f(x) \text{ is a non-zero square } \iff y^2 = f(x) \text{ has two solutions in } \mathbb{F}_q \\ -1 \text{ if } f(x) \text{ is a non square } \iff y^2 = f(x) \text{ has no solutions in } \mathbb{F}_q \end{cases}$$

Now let $N_0 = |\{x \in \mathbb{F}_q | g(x) = 0\}|$ and let $N_1 = |\{x \in \mathbb{F}_q | g(x) = 1\}|$. Then

$$N = N_0 + 2N_1$$

We can also let $N_{-1} = |\{x \in \mathbb{F}_q | g(x) = -1\}|$. Then we have

$$N_0 + N_1 + N_{-1} = q$$

which allows us to obtain another equation for N:

$$N = 2q - 2N_{-1} - N_0$$

As a slight generalisation, let $S_a = \{x \in \mathbb{F}_q | g(x) = 0 \text{ or } a\}$.

Take $a \in \mathbb{F}_q$ and l a natural number satisfying $m \leq l \leq q/8$. Our aim is to construct a non-zero polynomial $r \in \mathbb{F}_q[x]$ such that r has a zero of degree l at all points in S_a, and

$$deg(r) < cl + 2ml(l-1) + mq$$

From this it is then relatively simple to deduce our bound.

Suppose we have such an r. Then we have

$$|S_a| < c + 2m(l-1) + mq/l$$

Choose $l = 1 + [\sqrt{q}/2]$ (where $[x]$ is the integer part of x). Our assumption that $8m \leq \sqrt{q}$ and that $m \geq 3$ ensure that this choice of l satisfies $m \leq l \leq q/8$. We then get

$$|S_a| < c + 2m(\sqrt{q}/2) + \frac{mq}{\sqrt{q}/2} < c + 4m\sqrt{q}$$

Then we have $N_0 + N_1 = |S_1| < c + 4m\sqrt{q}$. So

$$N = N_0 + 2N_1 < 2(N_0 + N_1) < q - 1 + 8m\sqrt{q} < q + 8m\sqrt{q}$$

Similarly we have

$$\begin{aligned} N = 2q - 2N_{-1} - N_0 \geq 2q - 2(N_{-1} + N_0) &= 2q - 2|S_{-1}| \\ &\geq 2q - 2(c + 4m\sqrt{q}) \geq q + 1 - 8m\sqrt{q} > q - 8m\sqrt{q} \end{aligned}$$

We can therefore deduce our desired bound:

$$|N - q| < 8m\sqrt{q}$$

6.2 Constructing the polynomial

Now we wish to construct a polynomial with zeros of high order at the points in $C_f(\mathbb{F}_q)$. Our normal approach would to be consider the derivatives of the function and show that these are all zero at such points. Over a field of characteristic zero this is fine and we can use the following lemma [Mas16]:

Lemma 6.2. *Let $\mathbb{F}$ be a field of characteristic zero. Suppose we have a polynomial $f \in \mathbb{F}[x]$ such that $f^{(k)}(a) = 0$ for all $k \leq l - 1$ at some point a in $\mathbb{F}$. Then f has a root of order at least l at a.*

However over fields of non-zero characteristic we have to be more careful. Consider the polynomial $f(x) = x^p$ over a field of characteristic p. Then $f^{(p)}(x) = p!x^0 = 0$. So $f^{(l)}(0) = 0$ for $0 \leq l \leq p$ but 0 is only a zero of order p. To overcome this we introduce a modified form of derivative called the Hasse derivative.

The Hasse derivative is a linear operator $E^k : \mathbb{F}[x] \to \mathbb{F}[x]$ where

$$E^k(x^k) = \binom{n}{k} x^{n-k}$$

and we then extend the definition linearly to all of $\mathbb{F}[x]$. We can see that this definition does not have our previous problem, as $E^p(x^p) = \binom{p}{p}x^0 = 1$.

The Hasse derivative has several important properties which we state here [IKS04]:

Lemma 6.3. *Let f, g be polynomials in $\mathbb{F}_q[x]$. Then we have the following*

1. *Suppose $E^k(f)(a) = 0$ for all $k \leq l - 1$. Then h has a zero of order at least l at a.*
2. *Suppose $f(x) = h(x, x^q)$ for some $h \in \mathbb{F}_q[x, y]$. Then*
$$E^k(f)(x) = E^k_x(h)(x, x^q)$$
where $E^k_x(h)$ is the Hasse derivative of h with respect to x.
3. *Let $k \leq r$, then*
$$E^k(fg^r) = hg^{r-k}$$
for some $h \in \mathbb{F}_q[x]$ with
$$deg(h) \leq deg(f) + kdeg(g) - k$$

This now gives us the machinery we need to construct our r. We will consider polynomials of the form

$$r(x) = f^l \sum_{j=1}^{J} (r_j + s_j g)x^{iq}$$

for some J and polynomials r_j, s_j over $\mathbb{F}_q$ each of degree at most $c - m$. Then

$$deg(r) \leq lm + (c - m) + cm + Jq \leq (J + m)q$$

It is important to make sure that r is non-zero, so we need the following lemma:

Lemma 6.4. *$r = 0$ in $\mathbb{F}_q$ if and only if $r_j = s_j = 0$ for all $1 \leq j \leq J$.*

Proof. We shall prove by contradicting our assumption that $f(x)$ is not a square in $\mathbb{F}_q[x]$. We can assume without loss of generality that $f(0) \neq 0$.
Seeking a contradiction, suppose $r = 0$ but not all the r_j, s_j are zero. Let k be the smallest index such that either r_k or s_k is non-zero. Divide r by $f^l x^{kq}$ to get

$$\sum_{j=k}^{J} (r_j + s_j g)x^{(j-k)q} = 0$$

Let $h_0 = \sum_{j=k}^{J} r_j x^{(j-k)q}$ and $h_1 = \sum_{j=k}^{J} s_j x^{(j-k)q}$. We then have

$$h_0 + h_1 g = 0$$

We can write this as $h_0 = -h_1 g$. Squaring both sides and multiplying by f gives us

$$h_0^2 f = h_1^2 f^q$$

recalling that $g = f^c = f^{(q-1)/2}$. We can then note that $h_0 \equiv r_k \bmod x^q$ and that $h_1 \equiv s_k \bmod x^q$. So we have

$$r_k^2 f \equiv s_k^2 f^q \bmod x^q$$

Then as $f \in \mathbb{F}_q[x]$, we have that $f(x)^q = f(x) = f(x^q) \equiv f(0) \bmod x^q$. So

$$r_k^2 f \equiv s_k^2 f(0) \bmod x^q$$

But recall that we have bounds on the degree of r_k and s_k: $2deg(r_k)+m \leq 2(c-m)+m < q$ and $2deg(s_k) < 2(c-m) < q$. So we in fact have

$$r_k^2 f = s_k^2 f(0)$$

This contradicts our assumption that f is a non-square in $\mathbb{F}[x]$. □

To talk about the zeros of high order, we need to talk about the Hasse derivatives of r. Fortunately we can show that these take on a 'nice' form:

Lemma 6.5. *There exist polynomials r_j^k, s_j^k each of degree at most $c-m+k(m-1)$ such that*

$$E^k(r)(x) = f^{l-k} \sum_{j=1}^{J} (r_j^k + s_j^k g) x^{jq}$$

Proof. We can write $r(x) = h(x, x^q)$ where $h \in \mathbb{F}[x,y]$ is

$$\begin{aligned} h(x,y) &= f(x)^l \sum_{0 \leq j \leq J} (r_j(x) + s_j(x) f^c(x)) y^j \\ &= \sum_{0 \leq j \leq J} (f(x)^l r_j(x) + s_j(x) f^{l+c}(x)) y^j \end{aligned} \tag{2}$$

So by Lemma 6.3,

$$E^k(r)(x) = E_x^k(h)(x, x^q) = \sum_{0 \leq j \leq J} (E^k(f^l r_j) + E^k(f^{l+c} s_j)) x^{jq}$$

Then again by Lemma 6.3, there exist polynomials r_j^k, s_j^k such that $E^k(f^l r_j) = f^{l-k} r_j^k$ and $E^k(f^{l+c} s_j) = f^{l+c-k} s_j^k$ with $deg(r_j^k) \leq deg(r_j) + kdeg(f) - k \leq c - m + k(m-1)$ and similarly $deg(s_j^k) \leq c - m + k(m-1)$. □

We then want $E^k(r)(x) = 0$ for all $x \in S_a$ and all $0 \leq k \leq l-1$. Well, for x in S_a,

$$E^k(r)(x) = f^{l-k} \sum_{j=0}^{J} (r_j^k + s_j^k a) x^j$$

where we use that $g(x) = a$ or 0 and that x is in $\mathbb{F}_q$ so $x^q = x$. We can in fact by choice of coefficients of r_j and s_j make these derivatives identically zero as we shall now show. It is sufficient to make the polynomial

$$h_k(x) = \sum_{j=0}^{J} (r_j^k + s_j^k a) x^j$$

identically zero. The degree of h_k is at most $J + c - m + k(m-1)$ so to make h_k zero we must solve at most $J + c - m + k(m-1)$ linear equations where the variables are the coefficients of r_j and s_j. We must do this for all h_k for $0 \leq k \leq l-1$ so we have at most

$$\sum_{k=0}^{l} J + c - m + k(m-1) = l(J + c - m) + \frac{1}{2} l(l-1)(m-1)$$

equations.
We have $2J(c-m)$ variables so we can construct our r providing we have more variables than equations. Choosing large enough J should allow us to do so, and in fact taking $J = \frac{l}{q}(c + 2m(l-1))$ is sufficient using that $l \leq q/8$ [IKS04]. So we can prove that such r exists and hence that we have our bound.

Our method of proof here may at first seem quite different to those in the cap set problem and the sunflower problems. However, it is clear that this proof is the polynomial method. We constructed a polynomial with zeros at our desired points and then used the polynomial's degree to obtain a bound. There are a couple of interesting differences here. Firstly, we used the order of each zero also to obtain stronger bounds. And secondly, we did not actually explicitly construct our polynomial. Instead we just described one with enough coefficients that we could choose them to give the polynomial the properties we require. We shall see both these features again in our next section.

7 Sumset Decompositions of Roots of Unity

7.1 Roots of Unity

We now turn to considering the recent work of [HP19] and follow it closely. Over any field $\mathbb{F}$ we can consider the roots of unity. Quite simply these are just the elements of finite multiplicative order. We can be more specific and consider the $\mathrm{d^{th}}$ roots of unity:

$$Z_d = \{a \in \mathbb{F} | a^d = 1\}$$

This group has a natural multiplicative structure. Indeed, it is a multiplicative subgroup of $\mathbb{F}$. The sum-product phenomenon is the observation that subsets of rings tend to have either a large sumset or a large product set. So if R is a ring and $A \subset R$, we would expect either $|A + A| = |\{a_1 + a_2 | a_1, a_2 \in A\}|$ to be large, and hence A to have little additive structure, or that $|A \times A| = |\{a_1 a_2 | a_1, a_2 \in A\}|$ is large and that A hence has little multiplicative structure. Though we discuss it here very loosely, the sum-product phenomenon can be formulated very strictly. For example, the following result can be shown [Tao08]:

Theorem 7.1. *Suppose $\mathbb{F}_p$ is a finite field of prime order and $A \subset \mathbb{F}_p$. Then for any $\delta > 0$ there exists $\epsilon > 0$ and $c > 0$ such that*

$$|A + A| + |A \times A| \geq c|A|^{1+\epsilon}$$

for all non-empty $A \subset \mathbb{F}_p$ with $|A| < p^{1-\delta}$

With these intuitions we might then expect Z_d to not have much additive structure. One way to observe a lack of additive structure would to be consider sumset decompositions of Z_d, that is, sets $A, B \subset \mathbb{F}$ such that $Z_d = A + B$. We might expect that no such decomposition exists, or that it is not a very complex decomposition i.e. $|A||B|$ is quite 'small'.

We can see quite easily that this is the case in the complex numbers. Indeed, we know that in $\mathbb{C}$ all the roots of unity lie on the unit circle, which seems to have no additive structure.

Lemma 7.2. *Let $d > 4$ and Z_d be the d^{th} roots of unity in $\mathbb{C}$. Then Z_d admits no non-trivial sumset decomposition.*

Proof. Suppose we have a non-trivial sumset decomposition, so

$$Z_d = A + B$$

for some sets $A, B \subset \mathbb{C}$ with $2 \leq |A| \leq |B|$. So we may choose two distinct elements a_1, a_2 in A. Let S^1 be the unit circle. Then $a_1 + B, a_2 + B$ are subsets of S^1.
Next note that

$$a_1 - a_2 = (a_1 + b) - (a_2 + b)$$

for all $b \in B$. So $a_1 - a_2$ has at least $|B|$ representations as a difference of two elements on the unit circle. Any complex number can be written as the difference of two points on the unit circle at most 2 ways [HP19], so we conclude that $|B| = 2$ and that $d = 4$. □

In the case $d \leq 4$, Z_d is too small to admit a non-trivial decomposition.

Unfortunately this approach does not extend readily to finite fields of prime order. In $\mathbb{F}_p$, our roots of unity no longer lie on a circle and in fact all units are contained on the line $\{1, ..., p-1\}$. Over $\mathbb{F}_p$ the units are the $(p-1)th$ roots of unity and for $p > 5$ we can decompose them as so:

$$\{0, \frac{p-1}{2}\} + \{\ 1, ..., \frac{p-1}{2}\} = \{1, ..., p-1\} = Z_{p-1} = \mathbb{F}_p^\times$$

So we are left to consider the cases $4 \leq d < p-1$. An important case is where $d = \frac{p-1}{2}$ as then Z_d is the set of quadratic residues of $\mathbb{F}_p$ (see Lemma 6.1).

7.2 Bounds on $|A||B|$

Historically attempts to consider $|A||B|$ have considered multiplicative character sums. A multiplicative character ϕ is a group homomorphism $\phi : \mathbb{F}_p^\times \to \mathbb{C}$. Note this is slightly different to the characters we considered in Theorem 5.6 which considered the additive group $\mathbb{Z}/D\mathbb{Z}$. For a given multiplicative character ϕ we can consider the sum

$$S_\phi(A, B) = \sum_{a \in A} \sum_{b \in B} \phi(a+b)$$

Trivially $|S_\phi(A, B)| \leq |A||B|$ since $|\phi(x)| \leq 1$. In the case $|A||B| > p$, a theorem of Vinogradov gives a stronger bound [HP19]:

Theorem 7.3. *Let $A, B \subset \mathbb{F}_p$ and ϕ a non-trivial multiplicative character. Then*

$$|S_\phi(A, B)| \leq \sqrt{p|A||B|}.$$

This gives an important bound for $|A||B|$ [HP19]:

Corollary 7.3.1. *Let $A, B \subset \mathbb{F}_p$ such that $A + B \subset Z_d$ for some d properly dividing $p-1$. Then $|A||B| \leq p$.*

But in a new paper, Hanson and Petridis [HP19] were able to strongly improve on this bound by applying the polynomial method. As in our hyperelliptic curves proof, we use derivatives and the multiplicity of zeros.

Theorem 7.4. *Let p be a prime and let d be an integer properly dividing $p-1$. Let*

$A, B \subset \mathbb{F}_p$ such that $A + B \subset Z_d \cup \{0\}$. Then

$$|A||B| \leq d + |B \cap (-A)|$$

Proof. Let us assume that $|A| \leq |B|$ and enumerate A as $A = \{a_1, ..., a_m\}$ where $m = |A|$. In the case $m = 1$, we have

$$|A + B| = |B| = \begin{cases} d \text{ if } |B \cap (-A)| = 0 \\ d + 1 \text{ if } |B \cap (-A)| = 1 \end{cases} = d + |B \cap (-A)|$$

For $m > 1$, we apply the polynomial method. We first consider the following polynomial:

$$G(x) = \sum_{k=1}^{m} c_k (x + a_k)^{m-1} = \sum_{j=0}^{m-1} \left(\binom{m-1}{j} \sum_{i=1}^{m} a_i^{m-1-j} c_i \right) x^j$$

for some $c_k \in \mathbb{F}_p$. We wish to make G a constant polynomial by our choice of the c_k. So we're solving the $m - 1$ equations

$$\sum_{i=1}^{m} a_i^{m-1-j} c_i = 0$$

for $j = 1, ..., m - 1$. We can do this as we have $m - 1$ equations and m variables so can find a non-trivial solution. We next show that G is non-zero. Let $c = G(x)$. Then what we have done is solve the Vandermonde system

$$\begin{pmatrix} a_1^{m-1} & ... & a_m^{m-1} \\ ... & ... & ... \\ a_1 & ... & a_m \\ 1 & ... & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ ... \\ c_{m-1} \\ c_m \end{pmatrix} = \begin{pmatrix} 0 \\ ... \\ 0 \\ c \end{pmatrix}$$

The left hand matrix is non-singular as it is a Vandermonde matrix and our c_i are not all zero so the product cannot be zero. So c must be non-zero.
Now, as G is constant, $G^{(j)}(b) = 0$ for all $b \in \mathbb{F}_p$ and $j > 0$ where

$$G^{(j)}(x) = \sum_{k=1}^{m} c_k (m-1)...(m-j)(x + a_k)^{m-1-j}$$

We now use these c_i to construct the polynomial we will actually use to obtain our bound. Let $D = d + m - 1$. We know

$$m = |A| \leq |A + B| \leq |Z_d \cup \{0\}| = d + 1$$

So

$$D = d + m - 1 \leq d + d + 1 - 1 = 2d \leq p - 1$$

as $d \leq \frac{p-1}{2}$.
Define the polynomial

$$F(x) = -c + \sum_{k=1}^{m} c_k (x + a_k)^D = \sum_{l=0}^{D} \left(\binom{D}{l} \sum_{i=1}^{m} c_i a_i^l \right) x^{D-l}$$

Now, by our choice of c_k, the coefficients vanish for $l < m - 1$ but definitely don't for $l = m - 1$. So F has degree $D - (m-1) = d$ and is non-zero.

Take an element $b \in B \cup (-A)$ and an element $a_k \in A$, and consider $(b + a_k)^d$. As $A + B \subset Z_d \cup \{0\}$, either $b + a_k \in Z_d$ or $b + a_k = 0$. If $b + a_k \in Z_d$ then $(b + a_k)^d = 1$ so $(b + a_k)^{d+1} = b + a_k$. And if $b + a_k = 0$, then clearly $(b + a_k)^{d+1} = (b + a_k)$ also. Now consider

$$\begin{aligned}
F(b) &= -c + \sum_{k=1}^{m} c_k (b + a_k)^D \\
&= -c + \sum_{k=1}^{m} c_k (b + a_k)^{d+1} (b + a_k)^{m-2} \\
&= -c + \sum_{k=1}^{m} c_k (b + a_k)^{m-1} \\
&= -c + G(b) = -c + c = 0
\end{aligned}$$

Similarly, take $1 \leq j \leq m - 2$ and consider

$$\begin{aligned}
F^{(j)}(b) &= D(D-1)...(D-1+j) \sum_{k=1}^{m} c_k (b + a_k)^{D-j} \\
&= D(D-1)...(D-1+j) \sum_{k=1}^{m} c_k (b + a_k)^{m-1-j} \\
&= \frac{D(D-1)...(D-1+j)}{(m-1)...(m-j)} G^{(j)}(b) = 0
\end{aligned}$$

So F has a root of order $m - 1$ at each $b \in B \cap (-A)$.
Similarly if $b \in B \setminus (-A)$ then $(b + a_k)^d = 1$ for all $1 \leq k \leq m$. Repeating the above argument gives us that F has a root of order m at each $b \in B \setminus (-A)$. So

$$(m-1)|B \cap (-A)| + m(|B| - |B \cap (-A)|) \leq d$$

As $m = |A|$ we can rearrange to obtain

$$|A||B| \leq d + |B \cap (-A)|$$

□

We can clearly see the relations between this proof and our proof in Section 6. We used linear programming arguments to construct a polynomial with high order zeros to obtain a bound on our set. We can note here that we were looking to bound the product of the sizes of two sets. To achieve this, we placed the elements of B in the zeros of F, and then related the size of A to the order of the zeros. This then gave us the product of $|A|$ and $|B|$ in our bound.

This result has several interesting corollaries. It is a conjecture of Sarközy that for any prime $p > 3$ there is no non-trivial additive decomposition of $Z_{\frac{p-1}{2}}$. This theorem gives the following contribution towards this conjecture:

Corollary 7.4.1. *Let p be a prime and A, B be subsets of $\mathbb{F}_p$ such that $A + B = Z_d$ for some integer d properly dividing $p - 1$. Then*

$$|A||B| = d$$

and all sums $a + b$ are distinct. If d is prime and A, B are not singleton sets then no such decomposition exists.

Proof. If $A + B = Z_d$ then $|B \cap (-A)| = 0$ as $0 \notin Z_d$ so $|A||B| \leq d$ by Theorem 7.4. But $d = |Z_d| = |A + B| \leq |A||B|$. So $|A||B| = d$. □

A corollary proved by Shakan [HP19] shows that for almost every prime, a decomposition $A + B = Z_{\frac{p-1}{2}}$ does not exist.

Corollary 7.4.2. *As $x \to \infty$, the number of primes $p \leq x$ such that there exist non-singleton sets $A, B \subset \mathbb{F}_p$ with $A + B = Z_{\frac{p-1}{2}}$ is $o(\pi(x))$ where π is the density of the primes.*

7.3 Paley Graphs

One way to consider quadratic residues is to consider Paley graphs. A Paley graph is a graph that we can draw for every finite field $\mathbb{F}_q$ when $q \equiv 1 \bmod 4$. Each node on the graph corresponds to an element $x \in \mathbb{F}_q$. We connect two nodes if they differ by a quadratic residue i.e. we include the edge $\overline{xy}$ if and only if $x - y \in Z_{\frac{q-1}{2}}$. We note that the requirement that $q \equiv 1 \bmod 4$ is crucial here as this is equivalent to -1 being a quadratic residue in $\mathbb{F}_q$ and this ensures that our relation is symmetric (i.e. $x - y \in Z_{\frac{p-1}{2}}$ if and only

if $y - x \in Z_{\frac{p-1}{2}}$).
For $p = 5$ we just obtain a pentagon. For $p = 13$ we obtain the graph in Figure 2.

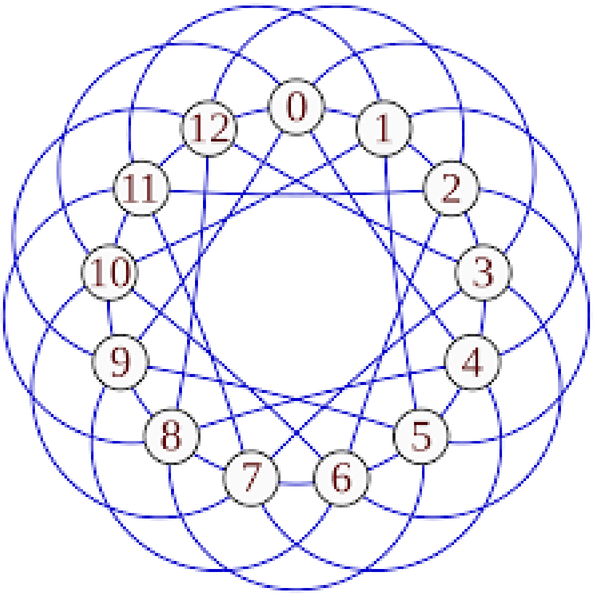

Figure 2: Paley graph with q=13 [Epp06]

For any graph we can consider a 'clique'. A clique is a connected subgraph. More formally, for a graph $G = (V, E)$ with vertices V and edges E, a clique is a subset of vertices V' such that $\overline{xy} \in E$ for all $x, y \in V'$. We can then ask what the biggest clique in a graph is? We say that the clique number of a graph G is the size (number of vertices) of its biggest clique. For a Paley graph, our theorem above gives us a strong bound on the clique number:

Theorem 7.5. *Let p be a prime congruent to* 1 *mod* 4*. Then the clique number* $\omega(G_p)$ *of the Paley graph* G_p *satisifies*

$$\omega(G_p) \leq \frac{\sqrt{2p-1}+1}{2}$$

Proof. Suppose $A \subset \mathbb{F}_p$ corresponds to a clique in G_p. So the difference of any two elements in A is a quadratic residue, so $A + (-A) \subset Z_{\frac{p-1}{2}}$. Then Theorem 7.4, we get

$$|A|| - A| \leq \frac{p-1}{2} + |A \cap -(-A)| = \frac{p-1}{2} + |A|$$

Rearranging gives us

$$|A|(|A| - 1) \leq \frac{p-1}{2}$$

Expanding and solving this quadratic equation gives us the bound:

$$\omega(G_p) \leq \frac{\sqrt{2p-1}+1}{2}$$

□

7.4 Generalisation of the bound

We proved our bound only for finite fields with a prime number of elements. But Paley graphs can be drawn for all finite fields, and in fact we can draw them for the ring $\mathbb{Z}/q\mathbb{Z}$ also. We may wonder then whether we can prove a similar bound for these cases? Indeed, does the proof of Hanson and Petridis automatically apply in these cases also?

Unfortunately our proof as it is fails in the general finite field case. The reason why is the same as why we had to use Hasse derivatives in our consideration of hyperelliptic curves. When we want zeros of high order, we show that the derivatives are zero at that point. But this only works when we differentiate at most $p-1$ times, as our example of $f(x) = x^p$ showed. In our proof we are implicitly using the following lemma [Mas16]:

Lemma 7.6. *Suppose $f^{(k)}(a) = 0$ for all $k \leq n-1 \leq p-1$ for some polnyomial $f \in \mathbb{F}_q[x]$. Then f has a root of at least order n at a.*

Our example shows that we cannot say anything stronger. This then gives us the hole in our proof. If we show that $F^{(j)}(b) = 0$ for all $j < m-1$, we can only say that F has a zero of at least order $min\{j, p\}$ at b. In the case of $\mathbb{F}_p$ this is fine as $m \leq p$. But this is not true in general $\mathbb{F}_q$.

We can ask whether Hasse derivatives can save us again? Indeed, the approach looks promising. We can find that

$$E^{(j)}(f)(b) = \binom{D}{j} \sum_{k=1}^{m} c_k (b + a_k)^{D-j}$$

and

$$E^{(j)}(G)(x) = \sum_{k=1}^{m} c_k \binom{m-1}{j} (x + a_k)^{m-1-j}$$

So then we can write:

$$E^{(j)}(f)(b) = \frac{D(D-1)...(D-1+j)}{(m-1)...(m-j)} G^{(j)}(b) = 0$$

Except sadly we can't quite, and we discover another flaw in our initial proof when extended. As m could be greater than p, the characteristic of our field, $(m-1)...(m-j)$ may be zero and we cannot divide through by it. So unfortunately the Hasse derivative does not allow us to easily extend the result.

8 Conclusion

In this report we have considered several different applications of the polynomial method. We have covered a diverse range of problems.

We began by considering the area of zero-sum problems and the Erdős-Ginzburg-Ziv theorem and how it can be proved using the polynomial method. Here we first saw some of the techniques common to the method. In the proof of the Chevalley-Warning theorem, we constructed a polynomial to count the zeros of our system. We then expanded this into monomials and argued using the pigeonhole principle to factor out factors of low enough degree. We then explicitly constructed polynomials of low enough degree to find our common zero. So here, degree was our measure of complexity.

We next considered the cap set problem and its generalisation. We introduced the concept of slice rank as our measure of complexity instead of degree, and showed how to relate it to the size of our set. We once again used a product of polynomials and properties of finite fields to count our special elements. Again, we expanded into monomials and used the pigeonhole principle to identify low degree factors. This time, we used combinatorial and probabilistic reasoning to obtain our explicit bound.

Afterwards, we saw the closely related problem of sunflower-free sets. We showed how the cap set result could be applied to obtain a bound, but then saw how using the polynomial method directly could obtain an even stronger bound. Our proof followed closely that of the cap set problem, using slice rank to measure complexity, counting our special elements in a product, expanding into monomials and using the pigeonhole principle to identify factors of low degree. This time we used combinatorial reasoning to obtain our explicit bound. We then also saw a very similar proof but using characters instead of polynomials.

Next we considered the area of hyperelliptic curves. Here we returned to using degree as our measure of complexity, but also for the first time considered the multiplicity of our zeros. We introduced the concept of Hasse derivatives and used these to ensure the zeros were of high enough mulitplicity. Also for the first time, we did not explicitly construct our polynomial. Instead, we just proved that such a polynomial must exist through the careful control of degrees and simple linear programming arguments.

Finally, we considered the problem of sumset decompositions of roots of unity. Our method was very similar to the hyperelliptic curves proof. We used degree as our bound on complexity and considered zeros of high multiplicity. We also did not explicitly construct our polynomial, but instead showed its existence using linear programming. We also considered how this result could be extended, and showed that introducing Hasse derivatives was not enough to generalise the result.

There are of course other applications of the polynomial method we could have considered, such as the Combinatorial Nullstellensatz, which generalises the factor theorem to polynomials in multiple dimensions [Tao13].

Theorem 8.1 (**Combinatorial Nullstellensatz**). *Let $\mathbb{F}$ be a field and $d_1, ..., d_n \geq 0$ integers. Suppose $P \in F[x_1, ..., x_n]$ is a polynomial of degree $d_1 + ... + d_n$ with a non-zero coefficient of $x_1^{d_1}...x_n^{d_n}$. Then P cannot vanish on any set of the form $E_1 \times ... \times E_n$ with $E_1, ..., E_n \subset \mathbb{F}$ and $|E_i| > d_i$ for $i = 1, ..., n$.*

Another example of the polynomial method is in proving bounds on Kakeya sets. A set $K \subset \mathbb{F}_q^n$ is a Kakeya set if it contains lines in every direction i.e. for any $x \in \mathbb{F}_q^n$ there exists $y \in K$ such that $L_{y,x} = \{y + rx | a \in \mathbb{F}_q^n\} \subset K$. The polynomial method can then be used to prove the following [Dvi09]:

Theorem 8.2. *Suppose $K \subset \mathbb{F}_q^n$ is a Kakeya set. Then $|K| \geq C_n q^n$ for some constant C_n depending only on n.*

We could also have developed the notion of slice rank further, and considered how our result for hyperelliptic curves relates to the Hasse-Weill bound.

We have demonstrated that the polynomial method can be applied in a wide variety of contexts. It has allowed us to prove strong bounds, often improving considerably on previous results as well proving various conjectures. This shows the power and versatility of the method. We have seen that the polynomial method is a powerful method of proof that can be applied in a diverse range of problems, and we can be sure it will continue to be applied effectively as we improve our understanding of it.

Bibliography

[Tao13] Terence Tao. *Algebraic combinatorial geometry: the polynomial method in arithmetic combinatorics, incidence combinatorics, and number theory.* 2013. arXiv: `1310.6482 [math.CO]`.

[AD93] Noga Alon and Moshe Dubiner. "Zero-sum sets of prescribed size". In: (1993).

[CLP16] Ernie Croot, Vsevolod Lev, and Peter Pach. *Progression-free sets in Z_4^n are exponentially small.* 2016. arXiv: `1605.01506 [math.NT]`.

[EG16] Jordan S. Ellenberg and Dion Gijswijt. *On large subsets of F_q^n with no three-term arithmetic progression.* 2016. arXiv: `1605.09223 [math.CO]`.

[Tao] Terrence Tao. *A symmetric formulation of the Croot-Lev-Pach-Ellenberg-Gijswijt capset bound.* URL: `https://terrytao.wordpress.com/2016/05/18/a-symmetric-formulation-of-the-croot-lev-pach-ellenberg-gijswijt-capset-bound/`.

[NS16] Eric Naslund and William F. Sawin. *Upper bounds for sunflower-free sets.* 2016. arXiv: `1606.09575 [math.CO]`.

[IKS04] H. Iwaniec, E. Kowalski, and American Mathematical Society. *Analytic Number Theory.* American Mathematical Society Colloquium Publications. American Mathematical Society, 2004. Chap. 11. ISBN: 9780821836330. URL: `https://books.google.co.uk/books?id=8i7wpzjSWrIC`.

[HP19] Brandon Hanson and Giorgis Petridis. *Refined Estimates Concerning Sumsets Contained in the Roots of Unity.* 2019. arXiv: `1905.09134 [math.NT]`.

[Hana] Yunghsiang S. Han. *Introduction to Finite Fields.* URL: `https://web.ntpu.edu.tw/~yshan/algebra.pdf`.

[Rot98] Joseph Rotman. "Polynomial Rings over Fields". In: *Galois Theory.* New York, NY: Springer New York, 1998, pp. 24–31. ISBN: 978-1-4612-0617-0. DOI: `10.1007/978-1-4612-0617-0_6`. URL: `https://doi.org/10.1007/978-1-4612-0617-0_6`.

[FO96] Carlos Flores and Oscar Ordaz. "On the Erds-Ginzburg-Ziv theorem". In: *Discrete Mathematics* 152 (1996), pp. 321–324.

[GG06] Weidong Gao and Alfred Geroldinger. "Zero-sum problems in finite abelian groups: A survey". In: *Expositiones Mathematicae* 24.4 (2006), pp. 337–369. ISSN: 0723-0869. DOI: `https://doi.org/10.1016/j.exmath.2006.07.002`. URL: `http://www.sciencedirect.com/science/article/pii/S0723086906000351`.

[Eng05] Miles at English Wikipedia. *Set Game Cards.* 2005. URL: `https://en.wikipedia.org/wiki/Set_(card_game)#/media/File:Set-game-cards.png`.

[Rah] Mustazee Rahman. *Roth's theorem on 3-term arithmetic progressions.* URL: `https://pdfs.semanticscholar.org/34d5/e5d802d1107b68f1aa76dff994e8b23341c2.pdf`.

[Ede04] Yves Edel. “Extensions of Generalized Product Caps”. In: *Designs, Codes and Cryptography* 31 (2004), pp. 5–14.

[TS] Terrence Tao and Will Sawin. *Notes on the ‘slice rank‘ of tensors.* URL: https://terrytao.wordpress.com/2016/08/24/notes-on-the-slice-rank-of-tensors/.

[JL01] Gordon James and Martin Liebeck. “Character tables and orthogonality relations”. In: *Representations and Characters of Groups.* 2nd ed. Cambridge University Press, 2001, pp. 159–167. DOI: 10.1017/CBO9780511814532.017.

[Hanb] Brandon Hanson. *Stepanov’s Method for Hyperelliptic Curves.* URL: http://personal.psu.edu/rcv4/stepanovs-method-hyperelliptic.pdf.

[Mas16] David Masser. *Auxiliary Polynomials in Number Theory.* Cambridge Tracts in Mathematics. Cambridge University Press, 2016. Chap. 6. DOI: 10.1017/CBO9781107448018.

[Tao08] Terence Tao. *The sum-product phenomenon in arbitrary rings.* 2008. arXiv: 0806.2497 [math.CO].

[Epp06] David Eppstein. *Paley Graph of order 13.* 2006. URL: https://en.wikipedia.org/wiki/Paley_graph#/media/File:Paley13.svg.

[Dvi09] Zeev Dvir. “On the size of Kakeya sets in finite fields”. In: *J. Amer. Math. Soc.* 22 (2009), pp. 1093–1097. DOI: https://doi.org/10.1090/S0894-0347-08-00607-3.