# Expanding Polynomials

A 2021 Polymath Jr. Project*

**Abstract**

This document describes a research project for the Polymath REU program of the summer of 2021. Please do not share it with anyone who is not part of the program.

**How to read this document.** You might like to first skim this document without dwelling too much on the technicalities. You can also focus on the specific sections that seem more relevant to you. For example, Section 5 describes a problem that fits people who are comfortable with basic abstract algebra. Section 4 describes a problem that fits people who like to play with geometric shapes.

If you decide to be part of the project, then you should carefully read this document and solve the exercises. If you are having a hard time with an exercise, ask for a hint or help. Ask for a solution only after spending enough time trying to solve the problem. Looking at solutions without spending time thinking on the problem tends not too be very helpful.

It might be hard to start some of the open problems on your own. We recommend chatting with a mentor when starting to work on an open problem. Even if you want to pursue this problem on your own.

This document relies on asymptotic notation, such as $O(n)$, $\Omega(n)$, and $\Theta(n)$. If you are not familiar with such notation, you can find a brief explanation at the bottom of the document. Alternatively, there are many online resources that introduce asymptotic notation, and you are also welcome to ask mentors about it.

*This project is mentored by Sara Fish, Sam Mansfield, Adam Sheffer, Sophie Stevens, and Audie Warren.

# Contents



# 1 Introduction: sum sets, product sets, and expanding polynomials

This section is an introduction to some objects and ideas that we will play with in this project. These objects and ideas belong to the mathematical subfield of *additive combinatorics*. This subfield lies at the intersection of combinatorics, number theory, and analysis, although we will not deal with the analysis part. After this introduction, we discuss the problems of our project in the following sections.

**Sum sets.** Let $A$ be a set of real numbers. The *sum set* of $A$ is defined as

$$A + A = \{a + a' \ : \ a, a' \in A\}.$$

We also consider the cases where $a = a'$. For example, if $A = \{1, 5, 9, 10\}$ then

$$A + A = \{2, 6, 10, 11, 14, 15, 18, 19, 20\}.$$

Recall that a set contains every element at most once. Thus, every sum appears at most once in $A + A$, even if the sum has many representations as $a + a'$.

We set $n = |A|$. The number of ways to choose a pair of elements $(a, a') \in A^2$ with $a \leq a'$ is $\binom{n+1}{2} = \frac{n^2+n}{2}$. This in turn implies that $|A + A| \leq \frac{n^2+n}{2}$, and this is tight when all the sums are distinct. If we build $A$ by taking elements of $\mathbb{R}$ at random, then we expect $|A + A|$ to be very close to this upper bound, since the probability of $a_1 + a_2 = a_3 + a_4$ is very small (where $a_1, a_2, a_3, a_4 \in A$). In other words, most sets $A$ of $n$ real numbers satisfy $|A + A| = \Theta(n^2)$.

Consider the set $A = \{1, 2, \ldots, n\}$, and note that$|A + A| = |\{2, 3, 4, \ldots, 2n\}| = 2|A| - 1$. The same bound holds whenever $A$ is an arithmetic progression. One of the main problems of additive combinatorics is characterizing the sets $A$ for which $|A+A| = \Theta(n)$. In other words, we are interested in sets $A$ that satisfy $|A+A| \leq k|A|$ for some constant $k$, when $|A|$ is asymptotically large. Intuitively, we wish to find sets that have a good additive structure that leads to a small sum set. (This problem is studied over many other fields and rings.) As a warmup, we begin with the following claim.

**Claim 1.1.** *Every set $A \subset \mathbb{R}$ of size $n$ satisfies the bound $|A + A| \geq 2n - 1$.*

*Proof.* Denote the elements of $A$ as $a_1 < a_2 < \cdots < a_n$. Then $A + A$ contains the following $2n - 1$ distinct elements:

$$a_1 + a_1 < a_1 + a_2 < a_1 + a_3 < \cdots < a_1 + a_n < a_2 + a_n < a_3 + a_n < \cdots < a_n + a_n.$$

□

Claim 1.1 establishes that, in $\mathbb{R}$, arithmetic progressions have minimum-sized sum sets. The following set is not an arithmetic progression, although it is defined in a similar way.

$$A = \left\{3k_1 + 100k_2 : k_1 \in \{1, 2, 3\} \text{ and } k_2 \in \{1, 2, 3, \ldots, n/3\}\right\}. \tag{1}$$

For simplicity, we assume that $n$ is a multiple of three. Then $|A| = n$ and $|A+A| = 5(2n/3 - 1) = 10n/3 - 5$. Thus, we still have that $|A + A| = \Theta(n)$. A *generalized arithmetic progression* of dimension $d$ is defined as

$$\left\{a + \sum_{j=1}^{d} k_j b_j : a, b_1, \ldots, b_d \in \mathbb{R} \text{ and with integer } 0 \leq k_j \leq n_j - 1 \text{ for every } 1 \leq j \leq d\right\}.$$

An arithmetic progression is a generalized arithmetic progression of dimension 1, and the set in (1) is a generalized arithmetic progression of dimension 2. The size of a generalized arithmetic progression of dimension $d$ is at most $n = n_1 n_2 \cdots n_d$, and it is not difficult to verify that it has a sumset of size smaller than $2^d n$. The following result characterizes the sets $A$ of $n$ real numbers that satisfy $|A + A| = \Theta(n)$.

**Theorem 1.2 (Freiman's theorem over the reals).** *Let $A \subset \mathbb{R}$ be a finite set with $|A + A| \leq k|A|$ for some constant $k$. Then $A$ is contained in a generalized arithmetic progression of size at most $cn$ and dimension at most $d$. Both $c$ and $d$ depend on $k$ but not on $|A|$.*

It is a major open problem to show that $c$ and $d$ are bounded by some polynomial in $k$. This is sometimes called *the polynomial Freiman-Ruzsa conjecture* over $\mathbb{R}$. It is too difficult for a summer research project.

**Exercise 1.** *Let $A$ be a set of $n$ real numbers.*[1] *Prove that there exists a subset $S \subset A$ such that $|S| = \Theta(n^{1/3})$ and every sum from $S+S$ has exactly one representation. In other words, every pair $(a, a') \in S^2$ has a different sum $a + a'$. (Hint: Repeatedly add to $S$ numbers that do not lead to multiple representations. How many numbers can you add without getting stuck?)*

**Product sets.** Let $A$ be a set of $n$ real numbers. The *product set* of $A$ is

$$AA = \{a \cdot a' : a, a' \in A\}.$$

Products sets are similar to sum sets. For example, by repeating the argument for $|A + A|$, we obtained that $|AA| \leq \frac{n^2+n}{2}$. When $A$ is a geometric progression, we get that $|AA| = 2n - 1$. For example, if $A = \{2^1, 2^2, \ldots, 2^n\}$ then

$$|AA| = |\{2^2, 2^3, 2^4, \ldots, 2^{2n}\}| = 2n - 1.$$

In this case we can do a bit better than $2n - 1$.

**Exercise 2.**
*(a) Find a set $A$ of $n$ reals for which $|AA| = 2n - 2$.*

*(b) For odd $n$, find a set $A$ of $n$ reals for which $|AA| = 2n - 3$.*

*(c) Prove that every set $A$ of $n$ reals satisfies $|AA| \geq 2n - 3$.*

We can similarly define generalized geometric progressions and a multiplicative variant of Theorem 1.2.

**Sums versus products.** While most sets $A$ of $n$ real numbers satisfy $|A + A| = \Theta(n^2)$, there are sets with an additive structure that leads to $|A + A| = \Theta(n)$. Similarly, while most sets $A$ satisfy $|AA| = \Theta(n^2)$, there are sets with a multiplicative structure that leads to $|AA| = \Theta(n)$. Erdős and Szemerédi [3] asked whether a set can have both an additive structure and a multiplicative structure. They made the following conjecture (originally only for integers).

**Conjecture 1.3.** *For any $\varepsilon > 0$ there exists $n_0$ that satisfies the following. For every $n > n_0$ and every set $A$ of $n$ real numbers, we have that*

$$\max\{|A + A|, |AA|\} = \Omega(n^{2-\varepsilon}).$$

Intuitively, Conjecture 1.3 suggests that at least one of $A + A$ and $AA$ has size arbitrarily close to the maximum $\Theta(|A|^2)$. One may say that this is a deep question about the nature of addition and multiplication.

[1] This exercise is more difficult than the following ones. You might want to first skip it.

We are far from proving Conjecture 1.3, and it is too difficult for an undergraduate summer project. The current best bound is

$$\max\{|A+A|, |AA|\} = \Omega(n^{4/3+2/1167}).$$

This bound is by one of the project mentors: Sophie Stevens. See [15].

**Expanding polynoimals.** Consider a function $f : \mathbb{R}^d \to \mathbb{R}$ and let $A$ be a set of $n$ real numbers. Generalizing the above, we write

$$f(A_1, \ldots, A_d) = \{f(a_1, \ldots, a_d) : \ a_i \in A_i, \ i = 1, \ldots, d\}.$$

For example, we can think of $A + A$ as defined by $f(x, y) = x + y$. We can think of $AA$ as defined by $f(x, y) = xy$.

If $A_1 = A_2 = \cdots = A_d$, then we write $f(A, \ldots, A) = f(A)$, remembering that $f$ is not necessarily univariate.

We say that a polynomial is an *expander* if $f(A)$ is large for every set $A$. We also refer to $f$ as an *expanding polynomial*. The meaning of *being large* is not well-defined, and changes according to the context, typically taking into account the number of variables involved in $f$. For example, we say that $f(x, y) = x+y$ and $f(x, y) = xy$ are not expanders, since both may satisfy $|f(A)| = \Theta(n)$. On the other hand, consider $g(x, y, z) = x + yz$. For every set $A$ of $n$ numbers, we have that $|g(A)| = \Omega(n^{3/2})$ (see Section 2). We thus say that $g(x, y, z)$ is an expander.

**Exercise 3.**
*(a) Show that $f(x, y, z) = x + y + z$ is not an expander.*

*(b) Show that $f(x, y, z) = x^2 + y^2 + z^2$ is not an expander.*

In this project we will study a variety of problems that involve expanding polynomials. This includes studying expanders over finite fields, studying applications of expanders, trying to extend common proofs for expander results, and more. Each of the following sections describes one aspect of the project. As a project participant, you may wish to focus on only one of those aspects of the project, or on multiple ones.

# 2 Problem 1: Incidences — a proof technique

We begin this section by studying the proof of an older sum-product bound of Elekes [1]. One might say that this result was the biggest breakthrough in the history of the problem. Elekes discovered that we can obtain stronger bounds for the problem by studying it geometrically. Since Elekes's paper appeared, almost all of the sum-product works have followed this approach.

We consider a point set $\mathcal{P}$ and a set of lines $\mathcal{L}$, both in $\mathbb{R}^2$. An *incidence* is a pair $(p, \ell) \in \mathcal{P} \times \mathcal{L}$ with the point $p$ being on the line $\ell$. For example, see Figure 1.

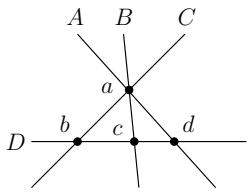


Figure 1: A configuration of four points, four lines, and nine incidences.

We denote by $I(\mathcal{P}, \mathcal{L})$ the number of incidences in $\mathcal{P} \times \mathcal{L}$. The following theorem is probably the best known incidence result.

**Theorem 2.1 (The Szemerédi-Trotter theorem).** *Let $\mathcal{P}$ be a set of $m$ points and let $\mathcal{L}$ be a set of $n$ lines, both in $\mathbb{R}^2$. Then $I(\mathcal{P}, \mathcal{L}) = O(m^{2/3}n^{2/3} + m + n)$.*

This theorem is tight, up to the implied constant: For every $m$ and $n$ there are point-line constructions with $\Theta(m^{2/3}n^{2/3} + m + n)$ incidences. The term $m^{2/3}n^{2/3}$ dominates the bound in the most interesting range of $m = O(n^2)$ and $m = \Omega(\sqrt{n})$. We are now ready to derive Elekes' sum-product bound.

**Theorem 2.2.** *Every set $A$ of $n$ real numbers satisfies*

$$\max\{|A + A|, |AA|\} = \Omega\left(n^{5/4}\right).$$

*Proof.* We consider the planar point set

$$\mathcal{P} = \{(c, d) : c \in A + A \quad \text{and} \quad d \in AA\}.$$

We define a line in $\mathbb{R}^2$ with an equation of the form $y = ax + b$. The line is the set of points $(x, y)$ in $\mathbb{R}^2$ that satisfy this equation. We consider the set of lines

$$\mathcal{L} = \{y = a(x - a') : a, a' \in A\}$$

We note that $|\mathcal{L}| = n^2$ and that $|\mathcal{P}| = |A+A| \cdot |AA|$. Figure 2 depicts such a point-line construction.

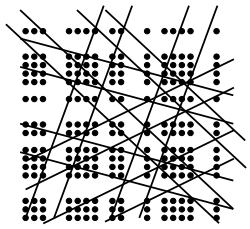

Figure 2: The point-line construction in Elekes' proof: A cartesian product of points and $n$ sets of $n$ parallel lines.

The proof is based on estimating the number of incidences $I(\mathcal{P}, \mathcal{L})$ in two different ways. First, we consider a line $\ell \in \mathcal{L}$ that is defined by $y = a(x - a')$. We note that $\ell$

contains every point of $\mathcal{P}$ of the form $(a'+b, ab)$ for every $b \in A$. That is, $\ell$ contains at least $n$ points of $\mathcal{P}$. By summing this over every $\ell \in \mathcal{L}$, we obtain that

$$I(\mathcal{P}, \mathcal{L}) \geq |\mathcal{L}| \cdot n = n^3. \tag{2}$$

On the other hand, Theorem 2.1 implies that

$$\begin{aligned} I(\mathcal{P}, \mathcal{L}) &= O\left(|\mathcal{P}|^{2/3}|\mathcal{L}|^{2/3} + |\mathcal{P}| + |\mathcal{L}|\right) \\ &= O\left(|A+A|^{2/3}|AA|^{2/3}n^{4/3} + |A+A| \cdot |AA| + n^2\right). \end{aligned} \tag{3}$$

By combining (2) and (3), we obtain that

$$n^3 = O\left(|A+A|^{2/3}|AA|^{2/3}n^{4/3} + |A+A| \cdot |AA| + n^2\right).$$

There are three terms in the right-hand side above. We separately check three cases, assuming that a different term dominates the bound in each. This leads to

$$|A+A| \cdot |AA| = \Omega\left(n^{5/2}\right).$$

It is thus impossible for both $|A+A|$ and $|AA|$ to be asymptotically smaller than $n^{5/4}$, which completes the proof of the theorem. □

To prove Theorem 2.2, we used a common combinatorial method called double counting. In this method we introduce and bound some quantity $X$ in two different ways, and then compare the two bounds. This leads to new information that does not involve $X$. In the proof of Theorem 2.2, we derived upper and lower bounds for $I(\mathcal{P}, \mathcal{L})$.

Similar incidence arguments are used to derive many bounds for expanding polynomials.

**Exercise 4.** *Let $A$ be a set of $n$ real numbers.*

*(a) Prove that*

$$|A + AA| = \Omega(n^{3/2}).$$

*(Hint: Consider the point set $A \times (A + AA)$.)*

*(b) Prove that*

$$|A(A+A)| = \Omega(n^{3/2}).$$

**Exercise 5.** *Let $n$ be a huge number.*

*(a) Prove that there exists a set $A$ of $n$ real numbers that satisfies $|A+AA| = O(n^2)$.*

*(b) Prove that there exists a set $A$ of $n$ real numbers that satisfies $|A(A+A)| = O(n^2)$.*

One problem in our project would be to explore various recent incidence bounds, trying to use those to obtain new bounds for expanding polynomials. For example, we are interested in the following recent results:

- An improved incidence bound when the points and lines do not form a specific structure: [8].
- An improved incidence bound when the point set is a lattice with additional restrictions: [14, Theorem 8].
- A much more involved recent result: [16].

Consider a point set $\mathcal{P} \subset \mathbb{R}^2$ and an integer $r \geq 2$. A line $\ell \subset \mathbb{R}^2$ is *$r$-rich* if $\ell$ contains at least $r$ points of $\mathcal{P}$. Similarly, given a set of lines $\mathcal{L}$ in $\mathbb{R}^2$, a point $p \in \mathbb{R}^2$ is called *$k$-rich* if at least $k$ lines from $\mathcal{L}$ pass through $p$. The following theorem is considered as equivalent to Theorem 2.1, since each can be derived from the other by using basic arguments.

**Theorem 2.3.** *Let $\mathcal{P} \subset \mathbb{R}^2$ be a set of $n$ points and let $r \geq 2$. Then the number of $r$-rich lines with respect to $\mathcal{P}$ is*

$$O\left(\frac{n^2}{r^3} + \frac{n}{r}\right).$$

**Exercise 6.** *Change the proof of Theorem 2.2, so that it relies on Theorem 2.3 instead of Theorem 2.1.*

# 3 Problem 2: Exceptional cases over the reals

Consider a polynomial $f \in \mathbb{R}[x, y]$. We say that $f$ is of a *special form* if there exist polynomials $g, h_1, h_2 \in \mathbb{R}[z]$ such that

$$f(x, y) = g(h_1(x) + h_2(y)) \qquad \text{or } f(x, y) = g(h_1(x) \cdot h_2(y)).$$

For example, $f(x, y) = x^2 + 4xy + 4y^2$ is of a special form, since we can write $f(x, y) = g(h_1(x) + h_2(y))$ with $g(z) = z^2$, $h_1(z) = z$ and $h_2(z) = 2z$.

**Exercise 7.** *Which of the following polynomials is of a special form? Explain your answers.*

*(a)* $f_1(x, y) = x^2 + 5x + 2xy + 5y + y^2 + 5y - \pi$.

*(b)* $f_2(x, y) = x^2 + y^2$.

The following very general result over $\mathbb{R}$ is from [11].

**Theorem 3.1.** *Consider a set $A$ of $n$ real numbers and a constant-degree polynomial $f \in \mathbb{R}[x, y]$ that is not of a special form. Then*

$$|f(A, A)| = \Omega(n^{4/3}).$$

Theorem 3.1 characterizes the expanding polynomials over the reals in two variables. This is not a perfect characterization, since some polynomials are of a special form and still expand. For example, $f(x, y) = x(y + 1)$ has a special form and expands. (It is an open problem to determine whether $|f(A, A)| = \Omega(n^{4/3})$ for this $f$.)

The *growth* of a function $f(x_1, \ldots, x_d)$ is the largest $\alpha \in \mathbb{R}$ that satisfies the following: For every set $A$ of $n$ real numbers, we have that $f(A, \ldots, A) = \Omega(n^\alpha)$. We note that $f$ expands if and only if the growth of $f$ is larger than one. Theorem 3.1 states every constant-degree $f \in \mathbb{R}[x, y]$ that is not of a special form has a growth of at least $4/3$.

We will explore questions such as:

- Which polynomials of a special form expand? Are there such polynomials with growth $1 < \alpha < 4/3$ ?
- What can be said about functions that are not polynomials? For example, what can be said about rational functions? (A bivariate function is rational if it is of the form $g(x, y)/h(x, y)$ for polynomials $g, h \in \mathbb{R}[x, y]$.)
- Which polynomials $f \in \mathbb{R}[x, y]$ have growth larger than $4/3$?
- Which polynomials $f \in \mathbb{R}[x, y, z]$ have growth larger than $3/2$? Growth smaller than $3/2$ but larger than 1?

# 4 Problem 3: Distinct distances between curves in $\mathbb{R}^3$

This section is about an application of expanding polynomials. The *distinct distances* problem was introduced by Erdős [2] in 1946. For a set $\mathcal{P} \subset \mathbb{R}^2$, let $D(\mathcal{P})$ denote the set of distances spanned by pairs of points in $\mathcal{P}$. Every distance appears in $D(\mathcal{P})$ at most once, no matter how many pairs of points span it. This is why we refer to $D(\mathcal{P})$ as the *set of distinct distances of* $\mathcal{P}$. See Figure 3 for an example. The distinct distances problem asks:

> What is the minimum number of distinct distances that can be determined by a set of $n$ points in $\mathbb{R}^2$.

In other words, the problem asks to find $\min_{|\mathcal{P}|=n} |D(\mathcal{P})|$. We denote this quantity as $d(n)$.

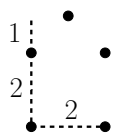


Figure 3: When $\mathcal{P}$ consists of these five points, $D(\mathcal{P}) = \{\sqrt{2}, 2, \sqrt{8}, \sqrt{10}\}$.

A set of $n$ points that are equally spaced on a line determines $n-1$ distinct distances. This implies that $d(n) \leq n-1$. An asymptotically better bound appeared in Erdős's original paper. Erdős considered the set

$$\mathcal{P} = \{(a,b) \in \mathbb{Z}^2 : 0 \leq a, b < \sqrt{n}\}.$$

The number of distinct distances determined by this set is an immediate corollary of the following result from number theory.

**Theorem 4.1. (Landau and Ramanujan)** *The number of integers smaller than $n$ that can be written as $a^2+b^2$ with $a, b \in \mathbb{Z}$ is $\Theta(n/\sqrt{\log n})$.*

Every distance in $\mathcal{P}$ is the square root of a sum of two squares between 0 and $n-1$. By Theorem 4.1, there are $\Theta(n/\sqrt{\log n})$ such numbers. Thus, $D(\mathcal{P}) = \Theta(n/\sqrt{\log n})$. This set of points immediately implies the following theorem of Erdős:

**Theorem 4.2.** $d(n) = O(n/\sqrt{\log n})$.

Erdős conjectured that $d(n) = \Theta(n/\sqrt{\log n})$, but deriving lower bounds for $d(n)$ turned out to be significantly more difficult. Better and better lower bounds for $d(n)$ appeared over the decades. In 2010, Guth and Katz [6] proved that $d(n) = \Omega(n/\log n)$, almost settling Erdős's conjecture. The proof of Guth and Katz is a deep result that combines tools from multiple mathematical subfields. It is not a good fit for an undergraduate summer project.

**Distinct distances between points on curves.** In the *bipartite* distinct distances problem we have two point sets $\mathcal{P}_1$ and $\mathcal{P}_2$. We denote by $D(\mathcal{P}_1, \mathcal{P}_2)$ the number of distinct distances spanned by pairs from $\mathcal{P}_1 \times \mathcal{P}_2$. That is, we ignore pairs of points from the same set. In other words,

$$D(\mathcal{P}_1, \mathcal{P}_2) = \left|\{|pq| \ : \ p \in \mathcal{P}_1, q \in \mathcal{P}_2\}\right|$$

(where $|pq|$ denotes the distance between the points $p$ and $q$).

We consider the case where $\mathcal{P}_1$ is a set of $m$ points on a line $\ell_1$ and $\mathcal{P}_2$ is a set of $n$ points on a different line $\ell_2$. Without loss of generality, we assume that $n \geq m$. When the two lines are parallel or orthogonal, the points can be arranged so that $D(\mathcal{P}_1, \mathcal{P}_2) = \Theta(n)$. Such constructions are depicted in Figure 4.

**Exercise 8.** *Let $\mathcal{P}_1$ be a set of $m$ points on a line $\ell_1$. Let $\mathcal{P}_2$ be a set of $n$ points on a line $\ell_2$. The lines $\ell_1$ and $\ell_2$ are either parallel or orthogonal, but not identical. Assume that $n \geq m$ and prove that $D(\mathcal{P}_1, \mathcal{P}_2) = \Omega(n)$.*

When $\ell_1$ and $\ell_2$ are neither parallel nor orthogonal, the problem significantly changes. The following result is from [18].

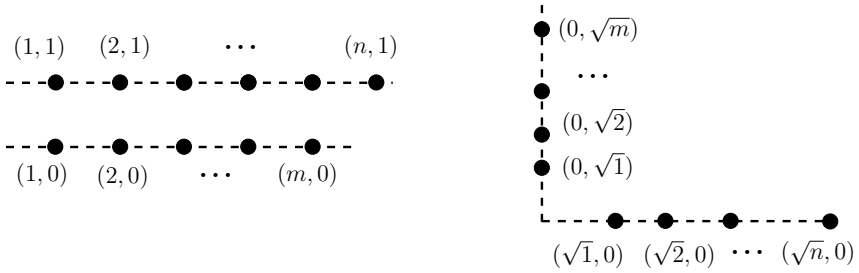


Figure 4: When the lines are either parallel or orthogonal, the points can be arranged so that $D(\mathcal{P}_1, \mathcal{P}_2) = \Theta(n)$. The distances in the parallel case are $\{2, 3, 4, \ldots, n\}$. The distances in the orthogonal case are $\{2, 3, 4, \ldots, n + m\}$.

**Theorem 4.3.** *Let $\mathcal{P}_1$ be a set of $m$ points on a line $\ell_1$. Let $\mathcal{P}_2$ be a set of $n$ points on a line $\ell_2$. The lines $\ell_1$ and $\ell_2$ are neither parallel nor orthogonal. Then*

$$D(\mathcal{P}_1, \mathcal{P}_2) = \Omega\left(\min\left\{m^{2/3}n^{2/3}, n^2, m^2\right\}\right).$$

For more intuition, consider the common case where $m = n$. When the two lines are parallel or orthogonal, it is possible that $D(\mathcal{P}_1, \mathcal{P}_2) = \Theta(n)$. Otherwise, we have that $D(\mathcal{P}_1, \mathcal{P}_2) = \Omega(n^{4/3})$.

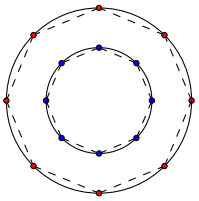

Figure 5: The case of two concentric circles.

We have a similar situation for circles. Let $\mathcal{P}_1$ be a set of $m$ points on a circle $C_1$. Let $\mathcal{P}_2$ be a set of $n$ points on a circle $C_2$, where $n \geq m$. If $C_1$ and $C_2$ are concentric, then it is possible that that $D(\mathcal{P}_1, \mathcal{P}_2) = \Theta(n)$. See Figure 5. Otherwise, we have that

$$D(\mathcal{P}_1, \mathcal{P}_2) = \Omega\left(\min\left\{m^{2/3}n^{2/3}, n^2, m^2\right\}\right). \tag{4}$$

**The problem: Distinct distances between curves in $\mathbb{R}^3$.** The work [8] considered the case of distinct distances between points on two circles in $\mathbb{R}^3$. In this case, there are more involved cases where one can obtain a linear number of distinct distances. These cases were characterized and the bound (4) was obtained for all other cases.

The proof technique of [8] is as follows. We consider a polynomial $f$ that describes the distance between two points, one on each circle. This polynomial has a huge number of monomials and is too complicated to be studied by hand. Instead, we construct $f$ with a Mathematica program. We then consider a variant of Theorem

3.1, stating that $f$ is either of a special form or expands. By inspecting several monomials of $f$ with Mathematica, we get that $f$ is of a special form if and only if the circles form one of the constructions that allow for a linear number of distances.

In this problem, we wish to study variants of the above. For example, we can study distinct distances between two parabolas in $\mathbb{R}^3$, or between one point and one parabola. Unlike some of the other problems, this problem has a well-defined plan:

- Read Sections 4 and 5 of [8]. The other sections of that paper are not relevant.
- Choose a variant of the problem and find the new polynomial that is obtained in this case.
- Adapt the original proof to the new polynomial and study the cases when it is of a special form.

???more exercises in this section???

# 5 Problem 4: Expanding Polynomials in $\mathbb{F}_q$

The problem in this section might be difficult to read without basic familiarity with abstract algebra. We begin by recalling many of the algebraic notions that we require. If you are not familiar with abstract algebra but still wish to work on this problem, the mentors can help you learn the relevant material.

**Finite fields.** A field is *finite* if it contains finitely many elements. The *order* of a finite field is the number of elements in that field. There is a finite field of order $q \in \mathbb{N} \setminus \{0\}$ if and only if $q = p^r$ for some prime $p$ and positive integer $r$. We say that such a $q$ is a *prime power*. For every such $q$ there is a unique field of that order, up to isomorphisms. We denote this finite field as $\mathbb{F}_q$ and as $\mathbb{F}_{p^r}$.

For a prime $p$, we can think of $\mathbb{F}_p$ as the set of integers $\{0, 1, \ldots, p-1\}$ under addition and multiplication mod $p$. To describe a field $\mathbb{F}_{p^r}$ with prime $p$ and integer $r > 1$, we consider an irreducible polynomial $f \in \mathbb{F}_p[x]$ of degree $r$. We think of $\mathbb{F}_{p^r}$ as the set of polynomials in $\mathbb{F}_p[x]$ under addition and multiplication modulo $f$. That is, when multiplying two polynomials in $\mathbb{F}_p[x]$ we first perform the standard polynomial multiplication, then replace each coefficient with its value mod $p$, and finally divide by $f$ and take the remainder. For example, by setting $f = x^2 + 1$ we get that $\mathbb{F}_9 = \{0, 1, 2, x, x+1, x+2, 2x, 2x+1, 2x+2\}$. The multiplicative group of $\mathbb{F}_{p^r}$ is cyclic, and the additive group of $\mathbb{F}_{p^r}$ is the direct product of $r$ cyclic groups of order $p$.

In this problem we work in the vector space $\mathbb{F}_q^d$, for some prime power $q$ and integer $d \geq 2$. We borrow the standard geometric notation from $\mathbb{R}^d$. For example, we refer to $\mathbb{F}_q^2$ as a finite plane and to an element of $\mathbb{F}_q^2$ as a point in the plane. Let $0_d$ be a vector of $d$ zeros. In other words, $0_d$ is the origin of a $d$-dimensional space.

A line in $\mathbb{F}_q^2$ is the zero set of a linear polynomial in $\mathbb{F}_q[x, y]$. Figure 6 shows that such a line might have a non-intuitive behavior. As in $\mathbb{R}^2$, an equivalent definition of

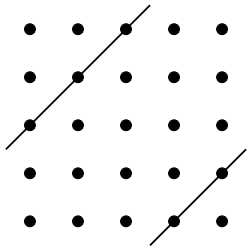

Figure 6: The line in $\mathbb{F}_5^2$ defined by $y \equiv x + 2$.

a line is $\{u + tv : \ t \in \mathbb{F}_q\}$ where $u, v \in \mathbb{F}_q^2$ and $v \neq 0_d$. This equivalent definition of a line is often easier to work with in a space $\mathbb{F}_q^d$ of dimension $d > 2$.

**Exercise 9.** *Let $q > 1$ be a prime power.*
*(a) Let $\ell \subset \mathbb{F}_q^2$ defined by $y = ax + b$, for some $a, b \in \mathbb{F}_q$. Find $u, v \in \mathbb{F}_q^2$ such that $\ell = \{u + tv : \ t \in \mathbb{F}_q\}$.*
*(b) Let $\ell = \{u + tv : \ t \in \mathbb{F}_q\} \subset \mathbb{F}_q^2$ where $u, v \in \mathbb{F}_q^2$. For which $u, v$ do there exist $a, b \in \mathbb{F}_q$ such that $\ell$ is defined by $y = ax + b$?*

**Expanding polynomials in $\mathbb{F}_q$.** When working in $\mathbb{F}_{p^r}^d$, our arithmetic is as described above: We first fix an irreducible polynomial $f \in \mathbb{F}_p[x_1, \ldots, x_d]$ of degree $r$. After each arithmetic operation, we take each coefficient modulo $p$ and the entire expression modulo $f$.

Expanding polynomials have been significantly studied over $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{F}_p$. However, much less is known over $\mathbb{F}_{p^r}$, where $p$ is prime and $r > 1$. For simplicity, we refer to this case as working over $\mathbb{F}_q$.

Since only a few works study expanding polynomials over $\mathbb{F}_q$, we believe that this is a good area to study. In this problem, we will play with expanding polynomials over $\mathbb{F}_q$, make conjectures, and try to prove results. We now provide a few warm-up examples for working on this problem.

The following exercise shows that polynomials can behave rather differently in $\mathbb{F}_q$.

**Exercise 10.** *We are given some function $f : \mathbb{F}_q^2 \to \mathbb{F}_q$. Show that there exists $A \subseteq \mathbb{F}_q$ with $|A| \to \infty$ as $q \to \infty$ that satisfies $|f(A, A)| \leq |A|$.*

The following is the current best bound for point–line incidences in $\mathbb{F}_q^2$. (Once again, by our own Sophie Stevens! [20])

**Theorem 5.1.** *For $q = p^r$, let $\mathcal{P}$ be a set of $m$ points and $\mathcal{L}$ be a set of $n$ lines, both in $\mathbb{F}_q^2$, such that $m^{7/8} < n < m^{8/7}$ and $n^{13}/m^2 = O(p^{15})$. Then*

$$I(\mathcal{P}, \mathcal{L}) = O(m^{11/15} n^{11/15}).$$

We can use the technique from Section 2 in $\mathbb{F}_q$.

**Exercise 11.** *Adapt the proof of Theorem 2.2 to $\mathbb{F}_q$, by using Theorem 5.1. What sum–product bound did you get? What are the restrictions?*

**Exercise 12.** *Repeat Exercise 4 over $\mathbb{F}_q$. What bounds did you get?*

The above point-line incidence bound in $\mathbb{F}_q^2$ actually has a better result in the case where the point set is a Cartesian product [20].

**Theorem 5.2.** *For $q = p^r$, let $A, B \subseteq \mathbb{F}_q$ be sets with $|A| \leq |B|$ and let $\mathcal{P} = A \times B$ be a set of points in $\mathbb{F}_q^2$. Let $\mathcal{L}$ be a set of $n$ lines in $\mathbb{F}_q^2$, such that $|A||B|^2 \leq n^3$ and $|A|n \ll p^2$ Then*

$$I(\mathcal{P}, \mathcal{L}) = O(|A|^{3/4}|B|^{1/2}n^{3/4} + n).$$

**Exercise 13.** *Repeat Exercise 11 over $\mathbb{F}_q$ using Theorem 5.2. What bounds did you get? How do the p-constraints compare?*

For more expanding polynomials over $\mathbb{F}_q$, see Section 7. For further reading, see [9].

# 6 Problem 5: Local properties

Let $A$ be a set of $n$ real numbers. The *difference set* of $A$ is defined as

$$A - A = \{a - a' \ : \ a, a' \in A \quad \text{and} \quad a > a'\}.$$

Note that we only consider positive differences. This is not a standard definition. However, it makes the following problem simpler and more intuitive.

Let $k$ and $\ell$ be fixed numbers and let $n$ be asymptotically large. We consider sets $A$ of $n$ real numbers, such that every subset $A' \subset A$ of size $k$ satisfies that $|A' - A'| \geq \ell$. Examples:

- When $k = \ell = 3$, we consider the sets that do not contain a 3-term arithmetic progression. Three numbers span three distinct differences, unless they form a 3-term arithmetic progression.
- When $k = 4$ and $\ell = 6$, we consider sets where no difference repeats twice. When a difference repeats twice, we can use the numbers that form these representations to obtain four numbers that span fewer than six differences.
- Consider the case of $k = 4$ and $\ell = 5$. Assume that there exist distinct $a_1, a_2, a_3, a_4 \in A$ such that $a_1 - a_2 = a_3 - a_4 > 0$. Then we also have that $|a_1 - a_3| = |a_2 - a_4|$. This is not allowed, since $a_1, a_2, a_3, a_4$ span at most four differences. However, 3-term arithmetic progressions are allowed in this case.

Let $g(n, k, \ell)$ be the minimum size of $|A - A|$, taken over all sets $A$ of $n$ real numbers that satisfy the above condition. That is, every $A' \subset A$ of size $k$ satisfies $|A' - A'| \geq \ell$. This problem was originally studied by Erdős and Sós.

Recall that the case of $k = \ell = 3$ consists of the sets $A$ that do not contain a 3-term arithmetic progression. *Behrend's construction* is a set of $n$ integers from $\{1, 2, 3, \ldots, n \cdot 2^{c\cdot\sqrt{\log n}}\}$ with no 3-term arithmetic progression (for some constant $c$). By definition, this set of numbers forms fewer than $n \cdot 2^{c\cdot\sqrt{\log n}}$ positive differences. Thus, we have that $g(n, 3, 3) \leq n \cdot 2^{c\cdot\sqrt{\log n}}$.

**Claim 6.1.** *For every $k \geq 4$, we have that*

$$g\left(n, k, \binom{k}{2} - \lceil k/2 \rceil + 2\right) = \Theta(n^2).$$

*Proof.* For simplicity, we assume that $k$ is even. It is not difficult to extend the proof to the case where $k$ is odd. Let $A$ be a set of $n$ real numbers that satisfies the local property. That is, every $k$ numbers from $A$ span at least $\binom{k}{2} - k/2 + 2$ differences.

Assume for contradiction that a difference $\delta \in \mathbb{R}$ has at least $k/2$ representation in $A$. That is, there exist $a_1, \ldots, a_k \in A$ such that

$$a_1 - a_2 = a_3 - a_4 = \ldots = a_{k-1} - a_k = \delta.$$

We set $S = \{a_1, \ldots, a_k\}$. If some of the elements $a_j$ are identical, then $|S| < k$. We arbitrarily add more elements from $A$ to $S$, until $|S| = k$. There are $\binom{k}{2}$ pairs of elements from $S$, and at least $k/2$ of these pairs form the difference $\delta$. Thus, $|S - S| \leq \binom{k}{2} - k/2 + 1$, which contradicts the assumption about $A$.

The above contradiction implies that every difference has at most $k/2 - 1$ representations in $A$. Since there are $\binom{n}{2}$ pairs of elements from $A$, and at most $k/2 - 1$ correspond to each difference, we conclude that

$$|A - A| \geq \binom{n}{2}/(k/2 - 1) = \Theta(n^2).$$

□

In the above problems, we assume a local property about small parts of $A$, and conclude a global property about $|A - A|$. For that reason, these are called *local properties problems*. We can also think about these as expanding polynomial problems, with a restriction on the set $A$. We show that $f(x, y) = |x - y|$ expands for sets of numbers with local properties.

As far as we know, the above local properties problems were only studied with respect to the $f(x, y) = |x - y|$. We suggest initiating the study of the expansion of other functions, under the above local properties. Or under different local properties. For example, what local properties affect the expansion of $f(x, y, z) = x + yz$ ?

We can also continue the study of the family of problems involving $g(n, k, \ell)$. Since these have been studied before, it would be more difficult to obtain new results. But that also means that new results would also be more impressive.

For further reading, see [4, 5, 10].

**Local properties in $\mathbb{R}^2$.** Instead of sets of $n$ numbers, we now consider sets of $n$ points in the plane. Instead of the difference of two numbers, we now consider the distance between two points. After these changes, the local property states that every $k$ points span at least $\ell$ distinct distances. Let $\phi(n, k, \ell)$ denote the minimum number of distinct distances determined by $n$ points that satisfy this local property.

Some local properties problems become more difficult when considered in $\mathbb{R}^2$. For example, while it is not difficult to show that $g(n, 4, 5) = \Theta(n^2)$, very little is known about $\phi(n, 4, 5)$.

**Exercise 14.** *Prove that $g(n, 4, 5) = \Theta(n^2)$.*

**Exercise 15.** *Prove that $\phi(n, 4, 5) = \Omega(n)$ and that $\phi(n, 4, 5) = O(n^2)$. (These are the current best bounds for $\phi(n, 4, 5)$.)*

Erdős asked whether $\phi(n, 4, 5) = \Theta(n^2)$. Obtaining new bonds for $\phi(n, 4, 5)$ is probably too difficult for a summer project. Instead, we suggest a new and hopefully simpler variant of the problem.

A set of points in the plane lies in *convex position* if they are the vertices of a convex polygon. (For a formal definition see `https://en.wikipedia.org/wiki/Convex_position`.) Can we prove stronger bounds for $\phi(n, 4, 5)$ when only considering point sets that lie in convex position?

Let $\phi_{conv}(n, k, \ell)$ denote the minimum number of distinct distances determined by $n$ points in convex position that satisfy the above local property. What can we say about $\phi_{conv}(n, k, \ell)$?

**Exercise 16.** *Prove the following bounds.*

- $\phi_{conv}(n, k, 1) \leq \lfloor n/2 \rfloor$
- $\phi_{conv}(n, k, 1) \geq \lfloor (n-1)/3 \rfloor$. *This part is challenging, so the following hints may be useful.*[2]
  - *Let $P$ be a set of $n$ points in convex position. Let $T = \{(a, p, q) \in P^3 : |ap| = |aq|$ and $a, p, q$ distinct$\}$. Our proof strategy will be to bound $|T|$ from above and from below.*
  - *For any $p, q \in P$, at most how many $a$ exist with $(a, p, q) \in T$? Use this to bound $|T|$ from above.*
  - *For each $p \in P$, let $x_p$ be the number of distinct distances between $p$ and all other $q \in P \setminus \{p\}$. Let $x = \max_{p \in P} x_p$. Let $C_{p,1}, \ldots, C_{p,x}$ be $x$ concentric circles around $p$ containing all points in $P \setminus \{p\}$. Set $n_{p,i} = |C_{p,i} \cap P|$.*

[2] This proof follows Lemma 3.1 from `https://arxiv.org/pdf/1406.1949.pdf`.

– *Express $|T|$ in terms of the $n_{p,i}$. Then, bound this expression from below using the Cauchy-Schwarz inequality.*

– *Combine the lower and upper bounds for $|T|$. What can we say about $x$? How does this apply to $\phi_{conv}(n,k,1)$?*

When $\ell = 1$, we are imposing no local property restriction. In other words, $\phi_{conv}(n,k,1)$ is the minimum number of distinct distances determined by $n$ points in convex position.

**Research directions.** Once you have read this section (and the Energy section), you could go in the following directions.

- Find a function $f$, so that when a set $A$ satisfies a certain local property (you can pick this), $f(A)$ expands in an interesting way. This is a very open direction—you can chat with mentors for more specific ideas.

  – One idea: fix the polynomial $f(x,y,z) = x + yz$. For all $A$ we know $f(A) = \Omega(n^{3/2})$ (Exercise 4). Can we get a better lower bound by imposing a local property on $A$? (By Exercise 5, our lower bound cannot be better than $\Omega(n^2)$.)

- Find lower or upper bounds for $\phi_{conv}(n,k,\ell)$. To better understand the convex position condition, it may be helpful to read [21, 22]. To better understand the local properties problem, you can start with [23]. Or, just draw points on a piece of paper and play around with the problem to get a feel for it!

# 7 Energy: a useful tool

This section is not about one of the problems of the project. Instead, it introduces an important tool, which could be used in many expander problems.

Let $A$ be a set of $n$ real numbers. The *additive energy* of $A$ is

$$E^+(A) = \left|\left\{(a_1,a_2,a_3,a_4) \in A^4 \ : \ a_1 + a_2 = a_3 + a_4\right\}\right|.$$

In other words, $E^+(A)$ is the number of solutions from $A$ to the equation $a_1 + a_2 = a_3 + a_4$. There are $n^2$ trivial solutions where $a_1 = a_3$ and $a_2 = a_4$. This implies that $E^+(A) \geq n^2$. After fixing the values of $a_1, a_2$, and $a_3$, there is at most one valid choice for $a_4$. This implies that $E^+(A) \leq n^3$.

For $x \in A + A$, we set

$$r^+(x) = \left|\left\{(a_1,a_2) \in A^2 \ : \ a_1 + a_2 = x\right\}\right|.$$

In other words, $r^+(x)$ is the number of ways in which $x$ can be represented as a sum of two elements of $A$. Every pair of $A^2$ contributes to exactly one $r^+(x)$, so $\sum_{x \in A+A} r^+(x) = |A|^2$.

The number of quadruples $(a_1, a_2, a_3, a_4) \in A^4$ that satisfy $a_1 + a_2 = a_3 + a_4 = x$ is $r^+(x)^2$. This implies that $E^+(A) = \sum_{x \in A+A} r^+(x)^2$. By the Cauchy-Schwarz inequality, we have that

$$\sum_{x \in A+A} r^+(x) \leq \left( \sum_{x \in A+A} r^+(x)^2 \right)^{1/2} \left( \sum_{x \in A+A}^{n} 1^2 \right)^{1/2} = \left( \sum_{x \in A+A} r^+(x)^2 \right)^{1/2} |A+A|^{1/2}.$$

Combining the above leads to

$$E^+(A) = \sum_{x \in A+A} r^+(x)^2 \geq \frac{\left(\sum_x r^+(x)\right)^2}{|A+A|} = \frac{|A|^4}{|A+A|}. \tag{5}$$

From (5), we get that a small sum set implies a large additive energy. For example, a minimal sum set $|A+A| = \Theta(n)$ implies a maximal energy $E^+(A) = \Theta(n^3)$. When seeing this, one may conjecture that there is a correlation between $|A+A|$ and $E^+(A)$. A few other examples that support this conjecture:

- If $A$ is an arithmetic progression then $|A+A| = 2n-1$ and $E^+(A) = \Theta(n^3)$.
- If $A$ is a random set, then we expect $|A+A| = \Theta(n^2)$ and $E^+(A) = \Theta(n^2)$. Indeed, the probability that two pairs of random numbers have the same sum is almost zero.
- Consider $0 < \alpha < 1$. Let $A = H + R$ where $H$ is an arithmetic progression of size $\Theta(n^\alpha)$ and $R$ is a random set of size $\Theta(n^{1-\alpha})$. Then $|A+A| \approx n^{2-\alpha}$ and $E^+(A) \approx n^{2+\alpha}$.

While $|A+A|$ and $E^+(A)$ are related, their relationship is not as simple as in the above examples. Consider $A = H \cup R$, where $P$ is an arithmetic progression of size $n/2$ and $R$ is a random set of size $n/2$. The random elements lead to $|A+A| = \Theta(n^2)$. The elements of the arithmetic progression lead to $E^+(A) = \Theta(n^3)$.

The *Balog-Szemerédi-Gowers theorem* states how $E^+(A)$ provides information about $|A+A|$. This theorem has many different variants. The following variant is by Schoen [17].[3]

**Theorem 7.1.** *Let $A$ be a set of $n$ real numbers that satisfies $E^+(A) = \delta \cdot n^3$. Then there exists $A' \subset A$ such that $|A'| = \Omega(\delta \cdot n)$ and*

$$|A' + A'| = O\left(\delta^{-8} \cdot n\right).$$

Intuitively, Theorem 7.1 states that large $E^+(A)$ implies small $|A+A|$, possibly after removing some noise. In the above example where $A = H \cup R$, the noise is the random set $R$.

[3]More precisely, it is a straightforward combination of a result of Schoen with Plünnecke's inequality.

Additive energy and its variants are main objects of additive combinatorics, and have many uses. For example, many local properties results are obtained by relying on energies (see Section 6).

**Energies in $\mathbb{F}_q$.** We now see how to derive more results in $\mathbb{F}_q$ by relying on energies. We also require the following point–plane incidence bound of Rudnev [13].

**Theorem 7.2.** *Let $q = p^r$ for some prime $p$ and positive integer $r$. Let $\mathcal{P}$ be a set of $m$ points and let $\Pi$ be a set of $n$ planes, both in $\mathbb{F}_q^3$, such that $n \geq m$. Assume that $m = O(p^2)$ and that no line of $\mathbb{F}_q^3$ contains $k$ points of $\mathcal{P}$. Then*

$$I(\mathcal{P}, \Pi) = O(n\sqrt{m} + kn).$$

Exercise 11 asked to prove a sum-product bound in $\mathbb{F}_q$ by replying on Theorem 7.2. We now prove a similar bound by using Theorem 7.2. This proof is an example of how we can use energies in proofs. It is taken from [12].

**Theorem 7.3.** *Let $q = p^r$ and let $A \subset \mathbb{F}_q$ satisfy $|A| \leq p^{5/8}$. Then*

$$\max\{|A+A|, |AA|\} = \Omega\left(n^{6/5}\right).$$

*Proof.* If $0 \in A$, then we remove this element from $A$. This does not change the asymptotic size of $A$. We can then set $A^{-1} = \{1/a : a \in A\}$.

Consider $a_1, a_2, a_4, a_5 \in A$ that satisfy $a_1 + a_2 = a_4 + a_5$. Each of the $|A|^2$ pairs $(a_3, a_6) \in A^2$ satisfy $a_1 + a_2a_3/a_3 = a_4 + a_5a_6/a_6$. This implies that

$$\begin{aligned} E^+(A) &= |A|^{-2} \left|\left\{(a_1, \ldots, a_6) \in A^6 \ : \ a_1 + a_2a_3/a_3 = a_4 + a_5a_6/a_6\right\}\right| \\ &\leq |A|^{-2} \left|\left\{(a_1, b_1, c_1, a_2, b_2, c_2) \in \left(A \times AA \times A^{-1}\right)^2 : \ a_1 + b_1c_1 = a_2 + b_2c_2\right\}\right|. \quad (6) \end{aligned}$$

With the above in mind, we define the energy variant

$$E'(A) = \left|\left\{(a_1, b_1, c_1, a_2, b_2, c_2) \in \left(A \times AA \times A^{-1}\right)^2 : \ a_1 + b_1c_1 = a_2 + b_2c_2\right\}\right|.$$

We study $E'(A)$ with a three-dimensional variant of Elekes's sum-product argument (see the proof of Theorem 2.2). Consider the point set

$$\mathcal{P} = \{(a_1, b_2, c_1) \in A \times AA \times A^{-1}\},$$

and the set of planes

$$\Pi = \{x + b_1z - c_2y = a_2 \ : \ (a_2, b_1, c_2) \in A \times AA \times A^{-1}\}.$$

A 6-tuple $(a_1, b_1, c_1, a_2, b_2, c_2) \in A^6$ contributes to $E'(A)$ if and only if the point $(a_1, b_2, c_1)$ is incident to the plane defined by $x + b_1z - c_2y = a_2$. Indeed, plugging

the coordinates of $(a_1, b_2, c_1)$ into the plane equation leads to $a_1 + b_1c_1 - b_2c_2 = a_2$. This implies that $E'(A) = I(\mathcal{P}, \Pi)$. We wish to derive an upper bound for $I(\mathcal{P}, \Pi)$ by using Theorem 7.2. We first check the conditions of this theorem hold.

We note that each triple $(a_2, b_1, c_2) \in A \times AA \times A^{-1}$ leads to a distinct plane of $\Pi$. We thus have that $|\mathcal{P}| = |\Pi| = |A|^2|AA|$. Let $m = |A|^2|AA|$. The maximum number of collinear[4] points in $\mathcal{P}$ is $|AA|$, since this is the size of the longest side of the cartesian product $A \times AA \times A^{-1}$. We may assume that $|AA| = O(|A|^{6/5})$, since otherwise we are done. Combining this with $|A| = O(p^{5/8})$ implies that $m = |A|^2|AA| = O(p^2)$. We can thus apply Theorem 7.2 on $\mathcal{P}$ and $\Pi$ with $k = |AA|$, to obtain that

$$\begin{aligned} E'(A) = I(\mathcal{P}, \Pi) = O\left(m^{3/2} + mk\right) &= O\left(|A|^3|AA|^{3/2} + |A|^2|AA|^2\right) \\ &= O\left(|A|^3|AA|^{3/2}\right). \end{aligned}$$

Combining the above with (6) leads to

$$E^+(A) \leq |A|^{-2}E'(A) = O\left(|A||AA|^{3/2}\right) = O\left(|A|^{14/5}\right).$$

The argument that led to (5) also holds over $\mathbb{F}_q$. That is, we have that

$$E^+(A) \geq \frac{|A|^4}{|A + A|}.$$

By combining the above lower and upper bounds for $E^+(A)$, we get that $|A+A| = \Omega(|A|^{6/5})$. □

Just like Elekes's argument, the proof of Theorem 7.3 is based on double counting incidences.

**Exercise 17.** *Let $q = p^r$ and let $A \subset \mathbb{F}_q$ satisfy $|A| \leq p^{2/3}$ and $0 \notin A$. Let $f(x, y, z) = (x + y)/z$. Prove that $f(A, A, A) = \Omega(n^{3/2})$.*

*Instructions: Define a new type of energy. Derive a lower bound for this energy by imitating the lower bound for $E^+(A)$ from* (5)*. Derive an upper bound for the energy by reducing the problem to a point–plane incidence problem, as in the proof of Theorem 7.3. In this case, the point set can be $A^3$.*

# References


[1] G. Elekes, On the number of sums and products, *Acta Arith.* **81** (1997), 365–367.

[2] P. Erdős, On sets of distances of $n$ points, *Amer. Math. Monthly* **53** (1946), 248–250.


[4] A set of points is said to be *collinear* if there is a line that contains all the points.

[3] P. Erdős and E. Szemerédi, On sums and products of integers, *Studies in pure mathematics, To the memory of Paul Turán,* 213–218, Birkhäuser, Basel-Boston, Mass., 1983.

[4] S. Fish, B. Lund, and A. Sheffer, A Construction for Difference Sets with Local Properties, *Eur. J. Combin.* **79** (2019), 237–243.

[5] S. Fish, C. Pohoata, and A. Sheffer, Local Properties via Color Energy Graphs and Forbidden Configurations, *SIAM J. Discrete Math.* **34** (2020), 177–187.

[6] L. Guth and N.H. Katz, On the Erdős distinct distances problem in the plane, *Annals Math.* **181** (2015), 155–190.

[7] S. Mathialagan and A. Sheffer, Distinct distances on non-ruled surfaces and between circles, arXiv:2011.08098.

[8] M. Mirzaei and A. Suk, On grids in point-line arrangements in the plane *Discrete & Computational Geometry* **65** (2021), 1232–1243.

[9] T. Pham, L. A. Vinh, and F. de Zeeuw, Three-variable expanding polynomials and higher-dimensional distinct distances, *Combinatorica* **39** (2019), 411–426.

[10] C. Pohoata and A. Sheffer, Local Properties in Colored Graphs, Distinct Distances, and Difference Sets, *Combinatorica* **39** (2019), 705–714.

[11] O. E. Raz, M. Sharir, and J. Solymosi, Polynomials vanishing on grids: The Elekes–Rónyai problem revisited, *Amer. J. Math.* **138** (2016), 1029–1065.

[12] O. Roche-Newton, M. Rudnev, and I. D. Shkredov, New sum-product type estimates over finite fields, *Advances in Mathematics* **293** (2016), 589–605.

[13] M. Rudnev, On the number of incidences between planes and points in three dimensions, *Combinatorica* **38** (2018): 219–254.

[14] M. Rudnev and I. D. Shkredov, On growth rate in $SL_2(\mathbf{F}_p)$, the affine group and sum-product type implications, arXiv:1812.01671.

[15] M. Rudnev and S. Stevens, An update on the sum-product problem, arXiv:2005.11145.

[16] M. Rudnev and J. Wheeler, Incidence bounds with Möbius hyperbolae in positive characteristic, arXiv:2104.10534.

[17] T. Schoen, New bounds in Balog-Szemerédi-Gowers theorem, *Combinatorica* **35** (2015), 695–701.

[18] M. Sharir, A. Sheffer, and J. Solymosi, Distinct distances on two lines, *J. Combinat. Theory Ser. A*, **120** (2013), 1732–1736.

[19] S. Stevens and A. Warren, On sum sets of convex functions, arXiv:2102.05446.

[20] S. Stevens and F. De Zeeuw, An improved point–line incidence bound over arbitrary fields, *Bulletin of the London Mathematical Society*, **49** (2017), 842–858.

[21] Nivasch, G., Pach, J., Pinchasi, R., Zerbib, S., 2013. The number of distinct distances from a vertex of a convex polygon. arXiv:1207.1266 [cs].

[22] Dumitrescu, A., 2004. On distinct distances from a vertex of a convex polygon, in: Proceedings of the Twentieth Annual Symposium on Computational Geometry - SCG '04. Presented at the the twentieth annual symposium, ACM Press, Brooklyn, New York, USA, p. 57. https://doi.org/10.1145/997817.997829

[23] Sheffer, A., 2018. Distinct Distances: Open Problems and Current Bounds. arXiv:1406.1949 [cs, math].

# A Asymptotic notation

Intuitively, saying that some quantity is $O(n)$ means that the quantity is at most some constant times $n$. For example, we may write $10n = O(n)$ since $10n$ is at most some constant times $n$. We may also write $10n = O(n^2)$ since $10n$ is at most some constant times $n^2$.

In asymptotic notation, we only care about very large values of $n$. We may write $10^{10} + n = O(n)$, since $10^{10} + n$ is smaller than $2n$ when $n$ is sufficiently large. It does not matter that when $n$ is small we have $10^{10} + n > 2n$. To recap: $f(n) = O(g(n))$ means that, for any sufficiently large $n$, the value of $f(n)$ is at most some constant times $g(n)$. In some sense, $f(n)$ is *upper bounded* by $g(n)$.

We write $f(n) = \Omega(g(n))$ when, for any sufficiently large $n$, the value of $f(n)$ is *at least* some constant times the value of $g(n)$. In some sense, $f(n)$ is *lower bounded* by $g(n)$. For example, $n^5/100 = \Omega(n^4)$. Note that $f(n) = \Omega(g(n))$ is equivalent to $g(n) = O(f(n))$

Finally, $f(n) = \Theta(g(n))$ implies that both $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$ hold. That is, $f(n) = \Theta(g(n))$ means that, for any sufficiently large $n$, the value of $f(n)$ is *about the same* as $g(n)$ (up to a constant).

If you are interested in a rigorous definition: $f(n) = O(g(n))$ implies that there exist constants $c, n_0$, such that for any $n \geq n_0$, we have $f(n) \leq c \cdot g(n)$. For example, $10n^2 + 1000 = O(n^2)$ holds since we can take $c = 100$ and $n_0 = 20$.

**Exercise 18.** *Find a function $f(n)$ which satisfies the following two properties:*

- $f(n) = \Omega(n^k)$ *for all* $k > 0$, *and*

- $f(n) = O(k^n)$ *for all* $k > 1$.

**Exercise 19.** *Let* $f(n)$ *and* $g(n)$ *be any two functions. Must we necessarily have* $f(n) = O(g(n))$ *or* $g(n) = O(f(n))$*? Prove this or find a counterexample.*